Urban Energy System Design from the Heat Perspective using mathematical Programming including thermal Storage

THÈSE N^{O} 6731 (2015)

PRÉSENTÉE LE 28 AOÛT 2015

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE D'ÉNERGÉTIQUE INDUSTRIELLE
PROGRAMME DOCTORAL EN ÉNERGIE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Jakob Moritz Fabian RAGER

acceptée sur proposition du jury:

Dr J. Van Herle, président du jury
Prof. F. Maréchal, directeur de thèse
Prof. A. Bardow, rapporteur
Prof. N. Shah, rapporteur
Prof. D. Favrat, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

It's not that I'm so smart,
it's just that I stay with my problems longer.
— Albert Einstein

To Véronique and Arthur,
Marie-Luise, Günther and Bastian ...

Acknowledgements

During the years spend at EPFL, many people influenced this work. The here presented work could not have done without the previous work.

I would like to thank the professors Daniel Favrat and François Maréchal for the fruitful discussions and input during this thesis. François, thank you for directing this thesis and your support during my years at the LENI and IPESE.

Also would like to thank my jury members André Bardow, Nilay Shah and the president of the jury, Jan Van Herle: You have helped me a lot to improve this work! Thank you for your comments and the discussion around it.

During this thesis, numerous collaborations lead to the here presented work. First, I'd like to thank the CREM for the collaboration throughout my thesis. Thank you Gaëtan, Gabriel, Loïc and Fabien for your feedback, ideas, patience and for reminding me of what is needed outside of academia. A big thanks also to the MEU project and its members for the years of active debate to reach a working web platform for energy management of cities. Massimiliano, thank you for heading this project for the Energy Center and enriching it with all sorts of play on words.

I would like to thank the Geneva Canton's Energy Office (Office cantonal de l'énergie) for getting this thesis started: Rémy Beck, your ability to collect the practical problems and form them into a research question helped me understand the challenges; Oliver Epelly for giving me the possibility of an internship back in the days where my french was not more than a collection of words; Guillaume Ferraris for the day to day work with me.

The collaboration with CERN introduced me to the world of international organizations, which works quiet differently than the local authorities.

I would also like to thank all my former students for the work they have done, with a special thanks to the following master project students: Ben, Mathias, Mathieu, Cédric, Sylvain. To my civil service support team Stéphane, Alex, Basile and Frédéric: Thank you for teaching me how to program. Ok, the road is still long, but you managed to show me why it is necessary and how it can be done. Working with you was a pleasure. A big thanks also for to the numerous secretaries which helped to organize all the different requests: Brigitte, Susanne, Sophie and Dagmara.

Nicolas Borboën for the technical support as well as the Scitas Team (especially the other Nicolas and Jean-Claude). Malik for the introduction to the great and endless world of python and linux. A big thanks to the open-source community for providing, maintaining and updating the following software:

- Xubuntu, a light weight linux distribution,
- the programming language Python in version 2.7 (mostly used the packages SciPy, NumPy, pandas, matplotlib, Seaborn, SALib, rpy2 and scikit-learn) ,
- the statistical computing environment R in version 3.2 (clustering package),
- the LaTeX-project for the typesetting,
- and the bibliography manager zotero.

To all my former office mates Nasibeh, Germain, Emanuele, Samuel: Thank you for supporting me and all the great time we had together sharing the same 4 walls. To my colleagues Matthias D., Stefano, Stéphane and Nils who have read part of this work: thanks for your courage going through my work and providing constructive feedback. And also tThe rest of the LENI/IPESE and Energy Center team for the great ambiance especially during shared pauses that were more than a great distraction.
My thanks to the outside world: With the my rowing team of the Lausanne Sports Aviron and the UNIL/EPFL rowing club, we had great time during the numerous trips, regattas or workouts on the lake during the morning hours. Pierre and Patrick for all the stupid ideas like "why not do La Patrouille des Glaciers?" or the few successful trials of just having "one" beer together!
To the two families, the Martis and the Ragers, who supported me during this thesis. Last but not least to my own family, Véronique and Arthur, who was born during the last days of this thesis. Surtout toi, Véronique chapeau ! Merci d'être bien plus que ma partenaire !

Lausanne, 31 July 2015 Jakob

Abstract

Energy planning recently received more attention in Switzerland through the new strategy phasing out nuclear energy by 2034. Often however the energy planning is only done from the electrical side. This work takes a different angel and helps communities and energy utilities planning tomorrows energy system from a heat based perspective.

After the data collection and structuring, the methodology presented here designs an energy system. Based on the quality of the collected data, the approach to define the energy demand should be chosen. In order to reduce calculation time, a data reduction approach is developed to reduce the input data without loosing significant information and precision.

In particular, the methodology focuses on the integration of a stochastic resource, in this case solar thermal heat production, in combination with thermal energy storage. The thermal energy storage can be used as a short or long term thermal energy storage. The framework compares design solutions for the two storage types considering either a total cost approach or a life cycle assessment approach using the cumulative exergy demand (CExD).

The proposed mathematical programming framework is based on a mixed integer linear programming (MILP) approach, that can work on different levels of detail between building to community or city level. The optimization problem can also be further simplified to a linear problem, increasing the size of problem that can be solved while reducing or keeping a constant computation time. The discussed cases show an interest in further investigating storage solution using both, the short and long term storage at once, because they allow to reduce the system's overall costs or CExD significantly.

The framework is then extended to consider buildings as an energy storage. The building's internal temperature can be raised from 20 °C up to 23 °C, giving a comfort temperature range that can be used for storing heat. The integrating of both storage types, the thermal energy storage and the building as an energy storage, show no significant impact on the energy system design. However, costs or CExD can be reduced. In addition, the heat demand can be modified through the decision of optimal energy retrofitting strategy for a group of buildings. The framework decides which of the building to refurbish based the overall CExD including CExD used for retrofitting the building.

Finally, a method is proposed to integrate uncertainty of the model's input parameters into the system design. A global sensitivity analysis evaluates the impact of each uncertain parameter onto the system, allowing to focusing on the outputs of interest. Robust optimization is applied with a simulation-based approach, the additional costs for a robust design are calculated, as well as the different unit sizes.

The low complexity of the developed models allows for an easy integration of new data collected during the development of a project, which is often the case in urban energy planning applications.

Key words: k-medoids clustering, mixed integer linear programming, thermal energy storage, building energy storage, solar thermal energy integration, global sensitivity analysis, robust optimization

Zusammenfassung

Durch die Energiewende und die dadurch eingeleitete Abschaltung der Kernkraftwerke bis 2034 rückt die Energieplanung auch in der Schweiz wieder in den Vordergrund. Allerdings wird dabei oft nur über die elektrische Seite der Energieplanung geredet. Die hier vorliegende Arbeit wählt einen anderen Ansatz: Sie hilft Energieversorger und Gemeinden, die Energiesysteme von morgen aus der Wärmeperspektive zu planen.

Nach dem Datensammeln und Strukturieren, wird eine Methode zum Energiesystemdesign zu vorgestellt. Die Datenqualität beeinflusst die Wahl der Ansatzes zur Bestimmung der Energienachfrage. Damit das Optimierungsproblem kleiner wird, aber die gleiche Lösung hat, werden Clustering Methoden verwendet um die Eingangsdaten auf ein Minimum zu reduzieren ohne wichtige Details zu verlieren.

Ins besondere der thermische Speicher und stochastischen Energiequellen wie die Solarthermie werden in das Energiesystem integriert. Es gibt einen kurz- und einen Langzeit thermischen Energiespeicher. Das Optimierungsproblem vergleicht verschiedene Designs ohne, mit einem der beiden oder beiden Speichern unter einem Gesamtkostengesichtspunkt oder einem Lebenszyklusanalyse/Ökobilanz basierten auf der kumulierten Exergienachfrage.

Das vorgeschlagene Framework löst das Designproblem mit einem gemischt ganzzahligen linearen Optimierungsansatz für kleine Nachbarschaften bis auf Gemeindeebene oder Städteebene. Das Optimierungsproblem kann noch weiter zu einem linearen Optimierungsproblem vereinfacht werden, wenn Parameter angepasst werden. Dadurch können größere Probleme gelöst werden, während die Zeit zur Lösung konstant bleibt oder reduziert werden kann. Die berechneten Fälle zeigen, dass sich gerade mit der Integration von lang und Kurzzeit Speichern gleichzeitig, Kosten und CExD sparen lassen.

Das Optimierungsproblem wird dann um einen Gebäudewärmespeicher erweitert. Dieser kann die Innentemperatur in der Komfortzone von 20 auf bis zu 23 °C heben, um zusätzliche Energie in der Gebäudestruktur zu speichern. Durch das Gleichzeitige Benützen von dem thermischen Gebäudespeicher und dem Kurzzeit Speicher können wieder Kosten eingespart bzw. die CExD reduziert werden ohne das Design des Energiesystem zu verändern. Wieder können Kosten eingespart oder CExD reduziert werden. Des Weiteren kann das Framework über energetische Renovierung entscheiden: Aus einer Gebäudegruppe wird die Optimale Menge an Gebäuden ausgewählt um die Systemkosten bzw. CExD zu reduzieren.

Im letzten Kapitel wird die Auswirkung von Unsicherheit im Energiesystemdesign untersucht. Mit einer globalen Sensitivitätsanalyse werden zwischen wichtigen und unwichtigen Parametern entschieden, damit nur die wichtigen weiterhin betrachtet werden. Danach werden

die Kosten und Auswirkungen einer robusten Optimierung auf das Energiesystem bestimmt. Dafür werden Schrittweise mehr Parameter auf den schlimmsten Wert gesetzt in einem Simulationsbasierten Ansatz. Die letztlich niedrige Komplexität der Modelle erlaubt nicht nur die Sensitivitätsanalyse sondern auch neue Daten bzw. einen Kenntnisstand in das Projekt zu integrieren. Gerade in urbanen Projekten kommt die häufig vor.

Stichwörter: k-medoids clustering, gemischte ganzzahlige Optimierung, thermischer Energiespeicher, Gebäudeenergiespeicher, Integration von Solarthermie, globale Sensitivitätsanalyse, robuste Optimierung

Short Contents

Contents

List of Figures

List of Tables

Nomenclature

Latin letters

ΔT_{min}	Minimum temperature difference [K]
$\dot{M}$	Mass flow [kg/s]
$\dot{E}$	Technical mechanical power [kW]
$\dot{e}$	Specific electric power [kW/kg]
$\dot{E}_q$	Heat exergy [kW]
$\dot{e}_q$	Specific heat exergy [kW/kg]
$\dot{E}_y$	Transformation exergy [kW]
$\dot{e}_y$	Specific transformation exergy [kW/kg]
$\dot{E}_{el,u}$	Consumed ($^+$) / produced ($^-$) nominal electricity by unit u [kW]
$\dot{E}_{f,u}$	Consumed ($^+$) nominal fuel by unit u [kW]
$\dot{L}$	Exergy losses [kW]
$\dot{Q}$	Heat power [kW]
$\dot{q}$	Specific heat power [kW/kg]
ω_p	weight of the period p [-]
A	Area [m^2]
c_p	Specific heat capacity [kJ/kgK]
c_{el}	Electricity price for import ($^+$) or export ($^-$) [CHF/kWh$_{el}$]
c_f	Fuel price [CHF/kWh]
$cexd_{f,u}$	Specific operating CExD [kWh-Eq./kWh]
$d_{p,t}$	Length of time slice t in period p [h]

f Multiplication factor [-]

f_A Annualization factor [-]

i Interest rate for investment [-]

$invC$ Specific investment costs [CHF]

$invCExD$ Specific construction and production CExD [kWh-Eq./kW]

k_{hl} Heat loss coefficient for storage tanks [kW/m^2K]

M Mass [kg]

n Expected life time of installations [year]

nf Number of different fuels [-]

nk Number of temperature intervals [-]

np Number of periods [-]

ns Number of streams [-]

nt Number of time slices [-]

nu Number of units [-]

nv Number of temperature levels of storage system [-]

Q Heat (energy) [kWh]

Ref_b CExD value for the retrofitting of building b [kWh-Eq.]

s Installation factor [-]

T Temperature [K]

T_a Ambient temperature [K]

U Global heat transfer coefficient [kW/(m^{2}K)]

$\dot{\boldsymbol{E}}_{\boldsymbol{el}}$ Total electricity demand ($^+$) or excess ($^-$) [kW]

$\dot{\boldsymbol{R}}_{\boldsymbol{k}}$ Cascaded heat to lower interval k [kW]

$\boldsymbol{f_u}$ Multiplication factor of unit u [-]

$\boldsymbol{M_{v,p,t}}$ Water content of storage tank v after period p and time step t [kg]

$\boldsymbol{M_{v,p=0,t=0}}$ Initial water content of storage tank v [kg]

$\boldsymbol{y_u}$ Integer variable representing the existence (1) or not (0) of unit u [-]

Greek letters

η_{COP}	Carnot / Exergy efficiency [-]
ρ	Mass density [kg/m^3]
Θ	Carnot Factor [-]
ε_{el}	Electrical efficiency [-]
ε_{th}	Thermal efficiency [-]

Indices

a	Heat source at ambient temperature
c	Cold stream
cog	Cogeneration unit
h	Hot stream
hp	Heat pump
in	Inlet
lm	Logarithmic
out	Outlet
$sink$	Heat sink
$source$	Heat source
k	Temperature interval / Quality interval
p	Period (often a day type)
t	Time step (often one hour)
u	Index for unit
v	Temperature level of storage tank

Superscripts

$*$	Corrected temperature domain / Instantaneous loads
$+$	Entering the system
$-$	Leaving the system
max	Maximum value

min Minimum value

Acronyms

CHP Combined heat and power

COP Coefficient of performance for $(_h)$ heating or $(_c)$ cooling

DV Decision variable

LP Linear programming

MER Minimum energy requirement

MILP Mixed integer linear programming

MINLP Mixed integer non-linear programming

Conventions

bold characters Optimization variables

1 Introduction

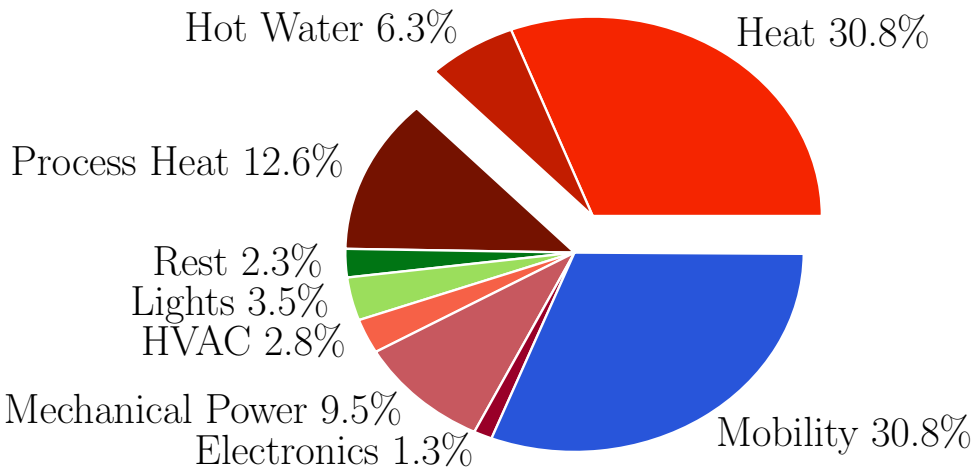
The introduction links sustainable development to urban energy planning and describes the aim of this work. The state of the art is reviewed to show shortcomings of currently existing tools and methods to derive the goals of this work.

1.1 Context: Urban Energy System Design and Sustainability

In the framework of sustainable development, the concept of the 2000-Watt society has been developed in Switzerland by the Board of the Swiss Federal Institutes of Technology [*Jochem et al.*, 2004]. The 2000-Watt society aims to reduce the consumption of the average Swiss citizen to 48 kilowatt hours a day or 2 $kW\ year/(cap\ year)$ without lowering standard of living by 2050. The 2000-Watt society picked this number because the total worldwide average rate of primary energy use equals 2000 Watts year/year in the year 2000. However, currently Switzerland is using 2.5 times more energy per person [*Kemmler et al.*, 2012], requiring a reduction to almost one third of the current consumption.

Haldi and Favrat [2006] review the challenges, finding that integrating new (renewable) energy sources and a more rational use of energy are the most important ones. This implies studying the energy or better exergy losses in the supply chain of the useful energy. The exergy losses allow to calculate the thermodynamic quality or degree of perfection of an energy system based on the Second Law of thermodynamics [see *Borel and Favrat*, 2010, p448]. Figure 1.1 shows the final energy to end-use services. Heat has the largest share. The heat share can be further sub-divided into space heating of about 30 percent and hot water production with 6 %. With around 75% of the Swiss people living in urban areas according to the *World Bank* [2015], most of useful energy demand occurs in cities. *Keirstead et al.* [2012] state that world wide two thirds of the energy consumption can be attributed to urban areas, which are responsible for up to 71 % of the global direct energy related greenhouse gas emissions.

Lowering heating requirements with measures in the built environment such as retrofitting a building or its energy conversion system can substantially impact the global emissions of the



Source: Prognos, TEP, Infras 2011

Figure 1.1 – Switzerland's final energy consumption in 2012

country. In order to find the best measure, all energy conversions including transport losses should be studied to improve the energy conversion(s) at the place with the highest impact: All energy conversions from the resource to the final energy use should be analyzed to choose the measures at the place in the energy conversion chain that have the highest impact on the overall energy conversion efficiency. Energy losses according to the official Swiss national statistics are about 50%. *Haldi and Favrat* [2006] show that the losses are significantly higher than the ones estimated, because the primary energy use is often underestimated. In addition, the loss calculation does not consider the temperature levels for the heat requirements. An accounting in terms of exergy could overcome this issue and show the usable energy or exergy correctly. When a clear target on energy use such as 2 kW is defined, it should also clearly distinguish between renewable and non-renewable primary energies. (Putting the same limit on both types of primary energy raises the question from an environmental point of view: why does the same limit apply to renewable primary energy?)

Switzerland decided to phase out its nuclear power plants in 2034. Therefore, choices have to be made for the country's electricity and more general energy strategy. Switzerland is autonomous on a yearly electricity balance, but exports electricity in the summer and imports in the winter. The new energy calculator Swiss EnergyScope [*Moret et al.*, 2014b] is addressing the Swiss energy transition on a national level. It shows that even if the nuclear power plants are replaced by renewable ways of electricity production (which is ambitious), it will not solve the problem of electricity imports in the winter. More measures need to be taken, one being the introduction of efficient energy conversion systems to avoid using electricity for direct heating. These kind plans need to include the communities and can be efficiently addressed with a bottom up approach.

Energy related indicators such as the energy use per capita represent often a basic link between sustainability and a energy consumption or emission reduction goal [*International Atomic*

Energy Agency et al., 2005] as shown with the "2000W-society". Achieving sustainable development means a "judicious use of resources, technologies, appropriate economic incentives and strategic policy planning at the local and national level" [*International Atomic Energy Agency et al.*, 2005, p.1]. The Brundtland Commission of the United Nations states on March 20, 1987 [*World Commission On Environment and Development*, 1987]:

> Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs.

A sustainable urban energy system can be defined as an energy system that uses the available resources efficiently. In addition, the system addresses the time gap between the energy demand and the availability of renewable energy reducing non-renewable energy use. The urban energy system considers its supply chain in terms of construction requirements and resource flows while delivering the required services in the most efficient manner. This can be measured with the help of an appropriate indicator such as an exergy indicator [*Favrat et al.*, 2008].

1.2 State of the Art

Energy system engineering in the field of urban energy systems is an interdisciplinary field in contact when designing and managing them. Urban energy system engineering covers different aspects:

- different simulation fields: especially building simulation, energy conversion technology simulation, energy transportation and storage technologies,
- optimization techniques in order to find feasible and optimal solutions,
- computer science in a larger sense to treat data, to store data and to interface between different tools,
- geographical information system for structuring and displaying data and
- life cycle considerations especially in the indicator design to compare solutions.

For the energy system design, finding optimal solutions means identifying the best energy conversion and distribution system, their sizes and the way they will be operated. Optimization techniques will be used to define the optimal strategy in operating and choice of equipment. When choosing indicators, it is important to propose key performance indicators to compare solutions. In addition, for the decision support methods in the field of urban energy systems, not only engineers are implicated, but a lot of other stakeholders with very different backgrounds and interests.

This literature review focuses on the use of optimization techniques in the context of the design of urban energy systems design. In this work, the term urban energy system is used to describe the build environment as a thermodynamic system. A key question in thermodynamics (and in research) is defining the system boundaries, [*Keirstead and Shah*, 2013, p.15]. Once

the boundaries are defined, a system performance can be calculated and different system configurations can be compared with each other, because it is clear which components to take into consideration. The boundaries are given through the perimeter of the urban environment studied.

1.2.1 History of Heat Demand Mapping and Planning in Switzerland

In the field of heat planning, heat mapping started already in the early 70s with *Jäckli* [1970] (see Figure 1.2). Heat planning links the planning of the heat infrastructure to the energy debate mentioned in Section 1.1. The Plenar working group [*Steiger*, 1979] came up with the proposal shown in Figure 1.3 to install Swiss wide district heating networks based on cogeneration and waste incineration plants instead of nuclear power plants in Switzerland. The working group already based their planning on recovering a maximum of heat from already existing infrastructures such as waste incinerator plants and proposed to connect areas with a high heat density demand. The Plenar work can be considered as a first nationwide integrated heat planning in Switzerland. The proposal was never realized, only cities with existing district heating networks still have a heat infrastructure today.

Similar ideas can be found in Denmark [*Kerr*, 2015], where energy planning and especially heat energy planning became a national interest starting with the First Heat Supply Law of 1979. The heat supply law introduced a tax on heat produced only through fossil fuel. As a consequence of the national heat planning, a growing share of people receive their heat through district heating: According to Euroheat, 61 % of the population are heated through district heating systems in 2011. In addition, about 600 000 m^2 of solar collectors are connected to district heating. More and more of the Danish installations include large solar thermal collector fields and between several thousand to hundred thousand of cubic meters of thermal energy storage.

In Switzerland, heat planning returned on local level after the proposal of *Köhler and Hanke* [2012] in 1979, leaving decisions up to every community or city. About 5 % of the Swiss population are connected to a district heating in 2011 (Euroheat). Neither large scale solar thermal collector fields nor large thermal energy storages have been installed. Small co-generation plants exist as well as a few heat pump driven district heating networks. Most of the heat is produced individually for each building based on combustion of non-renewable resources.

1.2.2 Stakeholder for Urban Energy System Planning and Model Classification

The stakeholder defines the criteria for the urban energy system design and realizes the design afterwards. The stakeholder during this work are the concerned communities or cities that provide one or multiple energy services via the urban energy system. Often the communities

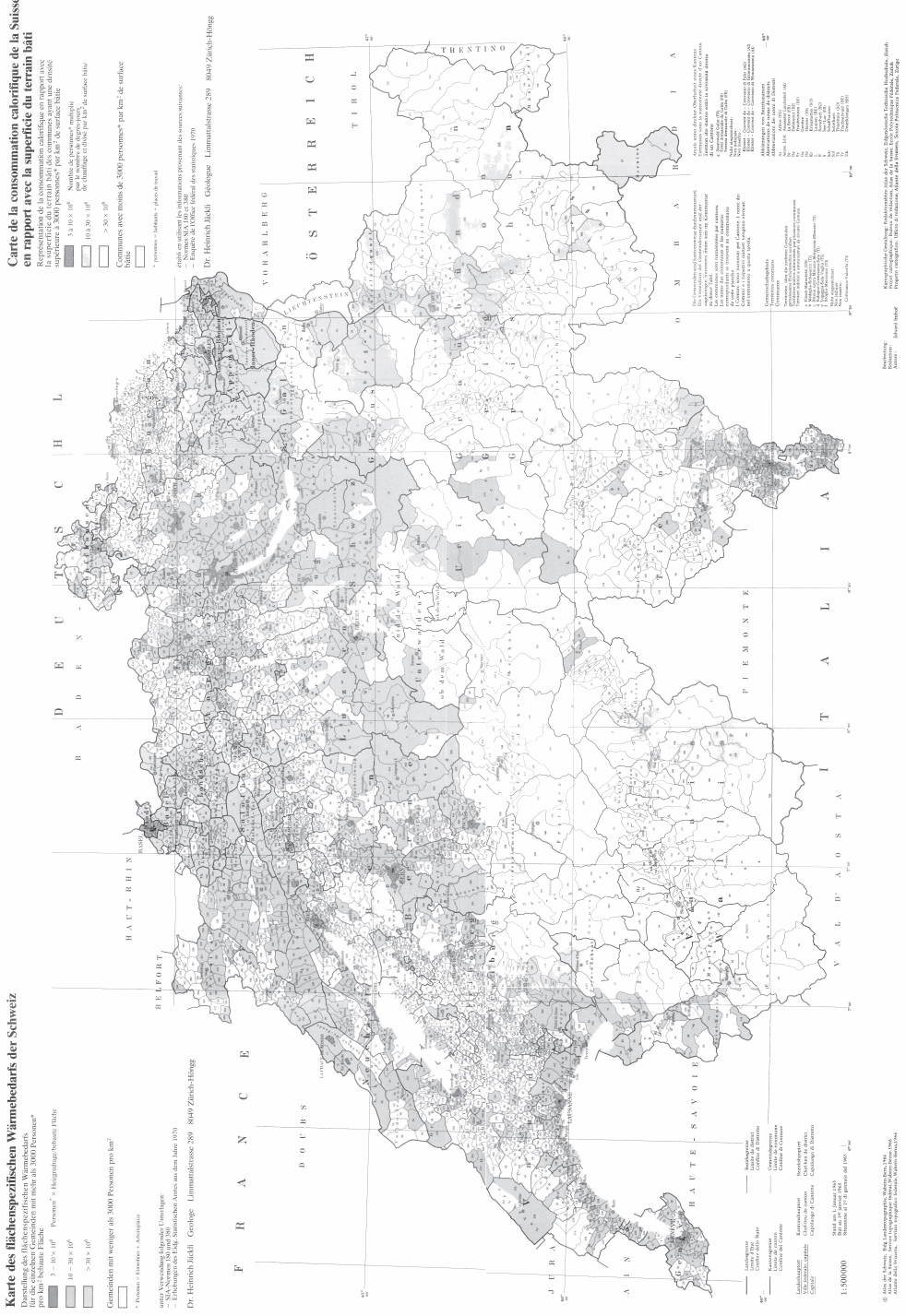


Figure 1.2 – Heat demand estimation from [*Jäckli*, 1970]

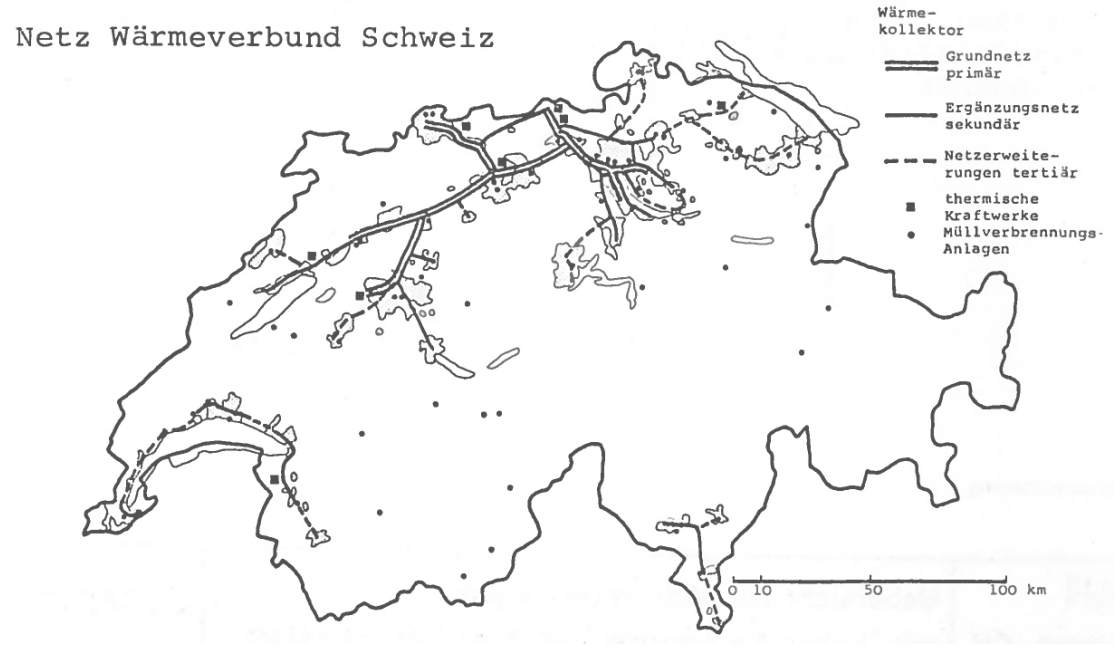


Figure 1.3 – Idea of a Swiss wide district heating networks based on cogeneration, waste heat incineration plants and waste heat recovery as a result of the Plenar working group in 1979 replacing of nuclear power plants from *Köhler and Hanke* [2012].

mandate a society that does that for them. In Switzerland, the utility companies are often still owned by the community.

The urban energy system design defines the equipment used to supply the required energy services in a given urban zone. It consists in defining the energy conversion units to be used, their size and location and the way they are interconnected one with the others using (urban) network infrastructures such as the electrical and/or gas grid, the district heating and or cooling networks. The design also includes the storage equipment used for the energy management. The design is based on different criteria such as the economic cost of the system and its operation or the environmental impact. Integration means profiting from synergies between the different units that are in the system allowing to share resources and equipment. The integration of renewable energy resources is one of the most important measures to mitigate the environmental impact. The design methodology should on the one hand consider the stochastic nature of the energy demand and on the other hand measures to integrate them with the stochastic and temporal variations of the different renewable resources like the solar or wind.

Urban energy system design and planning is a multi-scale problem. Table 1.1 shows different, quantitative levels of urban planning from the (inter) national level down to the construction project. The following actions are possible (abbreviations in parenthesis): building refurbishment (R), infrastructure construction (I), central utilities construction (CU), decentralized utilities construction (DU) and limiting emissions (E). Typically, emission control is rather

based on a top down approach, while other actions such as refurbishment are rather bottom-up approaches. The network infrastructure planning is done on all planning and spatial levels, the heat infrastructure is rather on the city to building level whereas the electricity and gas grids are planned on all levels (but often through different actors). Most of the utility planning is done on city or even smaller scale. Interestingly, efficiency could also be controlled or required from a top down level ensuring a minimum conversion efficiency for example. [*Favrat et al.*, 2008] show how this can be introduced into a law. For this reason, this work

Table 1.1 – Definition of urban planning: R = Refurbishment, I = Infrastructure, CU = Central Utilities, DU = Decentralized Utilities and E = Emission Control

		Spatial Levels				
		(Inter) National	Regional	City	Community	Building
Planning Levels	(Inter) National Plan	I, E	I, E	E	E	E
	City Wide Plan	-	I, E	I, E, CU	I, E, CU	E (DU)
	Community Plan	-	-	-	I, E, CU	E, (DU)
	Construction Project	-	-	-	I, CU, (DU)	R, CU, DU

focuses on the city to community level planning.

Various works such as *Libbe et al.* [2010] take a broader approach, because they discuss technical and social challenges of infrastructure and the options for their realization through urban planning. These methods often fail in generating practical solutions since they are typically made at higher levels and do not consider the local situation. Concerning energy system planning, they predict a trend of more co- or tri-generation systems (with local district-energy networks) and the increasing decentralization of the electrical energy conversion systems. They also suggest a stronger collaboration between actors confronted with the planning of thermal services to the ones with electrical services. In addition, they underline the importance of urban planning to integrate the potential in renewable energies through the use of storage mediums. These types of work provide general guidelines, however do not solve concrete problems. During this work, the gap between theoretical approaches and concrete solutions should be reduced.

Figure 1.4 shows *Kemfert* [2003]'s model classification for applied economic environmental energy modeling. The Figure shows that system optimization models work on a local geographical scale with a rather short term planning horizon, such as one typical year representing the lifetime of the system, using a bottum-up approach. In this thesis, system optimization models are used, which are limited to a rather short term planning and a local impact. System

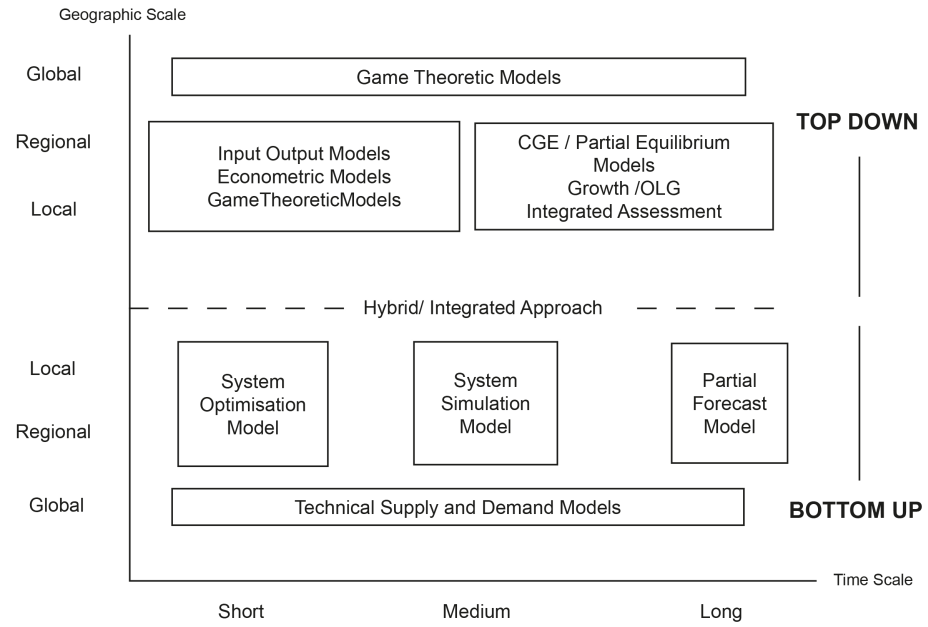


Figure 1.4 – Kemfert's model classification for applied economic environmental energy modeling [*Kemfert*, 2003]

optimization models are used in combination with technical supply and demand models to size the energy system. For the optimization, the objective function decides about the result.

1.2.3 Objectives of Energy System Design

The design of urban energy systems is not only a multi-scale problem but has also multiple objectives. The multi-objective analysis exists for the reason that different and possibly conflicting objectives arise during an optimization such as mitigation of the environmental impact and cost minimization. *Hugo and Pistikopoulos* [2005] developed a framework for supply chains based on a MILP formulation with different objectives. They combine the Eco-Indicator 99, which is a Life Cycle Analysis impact assessment method indicator measuring the potential environmental damages through multiple categories in cost minimization problem on European scale. *Gerber* [2012] demonstrates that when integrating environmental impacts into the decision making process, different system configuration based on the decision-maker preferences can be chosen.

Compared to costs which have the drawback of only considering elements that have a price, using a life cycle indicator has the advantage of extending towards a more environmentally conscious development. Costs often do not include the recycling or decommissioning of installations, neither are all emissions priced. In addition, the marginal price of a unit that is produced at high quantities profits from important economies of scale. New technologies that do not have yet a high market penetration, have a higher marginal price on the market, even if they perform better from an efficiency and/or environmental point of view. For strategic

planning, a single cost minimization objective might be a backwards oriented goal, that can undermine future opportunities, because only technologies with high market penetration are chosen due to their lower production costs. Many research applications try to avoid this problem by using a multi-objective (optimization) approach to show the trade-off between conflicting objectives such minimizing costs and minimizing environmental impact at the same time. The review of *Pohekar and Ramachandran* [2004] demonstrates the shift towards using multi-objective approaches from the early nineties on especially in renewable energy planning. *Gasparatos et al.* [2008] add that not a single indicator exists that covers all aspects, therefore working with a set of objectives is a valid option. This work is building up on the previous works in the domain from *Bürer* [2003], *Weber* [2008], *Gerber* [2012] and *Fazlollahi* [2014].

1.2.4 Optimization Techniques for Energy System Design

The energy field is of particular interest to researchers. Therefore a lot of different works with different optimization approaches have been proposed. In urban energy system design, the following problems need to be solved:

- superstructure level (also called synthesis level): A superstructure is used to include in a single problem the possible options for the energy system design and to describe their possible interactions. Solving the superstructure based optimization problem therefore results in the equipment selection and in the definition of the system configuration (i.e. how the units are interconnected),
- design level: sizing of the selected equipment,
- operation level: mass flows, temperature levels, pressures, part load behavior and ramp up and down times leading to an operational strategy.

Frangopoulos et al. [2002] examine the first two: synthesis and design. They underline the importance of optimization techniques for finding solutions, but also mention that solving all of them at once is methodologically and computationally difficult.

Curti et al. [2000] have explored the optimal design of heat pumping based district heating systems. Their approach mainly considers the design and the operational levels. They solve the problem using a single objective optimization problem. *Bürer* [2003] has considered a superstructure approach for the selection of the best configuration for the urban system using a multi-objective optimization approach.

The relevant topics in the field of mathematical programming are reviewed by *Grossmann* [2012], even though he restricts himself to enterprise-wide optimization where his paper is also (or especially) valid for system engineering:

- Solving MINLP problems remains a non-trivial tasks even though progress has been made and continues to be made in this area. Therefore a majority uses approximate MILP solutions in combination with different strategies.

- stochastic programming for the integration of uncertainty.
- Decomposition approaches use either a Lagrangean [Geoffrion, 1972], Benders [Benders, 1962], bi-level [Iyer and Grossmann, 1998] or rolling horizon approach[*Sethi and Sorger*, 1991].
- Multi-objective approaches are based on either transformation, non-pareto or pareto approaches where the latter two depend often on meta-heuristics (but could also use one of the decomposition approaches mentioned before).

Formulating a problem is generally easier than solving it. *Grossmann* [2012] review already points out that solving MINLP is not trivial, using an approximate MILP instead can be solved more easily.

The optimization problem can be solved either simultaneously or in a sequential way. The sequential way is typically using an explicit or pre-defined description of the different options extracted from the superstructure, the sizing is typically done by an heuristic algorithm and the performances are calculated using simulation tools in which the operation strategy is typically defined by a set of rules. The use of an optimization technique using a black box approach is then typically used to calculate the optimal sizes of the equipment.

In the simultaneous approach, the difficulty comes from the size of the problem and its non linear nature. A detailed review can be found in [*Fazlollahi*, 2014, Chapter 1].

Generally, a lot of different configurations (or scenarios) are possible, when choosing the equipment for an energy system and therefore it is not possible to calculate them all by hand. Often, the 3 problems listed above are solved in a step by step approach, which is practical but does not guarantee to give optimal solutions.

Using optimization techniques in order to systematically find best solutions is an obvious answer to theses problems. On the level of the synthesis, [*Voll*, 2013] proposes an automated optimization framework. Compared to by hand selection of scenarios, the superstructure free approach allows studying a high number of combination by defining who can be connected to whom. *Curti et al.* [2000], *Bürer* [2003] and *Weber* [2008] pre-define them.

On the design level and operational level, pinch analysis [*Linnhoff*, 1997] with the design of heat exchanger network is a proven methodology. Pinch analysis is generally solved with MILP formulation [*Papoulias and Grossmann*, 1983] and [*Maréchal and Kalitventzeff*, 1998] at relatively low cost in case of single period problems. The multi-period formulation [*Iyer and Grossmann*, 1998] of the same problem multiplies the number of binary variables as a function of number of periods. This formulation is therefore computationally heavy. [*Marechal and Kalitventzeff*, 2003] extended the method for the integration of energy conversion systems.

For big models, decomposition algorithms can be an elegant solution. [*Iyer and Grossmann*, 1998] propose a bi-level decomposition to solve their multi-period problem.

[*Grossmann*, 2012] does not mention the decomposition via genetic algorithm, described in different articles [*Yokoyama et al.*, 2002]. Proving optimality of this approach is difficult and can not be done. Because the genetic algorithm uses a black-box approach for solving the underlying steps, it can be slow in convergence. [*Rios and Sahinidis*, 2013] compares different derivative free algorithms and shows that better algorithms exist than a genetic algorithm. If the proof of optimality is not important, [*Fazlollahi*, 2014]'s multi-objective decomposition approach can be used and has successfully been demonstrated.

For the goal of thermal storage integration, pinch analysis [*Linnhoff*, 1997] is a proven approach [*Fazlollahi et al.*, 2012, *Varbanov and Klemeš*, 2011]. Pinch analysis [*Linnhoff*, 1997] was initially developed as a steady-state model for continuous industry processes. The introduction of a "dynamic constraint", [*Neumann and Morlock*, 2002, p.594], where the state $t+1$ depends on the state t, into a steady-state model allows for linking the heat cascades of different periods over the storage model together. This so-called time slice model[*Linnhoff*, 1997] considers variations as function of time. This enables to consider stochastically available resources such as solar energy as well as energy storage options within one model. In a simulation based approach, *Angrisani et al.* [2014] shows that considering the temperature levels is a key point for calculating the system performance when integrating a thermal energy storage.

When comparing formulations of the same problem, *Ommen et al.* [2014] shows that when using optimization of an energy system with storage in a dispatch problem, the LP can lead to similar solutions with only linear constraints in terms of cost than the MILP formulation with more constraints such as part load efficiency. However, the operating strategy changes based on the optimization approach in his model, where the heat pumps are used more in the MILP model than in the LP model. (Without knowledge of which solver settings are used for the MILP problem, there is a chance that a smaller MILP gap changes solution and might be closer to the one of the LP problem.) His MILP model solves in about a minute compared to the detailed NLP problem with about 14 hours of calculation time giving very similar results.

Even though academics provide methodologies for solving these typical problems, they are not very frequently used in industry or on a (local) energy utility level, because introducing all constraints to get to feasible solutions is either judged to be impossible or too time consuming [*Klatt and Marquardt*, 2009]. In addition decision makers often prefer a solution that they understand compared to a solution that they don't understand but that is better according to an optimization model. Instead, engineers still rely on their experience and simple scenario analysis with very few choices to identify the best solutions. When the problem is simple enough, experienced engineers can find the optimal or a close to optimal solution.

From a practical point of view, the multi-objective analysis is often criticized, because it does not give one solution but a Pareto-frontier of solutions. Therefore a multi-objective analysis does not tell which of the proposed solution is the best one, it helps however to identify the trade-off of the different objectives for the energy system. Additional methods are applied to identify the best solution such as clustering the solutions in groups or the introducing weights

for each objective. This means besides the fact that time needs to be spent on calculating the Pareto frontier, another algorithm chooses one point out of all solutions. Often industry partners have not been convinced about this approach even though they agree that conflicting objectives exist, they prefer using a single objective approach using costs. When they see a Pareto curve, they only look at the cost minimum. This implies for the results presentation that a few well chosen points are enough. With a multi-objective optimization, the result for each objective can be that point.

1.2.5 Exergy as Indicator

Exergy efficiency is used as an indicator to measure the thermodynamic degree of perfection of a given system. Numerous authors link exergy to sustainable development: [*Cornelissen*, 1997], [*Rosen and Dincer*, 2001], [*Wall and Gong*, 2001], [*Gong and Wall*, 2001] and [*Gasparatos et al.*, 2008].

The Second Law of thermodynamics introduces irreversible transformations. Compared to an energy approach, exergy can therefore be lost or destroyed. According to [*Cornelissen*, 1997], minimizing the loss of exergy leads to sustainable development. *Dincer and Rosen* [2004] state that the path from non-sustainable development to sustainable development passes by the three major steps of:

1. increase exergy efficiency,
2. reduce the exergy-related environmental degradation and
3. use of sustainable exergy resources.

Rosen and Dincer [2001] consider exergy as the confluence of energy, environment and sustainable development. Further, *Rosen and Dincer* [2001] state that systems with a high exergy efficiency have a lower resource use and therefore are automatically more sustainable than a similar system with a lower efficiency. *Wall and Gong* [2001] points out that the exergy can measure the difference in the environment when emission and pollutants are injected. *Gong and Wall* [2001] link exergy to an life cycle assessment approach based on *Szargut* [1987] and *Cornelissen* [1997] to use exergy as a quality indicator.

Gasparatos et al. [2008] point out that exergy analysis is universally accepted because it is based on thermodynamics. In an supply chain, the exergy efficiency of each step allows to define the weakest point. However, the definition of the required reference conditions introduce a relativity in the results. *Torío et al.* [2009] review reveals that different exergy efficiency definitions and different definitions of renewable exergy flows lead to difficulties when comparing studies and their results. Nevertheless, the stated authors conclude that exergy is the framework to use, when linking sustainable development and thermodynamic perfection. But the conclusion insist to standardized the exergy approach further to ensure comparability of studies. *Bösch et al.* [2007] address the missing standardization through a

consequent application of the exergy method on a whole database to ensure the best possible comparability.

In times where humanity is looking forward to identifying environmental and sustainable energy systems to reduce energy consumption, exergy analysis in combination with life cycle assessment offers a key framework to consider: The life cycle assessment approach allows to consider the path of the resources before entering the energy system. The equipment used within the system is also evaluated based on the same life cycle assessment approach taking into account the exergy used for their production and the decommissioning at the end of life time. The operational exergy can be reduced when using the heat cascade because the heat cascade allows to reduce the exergy used in each temperature interval.

An exergy based life cycle assessment framework can overcome the shortcomings of a simple monetary framework to judge an urban energy system on efficiency, because the exergy use throughout the whole supply chain and through the lifetime are considered. *De Meester et al.* [2009] demonstrate the use of the cumulative exergy demand as a decision help for the choice of a type of construction to perform.

Roosa [2010] connects exergy and sustainability in a larger framework going from buildings to (urban) energy systems. He then also addresses the policy effects sustainable planning. Exergy is mostly used in academia and rarely applied by practicing engineers. *Dewulf et al.* [2008] point out that this is probably linked to the fact that the concepts of exergy are often taught using the rather abstract entropy definition. Applying the concept of useful energy such as presented in *Haldi and Favrat* [2006] helps as it is easier to access. In the field of process engineering, it is an unopposed fact that exergy analysis is an effective tool to identify system inefficiencies. Frameworks extending to other fields such as the built environment existent but are rare. *Schlueter and Thesseling* [2009] developed a prototype tool for the integration of exergy analysis in early design studies. Legal obligation in time of changing energy politics could also become a driving force. The Canton of Geneva has included exergy evaluation of energy systems in a law in 2010, based on *Favrat et al.* [2008]. In addition, this work demonstrates a ready- and easy-to-use guide to the use of exergy efficiency as an indicator. *Rager et al.* [2011] apply this framework in a case study and propose it for selecting priority buildings for further investigation and energetic refurbishment, if the exergetic performance is confirmed.

Other frameworks try to convert exergy to monetary units such as the substitution of capital expenditures [*Müller et al.*, 2011], in an accounting framework [*Sciubba*, 2001] or via for the evaluation of a tax [see *Szargut*, 2005, Chapters 6 and especially 7]. The exact merits of the cited frameworks is still a lively debated.

1.2.6 Uncertainty in Urban Energy System Design

Dealing with energy system design (for planning on an urban scale) means taking uncertainty into account. Besides the uncertainty of input data, the system itself is frequently altering through changes in land use and/or energetically relevant changes to the existing building stock. *Hoffmann* [2001] defines three types of uncertainty:

> parameter uncertainty, model uncertainty, and decision rule uncertainty.

The parameter uncertainty in general can be accessed with the methods that *Saltelli et al.* [2008] propose. These methods execute the model systematically over the pre-defined uncertainty range of the parameter. The result is used among other for factor fixing, which means identifying the non-important uncertain parameters and fixing them to a given value. Because each model represents only a simplification of a real world phenomenon, each model contains also model uncertainty. Model uncertainty can often only be accessed by experts ensuring that the model is representing a phenomenon correctly and that the input-output combinations are correct. Model uncertainty is not considered because energy models are non-validatable and "doomed" to inaccuracy [*Hodges et al.*, 1992]. Decision rule uncertainty is very difficult to model as politics have a strong influence on it.

For this work, the parameter uncertainty is further considered, because perturbing a parameter can change the result of an optimization remarkably leading to (highly) sub-optimal or infeasible solution [*Bertsimas and Sim*, 2004]. Generally, parameter uncertainty can either be accessed with an stochastic or robust approach. Stochastic programming approaches [*Birge and Louveaux*, 2011] assume that the uncertainty has a probabilistic nature, e.g. a randomness exists. Therefore a stochastic approach requires further knowledge of the type of randomness and the distribution of the parameters. Through frequent model runs, an expected cost can be minimized (in the case of a cost minimization). In contrast, robust optimization [*Soyster*, 1973] tries to find a solution that fits all cases, the design case but also the one with the worst case parameters. Besides classifying the parameter's uncertainty into sets, it tries to remain computationally light.

In the field of energy system design, *Dubuis and Maréchal* [2012] use stochastic optimization and discuss different methods and conclude that they are quickly computational heavy. *Keirstead and Shah* [2013, Chapter 12] or also *Keirstead and Calderon* [2012] use a Monte Carlo method proposed by *Saltelli et al.* [2008]: the study allows for identifying key factors of the model that can be used for policy implications. *Fazlollahi* [2014] proposes a variance based analysis on the points of the Pareto-frontier. Varying systematically the economic input parameters create a wider Pareto-frontier allowing to choose the most reliable point defined as point on the Pareto frontier of the design case using average parameter values. *Tock* [2013, Chapter 8.3] uses a similar approach based on Monte Carlo simulations ranking the different solutions according to the number of times they appear on the Pareto frontier.

Urban energy system design studies often consider mean values for uncertain parameters without the quantification of the relevance of an individual uncertain parameter. The system might not perform as expected, because a key parameter might deviate during the project decreasing a system's performance. If the key parameters to energy system design are known, they could at least be communicated to each actor during the project realization ensuring that the importance is taken into consideration.

1.2.7 Existing Models and Tools for Urban Energy System Design

[*Pfenninger et al.*, 2014] resume the challenges that every model (and every tool) in the field of urban energy system modeling needs to address:

1. resolving time and space,
2. balance between uncertainty and transparency,
3. handle growing complexity of energy systems and
4. human behavior.

Keirstead et al. [2012] classify generally the existing urban energy models: they all try to find a coherent model complexity, rely on a specific data quality and uncertainty, link models more or less across different sectors (model integration) and might guide a policy makers. [*Klosterman*, 2012] concludes that whether or not a complex or simple model is used, it is most important to verify the model assumptions: when they are of poor quality, any model will produce nonsense.

For the field of urban energy systems design, *Keirstead et al.* [2012] deduce from their study that still a number of challenges remain to be tackled, especially the model complexity and (input) data uncertainty. *Kavgic et al.* [2010] focus on bottom up energy demand models in the UK. They stress the fact that only limited amount of input data is publicly available as well as the fact that the models are not open enough. Results can not be reproduced. In addition, not much information exists quantifying the importance of input data and its impact on the outputs. More and more specific models exist addressing subfields of urban energy models. In the more specific field of sustainable building design, *Evins* [2013] reviews methods for sustainable building design. *Kanters et al.* [2012] study the niche for architects integrating solar design.

Part of the models are used to create easy-to-use tools (for an audience also outside the academic world). Each tool has its own niche with its specialization and target audience. A comprehensive tool review is done by [*Connolly et al.*, 2010]. This review can be used as a guideline on how to choose the correct tool. Interestingly enough, the choice of temperature levels is not addressed in this tool review. *Manfren et al.* [2011] completes this review with a finer classification. Most tools on the market, commercial or free, offer system simulation but only few tools offer optimization. In addition, the definition of the word "optimization" in a

tool's description often refers to comparison of scenarios instead of mathematical optimization.

Linked to the fact that in an urban environment, a lot of different goals and actors, users and interests exist, a wide variety of tools can be found. [*Erhorn-Kluttig et al.*, 2011] propose a model classification under three main points:

- in which planning phase can the tool be used?
- which is the specific field of application?
- who is the user?

For 6 tools, this classification is completed in Table 1.2 and extended with the 2 options: integrate storage sizing and evaluate uncertainty. The 6 tools are: Citysim [*Kämpf and Robinson*, 2009] is an dynamic urban building simulator, TOPENERGY is an energy supply simulator, EnerGis [*Girardin et al.*, 2010] a static heating requirement estimator based on statistical information, EnergyPlan is a simulator for national energy systems, MEU is a planning and managing platform for communities and cities and the system design tool osmose.

Even though the classification is already detailed, it should be refined, because it is not possible to distinguish between simulation and optimization based approach for different applications. None of the tools proposes any ways to integrate uncertainty in the planning process besides MEU: it has a meta data system that judges the quality of each data set, however does not not aggregate it any further to a higher level. Also, the optimization methods are limited. When looking at the field of application for the last 3 points, storage, uncertainty and optimal energy system design, none if the existing tools can work on the 3 points at once. All tools need to be adapted.

With the rising number of tools that are available, it is difficult to keep track of all the developments ensuring that the latest model and/or latest tool are known to deploy them. Already in EPFL, 4 different commercial tools exist:

- CitySim [*Kämpf and Robinson*, 2009] for dynamic urban building simulations,
- MEU [*Rager et al.*, 2013c] and [*Capezzali et al.*, 2013] for the management of urban energy systems and decision support for communities and utilities,
- lesosai [*Lesosai*, 2013] for building certification and thermal balances and
- *epiqr* [*EPIQR and ESTIA*, 2004] specialized on energetic building refurbishment.

Numerous research tools in the field exist, that are not commercially available. Three examples from EPFL are stated here: Osmose is a system design tool, EnerGis [*Girardin et al.*, 2010] combines information on cities in a geographical information system and SIMAPE a multi-agent simulation of urban energy planning for a decision support system.

Table 1.2 – Characterization of energy models adapted from *Erhorn-Kluttig et al.* [2011, p.114] extended with storage sizing and uncertainty evaluation

		Decision Support					
		CitySim	TOPENERGY	EnerGis	EnergyPLAN	MEU	osmose
Planning Phase	Strategic Planning	-	X	-	X	X	(X)
	Predimensioning	-	X	X	X	-	(X)
	Detailed Planning	X	X	-	-	-	(X)
	Monitoring	-	-	-	-	X	-
Field of Application	Indiv. Building Energy Demand Estimation	X	-	X	-	X	-
	Indiv. Energy Supply Technology Choice	-	X	X	X	X	X
	Multiple Indiv. Building Energy Demand Estimation	X	-	X	X	X	-
	Multiple Energy Supply Technology	-	X	X	X	X	X
	Mixing Multiple Energy Supply Technology	-	X	X	X	X	X
	DHN	-	X	X	X	X	X
	Integrate Storage Sizing	-	X	-	X	-	X
	Evaluate Uncertainty	-	-	-	-	X	(X)
	Optimization (of the Energy System Design)	-	(X)	-	-	-	X
User	Urban Planner	X	-	-	X	X	-
	Energy Utility	-	X	-	X	X	-
	Construction Enterprise	X	-	-	-	-	-
	Engineering Company	X	X	-	X	-	-
	Architect	X	-	-	-	X	-
	Scientist	X	X	X	X	X	X

1.3 Conclusion

Based on the past of heat planning, the current situation in Switzerland with regard to urban energy system planning is discussed. To reach the nations goal fixed with a top down approach, bottom-up models are needed that can propose energy system designs considering demand and supply side options. This framework needs to work on a community to city level, because these systems will be implemented on this spatial scale.

In order to avoid backwards oriented planning by only cost based optimization, a second indicator should be used to judge the sustainability of the system. Exergy-based indicators should be considered because they can judge quantity and quality of exergy conversion at once.

Summarizing the current state of the art in optimization suggests that using an MILP problem formulation, when functions can be sufficiently linearized, is a promising approach that

guarantees optimal solutions (with respect to the hypotheses made for linearization). Energy integration offers an existing MILP framework that can be used for thermal storage integration.

There is currently no tool proposing an integrated methodology for designing energy systems under uncertainty. On the methodology side, enough independent methods exist to address the discussed problems independently, but rarely together. This dissertation tries to close the gap between the existing methods.

1.4 Objectives of this work

The following challenges are addressed in this work:

- energy demand modeling based on data availability, Chapter 3,
- deterministic data reduction with the help of clustering respecting sequences, if needed, Chapter 4,
- integration of a stochastic renewable energy sources with the example of solar thermal energy, Chapter 5,
- integration of (seasonal) storage, Chapter 5,
- combining stochastic renewable sources and storage sizing to a design method for urban areas, Chapter 5,
- using a building's mass as additional energy storage, Chapter 6,
- introducing retrofitting to the demand side management, Chapter 7 and
- a way of quantifying uncertainty, Chapter 8.

2 Problem Formulation for Urban Energy System Design Methodology

The second chapter describes the general optimization or mathematical programming approach chosen. The objective functions are explained as well as the additional constraints that exist for the every of the following chapters.

According to [*Keirstead and Shah*, 2013], the large number of choices for urban energy system design needs methodologies to systematically compare and evaluate different system configurations. Modeling, simulation and optimization can do this. Finding the best solution with the help of simulation can be very time consuming and is be limited to a certain number of pre-selected combinations. Optimization or mathematical programming is therefore a widespread tool.

2.1 Definition of Optimization Problem

In order to present real world phenomena such as heat exchanges correctly, the optimization problem of an urban energy system can be formulated as a non-linear optimization problem. However, nonlinear problems are generally harder to solve [*Grossmann*, 2012], especially if an exact solution is required.

In the here presented framework, non-linear problems are consequently estimated with the help of one linear function, if acceptable or a set of linear functions, if necessary. The optimization problem is then transformed into a linear and/or mixed integer linear problem which have the advantage of being generally easier to solve. The optimality for the solutions found can be proofed compared to heuristic methods that are often used to resolve non-linear problems. When a non-linear function is represented as a linear one, the error made through the simplification is kept below the uncertainty of the underlying data set that is used to define the non-linear function.

If one linear function is sufficient to estimate a non-linear function, the optimization problem can be formulated as linear problem. The simplex algorithm solves linear problems.

When one linear function is not sufficient for the estimation of a non-linear function, the form of the non-linear function decides whether integers have to be introduced or not: A convex non-linear function can be represented as a set of linear function with weights guaranteeing that only one function is active. A non-convex non-linear function can only be represented with a set of indicator integer variables leading to a mixed integer linear problem. Mixed integer linear problems are generally solved with a branch and bound algorithm that uses an implicit enumeration of all possible combinations of integer variables in a search tree. Choosing the nodes of the search tree is called "branching." The upper (or lower) bound of a branch can then help to avoid exploring further nodes in the same branch, when the bounds indicate that only suboptimal nodes (and further subbranches) exist. A MILP problem can be solved exactly in exponential time, heuristic approaches provide generally suboptimal solutions in polynomial time.

As identified in Section 1.2.4, an existing approach using MILP formulation consists in using energy integration with the heat cascade. A MILP model formulation is done in the following way:

$$\begin{aligned} &\text{Objective: minimize } c^T \boldsymbol{x} + d^T \boldsymbol{y} \\ &\qquad \text{Constraints: } A_1 \boldsymbol{x} + A_2 \boldsymbol{y} \leq b \\ &\qquad\qquad l \geq \boldsymbol{x} \geq u, \quad \boldsymbol{y} \in 0,1 \end{aligned}$$

x represents the continuous decision variables, y the integer (or binary) decision variable, c and b are vector of known parameters and A is the coefficient matrix. l gives the lower and u the upper bound for the continuous variable. For a linear problem (LP), only x exist in the problem formulation. In a mixed integer linear problem (MILP) at least one y exists also.

The objective function for urban energy system design is typically written as a minimization of utility investment and operation costs. The constraints in this work are of the following form:

- cost or life cycle based accounting,
- energy balance,
- heat cascade,
- time,
- technology availability and performance.

The following formatting rules are applied: **Bold** formatting indicates variables of the model, parameters are in normal font style. The variable $\boldsymbol{y}$ is the only binary integer variable of the model. All other variables described below are continuous ones, bigger or equal to zero, when not indicated otherwise. Based on the notation used in *Borel and Favrat* [2010], flows entering the system have the superscript +, the ones leaving the system a −, but a positive value. Existing upper limits are indicated, otherwise the variable has no upper bound. The lower bound is in general zero, but for the sake of completeness it is indicated. Total values are written in upper case, specific values are written in lower case.

2.2 Objective Function

Two different formulations of the objective function are used: The first one minimizes the annual costs. The specific investment costs $invC_u$ of unit u is multiplied with the installation factor $\boldsymbol{s_u}$ and the annualization factor f_A common for all units. The operating costs are calculated through summation over all periods np and all time slices nt of the specific fuel costs c_f which are multiplied by the unit multiplication factor $\boldsymbol{f_{u,p,t}}$ for each period p and time slice t and the nominal heat $\dot{Q}_u$ produced per unit. w_p represents the weight or number of repetitions of the current period, e.i. the number of times the day appears in a year. The duration of the time slice t in p is given through $d_{p,t}$. A certain quantity of electricity $\dot{\boldsymbol{E}}_{el}$ can either be imported ($\dot{\boldsymbol{E}}^{+}_{el}$) at a price of c^{+}_{el} or exported ($\dot{\boldsymbol{E}}^{-}_{el}$) at a price of c^{-}_{el}.

$$F_{obj,costs} = min\Bigg(f_A \cdot \sum_{u=1}^{nu} invC_u \cdot \boldsymbol{s_u} + \sum_{p=1}^{np} w_p \cdot \sum_{t=1}^{nt}\sum_{u=1}^{nu}\left(c_{f,u}\cdot \boldsymbol{f_{u,p,t}} \cdot \dot{Q}^{+}_{u} + c^{+}_{el,p,t}\cdot \dot{\boldsymbol{E}}^{+}_{el,p,t} - c^{-}_{el,p,t}\cdot \dot{\boldsymbol{E}}^{-}_{el,p,t}\right)\cdot d_{p,t}\Bigg) \tag{2.1}$$

The second formulation considers a minimization of Cumulative Exergy Demand (CExD). The formulation is the same, however the input values change. The construction and production of a utility u has a specific CExD value $invCExD_u$. Each fuel has also a specific CExD content $cexd_{f,u}$. The remaining parameters are the same as in (2.1).

$$F_{obj,cexd} = min\Bigg(f_A \cdot \sum_{u=1}^{nu} invCExD_u \cdot \boldsymbol{s_u} + \sum_{p=1}^{np} w_p \cdot \sum_{t=1}^{nt}\sum_{u=1}^{nu}\Big(cexd_{f,u}\cdot \boldsymbol{f_{u,p,t}} \cdot \dot{Q}^{+}_{u} + cexd^{+}_{el,p,t}\cdot \dot{\boldsymbol{E}}^{+}_{el,p,t} - cexd^{-}_{el,p,t}\cdot \dot{\boldsymbol{E}}^{-}_{el,p,t}\Big)\cdot d_{p,t}\Bigg) \tag{2.2}$$

The annualization factor f_A for $n = 25$ years and an interest rate of $i = 6\%$ has a value of 0.0782, because the same life time as for the CExD values is used for the sake of an equal comparison. The main stakeholders, cities or communities, can afford to calculate with such as a long term horizon. The annualization factor is calculated in Equation (2.3):

$$f_A = \frac{(1+i)^n \cdot i}{(1+i)^n - 1} \tag{2.3}$$

It is based on the hypothesis that each year the same amount of money has to be paid back and that all units have the same lifetime. Both formulations guarantee the balance between the unit investment costs or respectively the unit production CExD and the operating costs or the fuel's CExD content via the annualization factor: Multiplying the total investment cost with this factor results into the annual investment costs that can be put in direct relation to the annual operating costs. It is also applied to the CExD objective function ensuring that the

same approach is used for both objective functions. As an alternative, one could divide the CExD only by the lifetime in years.

2.3 From Exergy to the cumulative Exergy Demand (CExD)

The second law of thermodynamics implies that "all real process are irreversible" [*Szargut*, 2005, p.2]. [*Szargut*, 2005] further uses the exergy concept to measure thermal efficiency with the exergy efficiency. Compared to the first law efficiency (and the related indicator cumulative energy demand), it takes the quality of the transformation or irreversibility into account. It is therefore an indicator that measures and combines quantity and quality of a conversion chain. It does not cover toxicity, scarcity nor does it distinguish between basic energy needs or luxury energy needs. Neither is it a good distinction between renewable and finite or non-renewable resources. [*Borel and Favrat*, 2010, Chapter 10] define the exergy rate balance as follows:

$$\sum_k \left[\dot{E}_k^+\right] + \sum_i \left[\dot{E}_{qi}^+\right] + \sum_n \left[\dot{E}_{yn}^+\right] = \dot{L} \geq 0 \tag{2.4}$$

$\dot{E}_k^+$ as the work-power received by the system at the level of the engine k;
$\dot{E}_{qi}^+$ as the heat exergy received from the reservoir at the temperature T_i;
$\dot{E}_{yn}^+$ as the received transformation exergy rate and
$\dot{L}$ as the global exergy rate loss (always positive according to the Second Law).

The exergy linked to a transfer or a storage of energy is defined as the potential of maximum work which could ideally be obtained from each energy unit being transferred or stored (using reversible cycles with the atmosphere as one of the energy sources either hot or cold).

The exergy approach allows to quantify in a coherent way both the quantity and the quality of the different forms of energy considered.

The overall system exergy efficiency is defined as:

$$\eta = \frac{\text{Services delivered}}{\text{Exergy received}} = \frac{\sum[\dot{E}^-] + \sum[\dot{E}_q^-] + \sum[\dot{E}_y^-]}{\sum[\dot{E}^+] + \sum[\dot{E}_q^+] + \sum[\dot{E}_y^+]} \tag{2.5}$$

where all terms are positive, differentiating between the terms entering the system with "+" sign and the positive terms (services) delivered by the system with a "-" sign.

- Mechanical and Electrical exergy: $\dot{E}$,
- $\dot{E}_q$ represents the heat exergy with:

$$\dot{E}_q^- = \int \theta \delta \dot{Q}^- = \int (1 - \frac{T_a}{T}) \delta \dot{Q}^- \tag{2.6}$$

  Carnot Factor $\theta = (1 - \frac{T_a}{T})$ and the heat transfer rate $\dot{Q}$,

- $\dot{E}_y$ symbolizes the transformation exergy of the masses $\dot{M}$ being in the system or transiting through the system:

$$\dot{E}_y^+ = \sum_j \left[\int k_{czj} \dot{M}_j^+ \right] - \frac{d(J_{cz})}{dt}. \tag{2.7}$$

  k_{cz} represents the total specific coenthalpy and J_{cz} the coenergy defined through the sum of the total internal energy U_{cz}, the pressure volume work $P_a V$ and the entropy creation $-T_a S$.

In the field of Life Cycle Assessment, the ecoinvent database [*Althaus et al.*, 2010] is used because it offers standardized data sets especially for Switzerland. One of the indicators proposed is the CExD proposed by *Bösch et al.* [2007]. It is similar to the concept of cumlative consumption of exergy from *Szargut* [2005, Chapter 3] presented in [*Szargut*, 1987]: the chain of production processes from natural resource to the finished product are considered. However Szargut's concept is only linked to the process itself, the operational exergy demand. The CExD also includes construction and decommissioning exergy of the plant and looks at the supply chains used for resources entering the plant. It is defined as:

$$CExD = \sum_i M_i * ex_{cz,i} + \sum_j n_j * r_{ex-e(k,p,n,r,t),j}. \tag{2.8}$$

$CExD$ = cumulative exergy demand per unit of product or process (MJ-eq)
i, j = substance
M_i = mass of material resource i (kg)
$ex_{cz,i}$ = specific "embodied" exergy (MJ-eq/kg) = exergy used to produce and recycle a kg of substance i (MJ-eq/kg)
n_j = amount of energy from energy carrier j (MJ)
$r_{ex-e(k,p,n,r,i),j}$ = exergy to energy ratio of energy carrier j (MJ-eq/MJ) (based on *Szargut* [2005])
ch = chemical
k = kinetic
p = potential
n = nuclear
r = radiative
t = thermal exergy

The CExD is a method integrating the quality of the energy conversion during the production and construction of a utility, the operating exergy used during equipment life time and the recycling after end of use. The method is close to cumulative energy demand (CED), however it is more complete because it measures not only quantity but also quality of conversion in the supply chain.

Eight ecoinvent categories exist for CExD: fossil, nuclear, hydro power, biomass, other renewables, water, minerals and metals. The CExD value of one kilowatt hour heat produced through a gas boiler or a solar thermal panel have almost the same value. However the kilowatt-hour

produced with solar energy is almost entirely produced through the use of renewable energy. In order to minimize the use of non-renewable energy, only the non-renewables are considered. In the here presented work the non-renewable parts fossil, nuclear, minerals and metals are considered. Choosing to use only the non-renewable CExD is inline with the guideline of sustainable development suggested by [*Haldi and Favrat*, 2006].

When using CExD values directly from the data base, hypothesis about the electricity mix within the supply chain of the used technologies are included into the analysis. For this work, the database version 3.1 of ecoinvent is used.

2.3.1 Extrapolation of missing CExD Values

The CExD values for each technology are found with a specific size (and often in a specific system configuration) in the ecoinvent data base. The actual size of the equipment is however object of the optimization and most of the time the data is only available for one equipment. A way for data extrapolating is found in *Gerber et al.* [2011] based on the equation:

$$\frac{E_i}{E_{ref}} = \left(\frac{A_i}{A_{ref}}\right)^{k_{j,ref}} * c_i \tag{2.9}$$

with E being the emission of the Life Cycle Impact of the elementary flow i or of the known reference flow ref respectively.
A is a functional parameter related to the size such as heat exchanger surface or power.
k is the exponent, the correction factor c can only be determined when more than 3 data sets are available.
Otherwise, it is dropped.

Compared to the traditional LCIA, where extrapolation is done in a linear way, this approach considers the fact that an equipment's impact is measured proportional to the material used as in the cost functions. The three cases found in this thesis are:

- Case 1: Two data sets i and j are available

$$k_{j,i} = \frac{logE_j - logE_i}{logA_j - logA_i} \Rightarrow E_x = \left(\frac{A_x}{A_{ref}}\right)^{k_{j,i}} * E_{ref} \tag{2.10}$$

  The subscript x indicates the unknown elementary flow. The subscript ref indicates the reference equipment, which is closer to A_x.
- Case 2: Only one data set is available, extrapolation via costs

$$E_x = \frac{C_x}{C_{ref}} * E_{ref} \tag{2.11}$$

- Case 3: No data set is available, the value E_x is the sum of all individual equipment components u:

$$E_x = \sum_{u=1}^{n} E_u \tag{2.12}$$

This approach allows to extrapolate more precisely, however introducing the new costs values into the MILP requires to linearize them again.

2.4 Constraints

The size s_u of the equipment u is given through the maximum usage of the unit during all periods p and time steps t. The unit multiplication factor $f_{u,p,t}$ gives this information per unit u for all periods p and time steps t.

$$f_{u,p,t} \leq s_u \; \forall u, p, t \tag{2.13}$$

The binary unit variable y_u links the operational limits with the minimal sizing factor $f_{u,p,t}^{min}$ and the maximum one $f_{u,p,t}^{max}$ to the investment costs. It is $y_u = 1$ when the unit is purchased.

$$\boldsymbol{y_u} \cdot f_{u,p,t}^{min} \leq \boldsymbol{s_u} \leq \boldsymbol{y_u} \cdot f_{u,p,t}^{max} \; \forall u \tag{2.14}$$

The general heat cascade for each temperature interval k is given by:

$$\begin{gathered} \sum_{h_k}^{nh_k} \boldsymbol{f_u} \cdot \dot{Q}_{h,k,u} - \sum_{c_k}^{nc_k} \boldsymbol{f_u} \cdot \dot{Q}_{c,k,u} + \dot{\boldsymbol{R}}_{\boldsymbol{k+1}} - \dot{\boldsymbol{R}}_{\boldsymbol{k}} = 0 \quad \forall k = 1, \cdots, nk \\ \dot{\boldsymbol{R}}_{\boldsymbol{1}} = 0 \quad \dot{\boldsymbol{R}}_{\boldsymbol{nk+1}} = 0 \quad \dot{\boldsymbol{R}}_{\boldsymbol{k}} \geq 0 \quad \forall k = 2, \cdots, nk \end{gathered} \tag{2.15}$$

$\dot{Q}_{h,k,u}$ and $\dot{Q}_{c,k,u}$ are the nominal heat loads of the hot (h) and cold (c) streams, respectively, which exist in temperature interval k for unit u in period p and time slice t,
$\dot{\boldsymbol{R}}_{\boldsymbol{k+1}}$ is cascaded heat from the higher temperature level entering the temperature level k,
$\dot{\boldsymbol{R}}_{\boldsymbol{k}}$ is heat leaving this temperature level.

This formulation is adapted for a multi-period problem with multiple time slices t in each period p with the same restrictions on the cascaded heat $\dot{\boldsymbol{R}}$:

$$\sum_{h_k}^{nh_k} \boldsymbol{f_{u,p,t}} \cdot \dot{Q}_{h,k,u,p,t} - \sum_{c_k}^{nc_k} \boldsymbol{f_{u,p,t}} \cdot \dot{Q}_{c,k,u,p,t} + \dot{\boldsymbol{R}}_{\boldsymbol{k+1,,p,t}} - \dot{\boldsymbol{R}}_{\boldsymbol{k,p,t}} = 0 \quad \forall k, p, t \tag{2.16}$$

The heat cascade is used in process integration. Process integration and pinch analysis techniques minimize the energy consumption of industrial processes by maximizing the heat recovery. From the beginning of the mathematical formulation of pinch analysis, examples

show the relevance of pinch analysis to the field of urban energy system design [*Linnhoff*, 1997], [*Weber*, 2008] to [*Girardin*, 2012] and [*Fazlollahi et al.*, 2014b].

In energy integration, two different basic units are defined: utilities and processes. The process have utilization factor s_u fixed to 1 and represent the demand compared to utilities that are limited between f_u^{min} and f_u^{max} offering the required service. As later shown in Chapter 6, the utilization factor for a process can also be variable.

Each unit can have one or more streams attributed to it. A stream is either:

- a hot stream that requires cooling and therefore releases heat or
- a cold stream that requires heating and therefore absorbs heat.

Heat exchange between them ensures that the heating and cooling requirements are fulfilled. Under the hypothesis of steady state conditions without heat losses to the atmosphere or mechanical work, constant pressure and a constant specific heat capacity c_p between the temperatures T_{in} and T_{out} for a real single phase fluid, the heat load $\dot{Q}^+$ is given by Equation (2.17). The specific enthalpy difference of the fluid $\Delta h = h_{out} - h_{in}$ is equal to the temperature difference $\Delta T = T_{out} - T_{in}$ multiplied by the specific heat capacity at constant pressure c_p , when the mass flow $\dot{M}^+$ is constant.

$$\dot{Q}^+ = \dot{M}^+(h_{out} - h_{in}) = \dot{M}^+ \cdot c_p(T_{out} - T_{in}) \tag{2.17}$$

Variation of kinetic and potential energies are not taken into account.

The quality of the heat exchange between a hot and a cold stream can be described with the minimal temperature approach ΔT_{min}. This approach describes the trade-off between a high quality heat exchange with a low ΔT_{min} requiring a high heat exchanger surface thus a high investment cost but low external utility use. A high ΔT_{min} with a small heat exchanger surface thus a lower investment has a high external utility requirements because less heat is transferred. The value of $\Delta T_{min}/2$ is then associated to each stream:

- for each hot stream, the corrected temperature is $T_h^* = T - \Delta T_{min}/2$ and
- for each cold stream, the corrected temperature is $T_c^* = T + \Delta T_{min}/2$.

Only the corrected temperatures are introduced into the heat cascade.

2.5 Basic Technology Models

Three basic technologies are considered in the framework:

1. boilers,
2. heat pumps and
3. cogeneration.

With the heat cascade, additional technologies such as free-cooling or tri-generation can be integrated as well. For each additional technology, temperature-power profiles for each time step need to be known so that these profiles can be introduced into the heat cascade.

2.5.1 Boiler

The boiler can burn different fuel types: biomass, gas or fuel oil. A boiler can deliver heat at a fixed temperature level between 90 to 75 °Celsius. As in reality, the boiler delivers heat at 75 °Celsius.

For the boilers, the power $\dot{Q}^{+}_{Boiler}$ provided to system is divided by the corresponding efficiency ε to calculate the amount of resource $\dot{Q}^{+}_{Resource}$ used.

$$\dot{Q}^{+}_{Resource} = \frac{\dot{Q}^{+}_{Boiler}}{\varepsilon_{Boiler}} \tag{2.18}$$

The wood chips have a variable operating CExD value due to the origin of the wood chips: if they come from a local forest, the lower value can be used. If they are left over of industrial usage, the higher one should be used.

2.5.2 Heat Pump

The heat pump provides a range of temperature levels between 25 to 80 °Celsius to the urban energy systen: higher temperatures lead to lower coefficient of performance (COP) and therefore more electricity consumption. The number of available temperature levels can either be restricted to one so that only one temperature level at a time can be provided or is left unrestricted so that several levels up to all existing levels are possibly active at once. The optimization model decides which level(s) to consider based on the user's choice of maximal number of temperature levels at once allowed.

The COP is calculated according to *Girardin et al.* [2010] with the theoretical COP and the efficiency η_{COP} of the COP, a function of the resource and therefore of the technology. The theoretical COP is defined as the ratio of the hot side's logarithmic temperature difference over the logarithmic temperature difference of the hot and cold side's difference. η_{COP} varies according to the source used between 0.34 for an air/water heat pump, 0.43 for a water/water with a ground water source and 0.55 for a waste water treatment plant or lake water. The heat

delivered through the heat pump, $\dot{Q}^{+}_{HP}$, is divided through the COP_{HP} to obtain the electricity consumption of the heat pump.

$$COP_{HP} = \eta_{COP} \cdot \underbrace{\frac{\Delta T^{hot}_{lm}}{\Delta T^{hot}_{lm} - \Delta T^{cold}_{lm}}}_{\text{theoretical COP}} \tag{2.19}$$

$$\dot{E}^{+}_{HP} = \frac{\dot{Q}^{+}_{HP}}{COP_{HP}} \tag{2.20}$$

With Equations (2.19) and (2.20) the operating costs or CExD of delivering heat at different temperature levels can be calculated. The heat pump's operating CExD value is directly connected to the origin of the electricity. For the operating costs, the electricity costs of 0.25 $\frac{CHF}{kWh}$ are divided by the heat pump's COP in order to obtain the respective price per kilowatt hour of delivered heat at a specific temperature level. For the operating CExD optimization, local mini-hydro power has a value of almost 0 $\frac{kWh-Eq.}{kWh}$, the Swiss production mix is at about 2.0 $\frac{kWh-Eq.}{kWh}$ and the imported German electricity is at 3.0 $\frac{kWh-Eq.}{kWh}$.

2.5.3 Cogeneration

The cogeneration model provides heat at the same range as the boiler: 75 °Celsius. The cogeneration model is implemented as a heat driven cogeneration. The heat cascade defines how much heat $\dot{Q}^{+}_{Cogen}$ the cogeneration delivers. Dividing the heat $\dot{Q}^{+}_{Cogen}$ by the heat production efficiency $\varepsilon_{th,cogen}$ of the cogeneration, provides the amount of a resource used. The resource use divided by the electrical efficiency $\varepsilon_{el,cogen}$ results in the electricity production $\dot{E}^{-}_{el,Cogen}$ of the cogeneration.

$$\dot{Q}^{+}_{Resource} = \frac{\dot{Q}^{+}_{Cogen}}{\varepsilon_{th,cogen}} \tag{2.21}$$

$$\dot{E}^{-}_{el,Cogen} = \frac{\dot{Q}^{+}_{Resource}}{\varepsilon_{el,cogen}} \tag{2.22}$$

The bio-gas cogeneration engine has an operating value of 0.16 $\frac{kWh-Eq.}{kWh}$, the diesel cogeneration engine of around 0.47 $\frac{kWh-Eq.}{kWh}$. Both are implemented as a heat driven cogeneration, where the heat demand defines the operation of the unit over the heat cascade:

Additional constraints limit the electricity production to the total annual electricity demand of the concerned buildings $\dot{E}_{el,demand}$. In every time step, the unit u can produce the quantity $\boldsymbol{f_{u,p,t}} \cdot \dot{E}^{-}_{el,u}$ of electricity, where f is the unit's multiplication factor defined through the heat cascade and $\dot{E}_{el,u}$ the amount of electricity produced for each unit of heat.

$$\sum_{u=1}^{nu} \sum_{p=1}^{np} \sum_{t=1}^{nt} \boldsymbol{f_{u,p,t}} \dot{E}^{-}_{el,u} \leq \dot{E}_{el,demand} \tag{2.23}$$

The cogeneration unit can therefore only produce for the local electricity demand avoiding solutions where a cogeneration based system exists only for electricity export to the grid as a infinite "virtual" client outside of the given system. The additional constraints (2.24) and (2.25) are introduced to distinguish between the import and export price of electricity. Additional electricity can be imported ($\dot{\boldsymbol{E}}^{+}_{\boldsymbol{el,p,t}}$) or exported ($\dot{\boldsymbol{E}}^{-}_{\boldsymbol{el,p,t}}$).

$$\forall p, t:$$

$$\sum_{u=1}^{nu} \boldsymbol{f_{u,p,t}} \dot{E}^{-}_{el,u} + \dot{\boldsymbol{E}}^{+}_{\boldsymbol{el,p,t}} - \dot{\boldsymbol{E}}^{-}_{\boldsymbol{el,p,t}} \geq 0 \tag{2.24}$$

$$\dot{\boldsymbol{E}}^{+}_{\boldsymbol{el,p,t}} \geq 0 \quad \dot{\boldsymbol{E}}^{-}_{\boldsymbol{el,p,t}} \geq 0 \tag{2.25}$$

2.5.4 Technology Cost and CExD Overview

Summarizing the technology data for the use within the optimization model, two kind of data are collected for each technology. Based on the ecoinvent database, the CExD value for each technology are collected at all available sizes. For the here stated technologies only one size is currently available. All CExD values in Table 2.1 are extrapolated according to [*Gerber et al.*, 2011] as shown in Section 2.3.1 except the gas boiler. Two data sets at different sizes exist for the gas boiler. For the oil boiler, ecoinvent states that in terms of production CExD, gas and oil boiler are to be treated as similar equipment.

The cost data is based on the CREM's CostDBCREM [*Poumadère et al.*, 2015] data base. The data contains detailed costs for the investment based on realized projects in Switzerland, such as the civil engineering and the material costs. The operating costs contain the resources cost and the maintenance costs, also based on realized projects. Due to confidentiality, no detailed cost break down can be provided. Only aggregated values can be shown. Just like ecoinvent, the cost database does not contain oil boilers neither, therefore the investment cost for gas boilers are used for them. All functions are non-linear. They are approximated and represented in a peace-wise linear function that does not add an additional error to the already existing uncertainty of the data set. Throughout this work, the gas boiler is used as an example, its investment cost function is defined as: $379.42 \cdot \boldsymbol{s_u^{-0.437}}\ CHF/kW$.

Table 2.1 – Non renewable CExD *Dones et al.* [2007] and cost data based on CostDBCREM [*Poumadère et al.*, 2015] for utilities for a reference size of 1 MW: all CExD values are extrapolations except the gas boiler and in ecoinvent gas and oil boiler are equal

Utility	CExD		Costs		Efficiency	
	Operating $\frac{kWh-Eq.}{kWh}$	Production $\frac{kWh-Eq.}{kW}$	Operating $\frac{CHF}{kWh}$	Investment $\frac{kCHF}{kW}$	ε_{th} –	ε_{el} –
Gas Boiler	1.14	47	0.14	500	0.98	–
Wood Chips Boiler	0.1-0.2	242	0.07-0.12	1500	0.85	–
Heat Pump	$0-3.0/COP$	123	$0.25/COP$	746	3.5	–
Co-generation (Biogas) Engine	0.16-0.47	544	0.24	6000	0.55	0.35

3 From Data to Heat Demand of Urban Areas

The optimization model requires input data for the design of an urban energy system. Characterizing the heat demand is the first step. The heat demand of an urban area can be defined through available data. It is the starting point to a typical bottom up problem for the design of an urban energy system: For the sizing of an energy system, the energy demand needs to be known. For an urban energy system, the buildings represent the demand: how can the demand be estimated? Looking at the available data, the demand can only be estimated with different approaches.

This chapter gives an overview over common modeling approaches first. Then the data for the models is discussed, pointing towards the available data sources in Switzerland. A validation study shows the difficulty to represent the reality with models.

3.1 Modeling Approaches

Defining the energy demand of any given building can be done based on different methods. Based on the amount of knowledge about a building, models of different precision can be used. This chapter will examine the data collection and structure phase in order to systematically feed models with the necessary input data.

In modeling, *Swan and Ugursal* [2009] distinguish between the two general approaches: top-down and bottom-up. The top-down approaches do not need much data and the data can be given in aggregated form. They lack the possibility to model advances in technologies or individual end-use of energy. The bottom-up modeling can either use an engineering approach such as building simulation models that represent physically a building with a certain level of detail or statistical models fitted on available data. *Foucquier et al.* [2013] take a different angle by classifying approaches according to the three methods: physical (white-box), statistical (black-box) and hybrid (grey-box). The grey-box model try to take advantage of the both approaches, the statistical one and the physical one. It is the most recent approach developed of the three.

During this thesis, three different models with different input data requirements are used. First and only in this chapter, the energy-signature is used. It is a statistical model that establishes a correlation between the power requirement and the outdoor temperature. *Hammarsten* [1987] explains the energy signature and critically reviews it: he concludes that when moving from yearly towards hourly estimations of the heat load with the energy signature, the method becomes very unreliable. Second, the dynamic building simulation software bSol, a tool with a reduced physical building model, used in [*Page et al.*, 2014] and described by [*Bonvin*, 2004] can calculate the heat requirements. It is not used for the case study of this chapter, but for the case studies with Verbier in Chapters 4, 5, 7 and 8. Third, CitySim [*Kämpf and Robinson*, 2009], a tool for urban energy simulations is used for a detailed urban building simulation. Compared to bSol, CitYSim uses a more detailed physical building, that also considers interactions between buildings especially through its own irradiation model.

Considering the interactions of buildings requires additional information that a building simulation does not automatically considers such as the distance between buildings and their individual horizon using a 3-D model based on the whole scene. CitySim is therefore the tool that needs the most information in order to calculate an energy demand. bSol requires less input data than CitySim on the physical building properties and also ignores the position of the buildings to each other, the individual horizon. The energy signature uses a so-called backwards oriented modeling approach. The energy signature fits the annual consumption using the heating degree days to estimate the losses of the building, that are regrouped in a single term. The power is a linear function of the outdoor temperature: heating power is required when the outdoor temperature falls below the cut of temperature T_{cut} for heating. k_1 represents the slope of the signature or the specific heat losses proportional to the external temperature T_{ext}, k_2 is the y-intercept or the power required at 0 °Celsius.

$$\dot{Q}(t) = \begin{cases} k_1 \cdot T_{ext} + k_2 & \text{if } T_{ext} < T_{cut} \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

3.2 Available Data and its Sources

All previously discussed tools that use one of the discussed modeling approaches in Chapter 3.1 need to be fed with information. For this, yet another set of tools exists that combines available measurements and statistical data systematically on an urban scale. Based on the knowledge gained from studying Geneva's thermal energy strategy with EnerGIS [*Girardin et al.*, 2010], the tool Planeter [*Blanc et al.*, 2013] was developed. It is an implementation of EnerGis, helping to quickly gather data of a given urban environment through the combination of available annual measurements and statistical data. Both tools use the energy signature. In addition, the project MEU was launched mainly as an energy management platform for bigger communities and their utilities, which can also monitor performance on an annual basis. MEU contains a data model based on the input data for the dynamic building simulation software CitySim [*Perez*, 2014]. Fuel consumption measurements of all available years can

be introduced and shown on a map. The fuel consumptions over several years vary mostly because of different meteorological conditions. Considering the meteorological conditions with the heating heating degree days [*Day*, 2006], allows to compare the consumption of different years with each other.

The previously described tools help to define the energy demand, because they gather, combine and complete the collected data, which will be described in further detail now. A data source is written in *italics*, when it is used during the validation case study.

The following data categories exist: data about the user, geometrical data, physical building data and data about the technical installations. Table 3.1 shows the different data sets that are necessary for the characterization of the energy demand on the building and city level. The table does not claimed to be complete, but contains the key elements needed for the energy demand definition. It shows the wide range of details, that can be available during a project for which real data is used.

Starting with the building's user or inhabitant, the building usage gives a static information about the expected user. Measurements or presence profiles can help to refine the information. On the aggregated level, the total numbers of people working and/or living in the building are often known by the administration. The geometrical data is normally easy to obtain: recent efforts of the Swiss Federal Office of Topology allow to recover very detailed geometrical data of buildings leading to an accurate 3-D model. With the horizon for a given city and a 3-D model of the buildings including all other obstacles such as trees, an individual, precise horizon for each building can be derived from it leading to a precise calculation of solar irradiation. Heated surface per building is a minimum requirement for the heat demand estimation. The physical building data is generally used to create archetype buildings. The construction year refers to the norms in place at the time and thus to the material used to provide an information about the potential energy demand. For building simulation, the exact physical building characteristics help to find a precise result. The minimal information about the technical installations of the energy system are the heating system type and the annual fuel consumption for the archetype buildings. The more additional details are known, the finer the distinction between fuel consumption and the building's energy demand can be made.

On an urban level, often the energy system consumption of the whole area in an aggregated form is known. Through the advances in image and radar data processing, the geometrical data can be considered as precise. The remaining categories, the user's behavior and the physical building characteristics, remain a large source of error. Often, only the minimal set is available.

The infrastructure options are a key information for the planning of future infrastructure while considering the existing ones. Infrastructure includes storage installations, i.e. for the use of renewable energy integration or/and peak shaving options. If possible the energy bill should be available per service, either for heating, hot water production, cooling or electricity.

Table 3.1 – Required information to define the energy demand of a building in pre-dimensioning study. The minimum information is contained in the maximum data set

	spatial levels		
	building data requirement:		neighborhood to city
	maximal	minimal	
data about the user	presence, behavior (hot water needs)	building usage	number of workplaces and inhabitants
geometrical data	2.5 to 3-D building model, (windows') orientation, horizon, shadows of other obstacles	heated surface	surface, horizon
physical building data	U and G-values of all parts, air tightness, type and year of refurbishment	construction year	building stock studies and programs
energy system	gains of appliances, technical equipment installed, distribution temperature, radiator type, power of heating system, indoor temperature	heating system type, annual fuel consumptions	aggregated (annual) fuel consumptions and meteorological data over the same period, infrastructure

Detailed information about the different energy services are in particular important if only electricity is used to deliver all services, because it is difficult to separate the consumption per service without additional knowledge.

With the available fuel consumption, the demand can be characterized when the efficiency of the energy supply system is known. Table 3.2 shows the minimal and maximal information needed for this analysis of energy supply systems at urban scale. For an individual building in an urban energy study, at least the efficiency of the technical installations and their size should be known.

The data for the use within the described models and the needed data sets can be found in different sources. One basic source are norms, especially building construction norms, which are evolving over the years. However, the norms are not always respected, neither are all the buildings maintained to ensure a constant fuel consumption while aging during the same meteorological conditions.

For and in Switzerland, different data sources for the different levels exist: The Swiss Statistical Office provides the *Federal Building Registry* that provides the minimal data set of Table 3.1. More and more efforts are made on obtaining a nation wide 3-D model with all buildings. On cantonal level the same information exists with a different structure in the cantonal building registry. Some communities and cities, especially the ones that own their own utilities, are the ones with supplementary information: annual fuel consumption data can be found based on

Table 3.2 – Required information level to define the energy supply side for a given urban area. The minimum information is contained in the maximum data set.

	spatial levels		
	building		neighborhood to city
	maximal data	minimal data	
technical installation	biomass, wind and solar potential, (seasonal) storage installation	efficiency of (existing) installation(s), size of installations	potential of biomass, wind, solar, (seasonal) storage installation, heat pump , available power on existing infrastructure, inventory of thermal discharges, source(s) of electricity, political goals

the billing of it. On the technical and physical level, the construction permit can sometimes be accessed to gain information that completes the maximal data set. Often however, this information can not be accessed directly due to data protection laws. *Interviews* with the cities or utility companies, for example the construction permit department, can complete the data sets. The interviews with some of the utility companies revealed that often a low amount of data is recorded (sometimes still manually) on how their installations are currently running for the monitoring purposes. Others have already automated systems set up that give alarms, when the efficiency is dropping below a certain threshold.

The chimney sweeper's data can help for all installations requiring chimneys: the age of the installation and the size is often noted. For solar installations, PV or thermal collectors, image treatment of areal photos can help to get an idea of the size. Efficiency remains however unknown and also information about the combination with a storage of any kind and size.

The *Federal Commercial Registry* shows the branch and the number of equivalent full time jobs and the type of industry or businesses. Obtaining systematically more information about the enterprises and their fuel or electricity consumption remains difficult.

In addition to the named sources, historical data sets might be available such as photos, studies oriented towards fuel consumption, best practice construction guides from a city, cantonal or federal office. Exploiting them can be useful, however is also very time consuming.

For the use of the heat cascade, *Kemp* [Chapter 9.6 2007, pp 369] suggests to create a list of all services requiring heating or cooling. In urban environments, these are space heating, domestic hot water production and (free) cooling. Sometimes refrigeration is also needed. Each service should be defined with a temperature-power profile as a function of time. Available resources are classified in the same way. However, no data set on the urban level exists that systematically contains this information per building. *Girardin* [2012] derived the temperature from a combination of factors, mainly the building's heating system and the buildings age.

3.3 Data Validation Case Study

For this study, all available data sets discussed in the previous section at urban scale for the 3 cities Neuchâtel, La Chaux-de-Fonds and Martigny are used. The data model structuring the available data sets is based on an energy flow graph for the planning, monitoring and evaluation of urban energy systems. It contains a spatial dimension for map creation and a temporal dimension for monitoring over all available years.

Based on annual energy balances, the oriented graph traces all energy flows within the boundaries of the chosen district that concern the following three services: heating, hot water production and electricity. From source nodes that represent resources, the energy flows towards energy distribution or conversion system nodes. Each distribution network node contains information about the physical connections and additional information about the type of contract. From there on, further edges can either connect to network nodes, energy conversion system nodes or directly to sinks that represent the building's energy demand. Each node has one or many input flow edges and output flow edges. Each flow-edge contains a fraction equal to the amount of the start and stop node's total transported energy. A network node has a loss factor, therefore the input does not equal the output. Energy conversion system nodes rely on a separate set of simple energy conversion models for different technologies. When measurements and simulation values exist for the same node, the measurement value is used. *Perez* [2014, Chapter 3] describes the conceptual data framework in detail.

With the help of this graph, about 50 percent of the final energy use within the system's boundaries can be traced. After the integration of the above stated data sets into the data base of MEU, measurements and simulations can be systematically compared.

First, CitySim, the dynamic urban scale building simulator, estimates the heating needs without the hot water requirements based on the geometrical and physical building data. The only very accurate data set is the building footprint from the land registry office. With the Swiss norm [*Swiss Society of Engineers and Architects*, 2006], which contains the standard use of buildings and statistical presence data, presence profiles for each building type are created and used during the simulation. The physical building data is already integrated into CitySim, relying on past building studies made.

The efficiency of a given technology in a building is either contained as a measured value in the data or estimated based on the Swiss norm [*Swiss Society of Engineers and Architects*, 2009a], which provides a conservative annual efficiency estimate. Most of the data concerning the energy conversion systems is based on default values defined for each technology type. During this case study, only technical data from central district heating facilities or public buildings could be easily accessed. Therefore mainly the default values of Table 3.3 are used. No data is available concerning solar thermal nor photovoltaics production, wood use in stoves or installed energy storage. Because the data about renewable energy or storage systems is missing, only rarely systems with multiple energy conversion systems, multi-energy system, exist in the case study. If the energy produced with renewable resources is locally consumed,

the overall energy bill will be reduced. As a consequence, buildings might be classified better than they should be. In addition, a large difference between the estimated energy demand via simulation might appear. In almost all cases, a combination of a wood boiler or stove with another boiler lead to a much lower energy bill and are responsible for an inexplicable gap between energy demand and energy delivered.

Table 3.3 – First Law annual technology efficiency

Technology	Efficiency
Wood Boiler	0.65
Wood Pellets Boiler	0.70
Gas and Oil Boiler (condensing)	0.85
Gas and Oil Boiler	0.80
Electrical Boiler	0.93
Heat Exchanger	0.93
Heat Pump	3.4 (COP)

Measurements in form of the annual energy consumption are almost citywide available. Most measurements are available as a sum of heating and hot water demand for residential buildings. With the default values stated in Table 3.4, the fraction of the heating requirement can be estimated and subtracted from the other services combined in the annual energy consumption. The norm provides a guideline for the separation of the services. They do not provide information about a combination of technologies that could offer this service efficiently.

Measurements, namely the annual energy consumption Q, and simulations of the annual heating demand Q_{sd} can be compared to calculate the percentage error (PE) with Equation (3.2). This allows to compare each service for each building in terms of annual energy demand. Equation (3.2) can be applied to each building with an individual measured heating fuel consumption Q that does not contain the hot water production: the annual energy consumption Q is multiplied with the efficiency $\varepsilon_{technology}$ of the installed technology to obtain the final energy demand. The simulated heating demand by CitySim Q_{sd} represents the estimation of it.

$$PE = \frac{Q_{sd} - Q \cdot \varepsilon_{technology}}{Q \cdot \varepsilon_{technology}} \cdot 100\% \tag{3.2}$$

Equation (3.4) is used when a service needs to be excluded such as the hot water use Q_{hw}. The hot water use is estimated with the Swiss norm [*Swiss Society of Engineers and Architects*,

Table 3.4 – Default values used based on [*Swiss Society of Engineers and Architects*, 2006] to derive heating requirements from measurements which combine services such as electricity and/or hot water need with heating requirements for different building types

Category	Electricity Need $[MJ/(m^2an)]$	Hot Water Need $[MJ/(m^2an)]$	T min set point $[°C]$
apartment	100	75	21
building individual home	80	50	21
administrative	80	25	21
schools	40	25	21
sales	120	25	21
restauration	120	200	21
meeting venues	60	50	21
hospitals	100	100	23
industry	60	25	18
stores	20	5	18
sports installations	20	300	18
indoor swimming pools	200	300	28

2009b] which is providing specific annual hot water demands q_{hw} that only needs to be multiplied by the heated surface A_h of building.

$$f_{heating} = \frac{Q - Q_{hw}}{Q} \qquad Q_{hw} = q_{hw} * A_h \tag{3.3}$$

$$PE = \frac{Q_{sd} - Q \cdot \varepsilon_{technology} \cdot f_{heating}}{Q \cdot \varepsilon_{technology} \cdot f_{heating}} \cdot 100\% \tag{3.4}$$

When the technology is connected to several buildings, such as in a district heating system, only the building's fraction part $f_{Building}$ is taken into account (Equation (3.5)).

$$PE = \frac{Q_{sd} - Q \cdot \varepsilon_{technology} \cdot f_{heating} \cdot f_{Building}}{Q \cdot \varepsilon_{technology} \cdot f_{heating} \cdot f_{Building}} \cdot 100\% \tag{3.5}$$

The fraction $f_{Building}$ is the proportion of the heat demand for the given building over the sum of all connected buildings.

$$f_{Building} = \frac{\text{Building Heat Demand}}{\sum \text{Building Heat Demands}} \tag{3.6}$$

The building's heating demand is also estimated with the energy signature using the same data and the meteorological weather data for the city of La Chaux-de-Fonds, Switzerland. The calibration of the energy signature model is shown in detail in *Girardin et al.* [2010]: the heating cut-off temperature is estimated based on experience or measurements allowing to integrate the annual energy consumption with the outdoor temperature.

3.3.1 Results

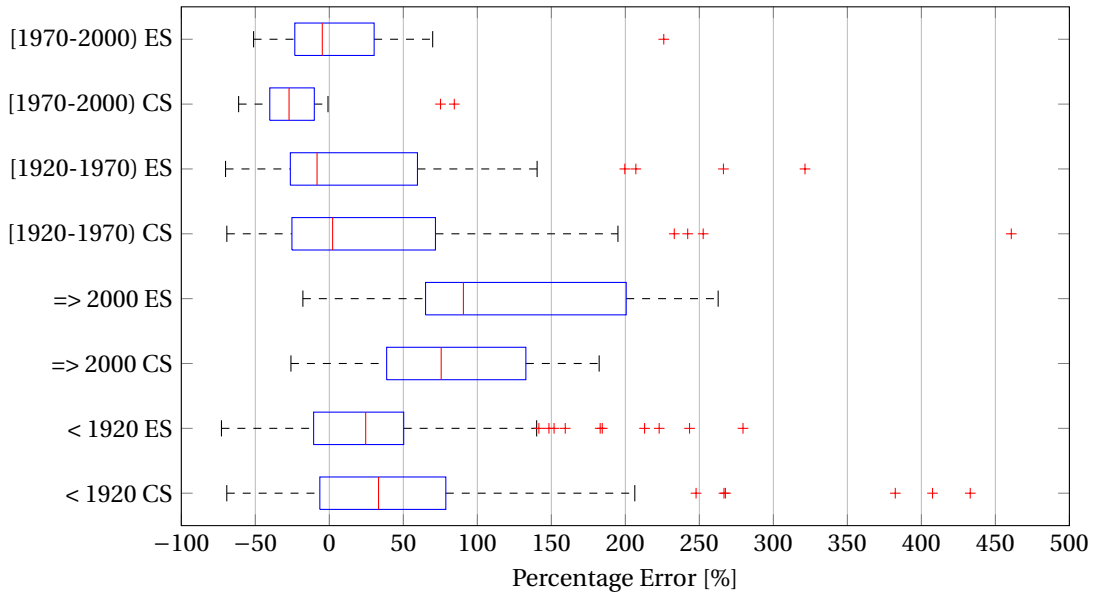


Figure 3.1 – La Chaux-de-Fonds: boxplot with outlier, marked through +, and mean, marked through vertical the red line in the box, of the error between measurement and estimation of the energy demand with either the energy signature (ES) or CitySim (CS) for different construction periods

When studying Figure 3.1, both methods generally over-estimate the energy demand. Very old buildings as well as very new residential buildings seem to be difficult to estimate correctly as they are highly overestimated. New buildings represent only five percent of all available buildings, it is therefore difficult to evaluate these buildings in more detail. Residential buildings constructed between 1920 to 2000 are in average correctly simulated, however high standard deviations indicate that more knowledge of the building could help.

Comparing these results to other cities about provided data, show that high standard deviations and high mean errors. More work on collecting the individual building data is needed. For the moment, it is highly likely that the error between the real and the estimated demand for the buildings that do not have been verified individually, show the same differences. Therefore it is very likely to overestimate the total energy demand by about at least 20% with the energy signature and 50% with the physical building simulation shown in table 3.5 for each city. No systematic error was identified, so that the source leading to the gap between estimation and simulation needs to be assessed with an individual building per building approach. If the

Table 3.5 – Results across all cities for which measured data is available

City	Number of measurements	Mean consumption	PE error ± standard deviation	
			CitySim	Energy signature
La Chaux-de-Fonds	237 of 307	$126 \frac{kWh}{year*m^2}$	50% ± 131	33% ± 75
Neuchatel	104 of 152	$166 \frac{kWh}{year*m^2}$	74% ± 97	19% ± 54
Martigny	69 of 76	$78 \frac{kWh}{year*m^2}$	103% ± 88	20% ± 48

physical building data base could be taken as a correct reference, MEU can systematically identify faulty measurements or performance changes of buildings.

After the previous study, *Perez et al.* [2013] repeated the same effort on a bigger sample while using a different error measure, the logarithmic discrepancy factor (DF):

$$f_2 = log_2\left(\frac{\text{simulated value}}{\text{adapted or measured value}}\right) \tag{3.7}$$

The logarithmic error has the advantage of being a symmetrical error, 0 means it is a perfect match. A result of -1 corresponds to a simulated value twice smaller than the adapted value, 1 means the opposite, a simulated value that is twice as big. The authors spent a considerable amount of time searching for different default values than given by the stated norms to find a better predictions for the dynamic building simulation and efficiences of the energy conversion technologies. The results in Figure 3.2 shows that the situation is clearly improved compared to the original version which only used default values.

The variance can not be improved in most cases, however the average can be centered closer to 0, where it should be when the simulation correctly estimates the annual energy demand. According to *Perez et al.* [2013] the most influential factor remains the behavior, which can not be addressed within this framework. On an urban level, average buildings can be calculated, whereas the individual building remain much less reliable. Additional energy systems such as solar thermal systems or wood stoves were not studied, because no detailed information about them exist concerning their existence, size and usage.

Nevertheless this approach prioritizes the order of buildings to study in more detail by combining the error measure and the total energy demand of the buildings. The consumers with a high energy demand that are not well understood yet should be studied first.

During the study, *Perez et al.* [2013] notes that each community has individual historical architectural styles of buildings. Default data sets therefore cannot be transferred from one community to another even though the geographical distance between them is with only between 20 to 150 km rather small. The different environmental conditions, La Chaux-de-Fonds is situated at $1000m$ altitude in the Jura mountain chain compared to Neuchâtel facing south east on a lake at only $430-500m$ altitude and Martigny ($470m$) in the Rhone valley surrounded by the Swiss Alps, can explain this.

In addition, this study does not resolve the problem for individual buildings: it introduces for all model input parameters a most likely value, therefore still high variations for individual buildings are possible. Besides the high variation for an individual building, a second case is possible: One might find no difference between an estimation and a measurement for an annual energy consumption, but still not represent the reality, because several sets of input parameters for the simulation can lead to the same result. Errors in the input parameters can cancel each other out, when only looking at the annual energy consumption. For example, in the case of a partial energetic retrofit, with the given data during this case study, it will be difficult to distinguish between the heat losses through windows, walls and ventilation. A visit of the building is still highly recommended to ensure that the best measure is chosen.

3.4 Conclusion

The results show that the model and the input data should both work on the same level of detail. Here two models are compared: a detailed physical building model and a statistical model.

On one side, the draw-back of a statistical model, the energy signature model, are evident when going into detail, e.g. when using more detailed physical building data, when trying to calculate hourly values or when the building's thermal capacity should be addressed. On the other side, very detailed models such as CitySim need a large quantity and high quality of input data to explain annual fuel consumption. The results strongly depend on the presence profile and a fixed internal temperature for the building besides the information about the physical building properties. All information are rarely available at high precision. When taking only default values, hardly no difference between the energy signature and detailed building simulation appears in terms of result's precision. Complex physical models can be calibrated to fit to the data better based on additional physical information as shown during the work of *Perez et al.* [2013]. The energy signature can only be calibrated, when the building's correlation between outdoor temperature and heat load remains linear. Either time is spend on completing the input data on an urban level at the necessary detail or simpler models should be considered.

Until now, most of the time only annual measurements exist. Therefore, the energy demand can only be expressed as a function of the time with a more precise resolution through the help of more detailed models. From the available models, only the physical models can predict based on the annual fuel consumption values an estimated hourly heat demand profile. However the variation at urban scale remains large. Even though the data collection is very detailed, the nominal power of installations cannot be addressed because of a lack of data: When the nominal power of individual technologies remains unknown, it is difficult to address the difference between the annual efficiency and the operating efficiency to provide a more detailed energy system design suggestions. Frequent on-off switching of a technology is diagnosed as a low annual efficiency. Thermal energy storage integration into the energy system

might already be able to address this issue, because it might reduce on-off switching. However no information about existing storage could be found neither whether the given building has the space to install one. For the successful integration from an exergetic or economic point of view, the time dependent power-temperature profiles of the studied buildings should be known.

After the heat demand characterization, when studying the available norms one can see an additional problem: No information about the sizing of multiple technology systems is given in the existing norms. For example, when integrating solar energy, energy storage and an additional technology ensuring the desired services, need to be sized.

Finally, when looking back on this study and other similar studies at urban scale, it can be concluded that often around 50% of the project's time is spend on data collection, structuring, understanding, verification and searching for complementary information. On one side, for research in field of urban energy design using real data sets allows closing gaps in existing methodologies. On the other side, it is also a lot of time spent on preliminary work not directly related to developing urban energy design methodologies.

Based on the findings of this chapter, in the next ones the following questions are addressed:

- Based on building simulation, how can a time dependent power profile be integrated into the system design while keeping a MILP problem solvable?
- How can the design of complex energy systems be addressed including thermal storage and solar thermal panels?
- Finally, how can uncertainty be integrated in the design process?

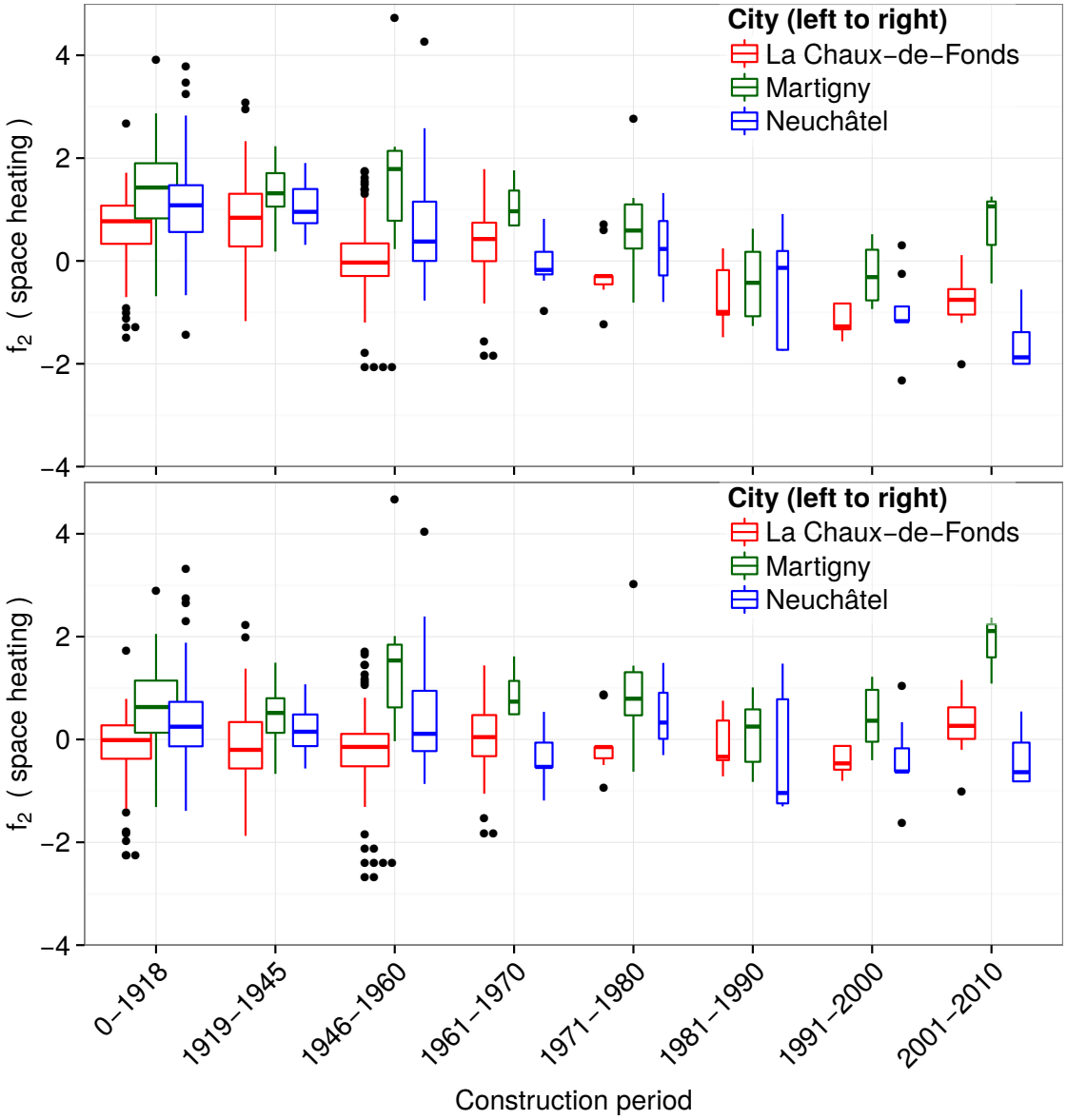


Figure 3.2 – Symmetric discrepancy factor La Chaux-de-Fonds, Martigny and Neuchatel from [*Perez et al.*, 2013]. The boxplots show the error between measurement and CitySim simulations for different construction periods. Figures from [*Perez et al.*, 2013]

4 Reduction of the Input Data Set

> Humans are good at discerning subtle patterns that are really there, but equally so at imagining them when they are altogether absent.

Carl Sagan

The motivation of this chapter is described by the following observation: when sizing equipment of an energy system, practitioners identify extreme conditions in order to size their energy systems (that contains often only one technology), despite the fact that these conditions occure only several hours during a year. A security margin is often added to take into consideration the aging of buildings and equipments. Furthermore, equipments are only available in discrete sizes, meaning that the next available higher capacity is typically chosen. As an advance, normal operating conditions are slowly starting to be taken into consideration. For sizing more and more complex energy system, all conditions from extreme, whether maximum or minimum load, to average conditions should be represented as input for the sizing ensuring that the system can respond to them accordingly to the design ideas.

Commonly during the energy system design for buildings, the energy demand is defined for a typical year. Often already the demand is increased: *Porges* [2001, p.91] states that the margin added to the heat loss calculations are between zero to twenty percent for buildings (up to 50 % if not in daily use) that have exposed north surfaces. When the equipment is sized to satisfy the demand, an additional safety margin of 10 to 15 % is included [*Porges*, 2001, p.119]. Because the buildings represent the energy demand in urban areas, the overestimate is highly likely to be found in the urban energy system. Ensuring that not simply all safety margins are added, so that the actual system might be over sized, is a responsibility of the energy design process and implies to understand all steps of the design process.

Simplifying the 25 years of lifetime to an annual demand simplifies the approach. But it neglects that the actual problem is a design problem over 25 years (or the lifetime of the energy

system that can also be longer). However, the hypothesis of constant annual demand for each year thanks to the typical year allows to only calculate one year reducing significantly the overall problem size. The cyclic climatic conditions during the year, which are the main driving force for the energy system, further encourage to use only one year. Using a typical year ensures representing also the extreme events in their frequency of appearance.

This typical year contains a succession of typical days. Given the mathematical and computational complexity of the problem, a trade-off must be made between the amount of data and the expected quality of the results. Through this approach, we are able to reduce the overall amount of data while preserving its key elements in order to ensure the proper sizing of equipment while significantly reducing the amount of data to be handled by in the model. This is particular important for the use in an MILP problem, where integer variables are created as a function of the number of periods. Then reducing the number of periods has a direct impact on the problem size and also impacts the ability of being able to resolve it.

The here presented approach identifies representative days that can be used for sizing of energy systems. Choosing representative days for the demand and supply when both vary over time, impact the energy system design. The size of the input data set has an important impact on the solving time (or solvability) of any optimization problem. Therefore redundant data should be reduced to the minimum required data for the task to be performed. As the input data for the design of urban energy systems contains spatial and temporal data sets, different levels of reduction can be applied to obtain a smaller data set without changing the overall result of the optimization.

Firstly, the hourly annual heat demand is reduced to several operating periods representing the year with a deterministic approach. They can be reduced again if necessary. Secondly, the composite curves of a group of buildings can be reduced limiting the amount of individual streams considered in the heat cascade through aggregation on a district level without loosing accuracy.

4.1 State of the Art

The goal of clustering is the reduction of input data, a numerical way of finding groups in data. *Tan et al.* [2005, Chapter 8] states that clustering techniques can reduce the number of data points significantly. Different clustering methods exist. The challenge is the identification of the right number of clusters k, that is from the computationally point of view while maintaining its critical information.

The technique was first proposed by *Lloyd* [1982] in 1957 even though he published it only in 1982. The goal of clustering is an objective and stable classification of data into groups or clusters [*Everitt et al.*, 2011d, p.4]. A cluster is defined by its internal cohesion, or homogeneity, and external isolation, or separation [*Everitt et al.*, 2011d, p.7]. In the field of artificial intelli-

gence, these techniques are used for machine learning. In other fields of research such as a bio-medical research [*Xu and Wunsch*, 2010], they are also heavily used.

In the field of energy system engineering, *Ortiga et al.* [2011] uses k-means to demonstrate the possibility of reducing an annual hourly data to several key days. k-means has the disadvantage of being a heuristic approach that starts by randomly picking a k days out of the data and grouping around them. Therefore, the clustering needs to be repeated often enough to ensure that not only a local minimum is found but a global one. In addition, k-means creates new data points, because a cluster is characterized through the mean value of all points within the cluster. The authors use the cumulative energy demand curve to judge whether the representation is meaningful. Peak load periods are often added manually because they do not appear automatically within the mean values representing the data set. *Rager et al.* [2013a] follow the same approach.

Fazlollahi et al. [2014b] builds up on this approach formalizing indicators that describe the quality of the clustering. Most of the indicators are meaningful in the field of clustering, however they are difficult to use for the energy system designer who is not that exposed to clustering. For their first case study, two strongly correlated variables, outdoor temperature and the heat demand calculated with the energy signature, are clustered. Testing for correlation would probably show that it is efficient to work with one of the two variables.

Domínguez-Muñoz et al. [2011] works with k-medoids, a method from the same clustering family. The major difference between k-medoids and k-means is that k-medoids selects cluster centers from within the cluster points instead of using an average. They start by selecting manually the peak load days and cluster the remaining days. With the help of scaling factors on the cluster days, they ensure the annual energy balance remains correct.

Additional approaches using a genetic algorithm for the same problem are proposed by *Marechal and Kalitventzeff* [2003] and *Bungener et al.* [2013]. For the case studies shown, they provide good results, with the drawback of depending on a certain number of runs from the genetic algorithm to ensure convergence to an acceptable solution.

The literature review reveals that the clustering methods used rely on repetition to ensure finding an acceptable or even optimal solution of the clustering. It would be interesting to study whether a deterministic method can get to the same result. All approaches add the extreme periods manually to the clustering, are there ways to include it directly? Also, the combination of indicators judging the quality of the clustering could be improved to a more relevant combination for system design: it should rely on the combination of numerical and graphical indicators at the same time. In addition, the number of day alternation in a sequence is not addressed in the literature, because in the shown applications no storage unit with temperature levels exist that interconnects the different typical days.

4.2 Deterministic Selection of typical operating Periods

The clusters are judged with a set of measures ensuring that only meaningful clusters are retained to guarantee a certain quality. This ensures to reduce the data points while keeping a strong fit between the data and the created clusters.

In almost all clustering application, the number of groups or clusters that should be identified in the data is an input variable of the clustering algorithm. This bears the risk that clusters are created that may not represent a physical reality because the clustering algorithms always provide an answer with the number of groups requested. The algorithms will output the number of groups requested, even, if the request number of groups does not make sense.

Based on *Everitt et al.* [2011b, p.261-262], a typical cluster analysis contains the following steps:

1. identification of variables (and with correlation tests between them)
2. choice of clustering method and algorithm
3. standardization of variables
4. choice of proximity measure
5. identification of the best number of clusters
6. interpretation of the results

These mentioned steps are developed in the following sections. There is also an interest in defining outlier points in order to not include them in an optimization, though this is not done in this work.

4.2.1 Variables

The input set considers two key values for the pre-design of energy systems:

1. Heat load (or outside temperature as they are highly correlated),
2. Production of heat with solar thermal panels (or irradiation on a given building).

More variables can be added, if the (pre-)design model can use more information. Examples can be the electricity price or demand, if time dependent functions or measurements exist.

Furthermore, it should be assured that the variables are tested for correlation. If the variables are correlated, the cluster analysis has a high chance of following this pattern. Therefore, in some application it might be possible to eliminate variables from the analysis. Pearson's product-moment coefficient or Spearman's rank correlation coefficient offer such tests based on the covariance.

Applying this to energy system design means that either heat load or outside temperature is chosen as a value to be clustered. Figure 4.1 shows that they are inversely proportionally linked: When it is very cold outside, the heat load to be delivered to a building is increasing.

Figure 4.1 is based on the dynamic simulation of a multi-family house by *Dorer et al.* [2005] for the Norm 380/1 [*Swiss Society of Engineers and Architects*, 2009a]. The data has a Pearson Coefficient of −0.90 and an adjusted r^2 values of 0.81. The test of correlation shows that one variable can be used to explain the other one, when accepting the threshold found between the outdoor temperature and the heating load. If the same observations are made for passive

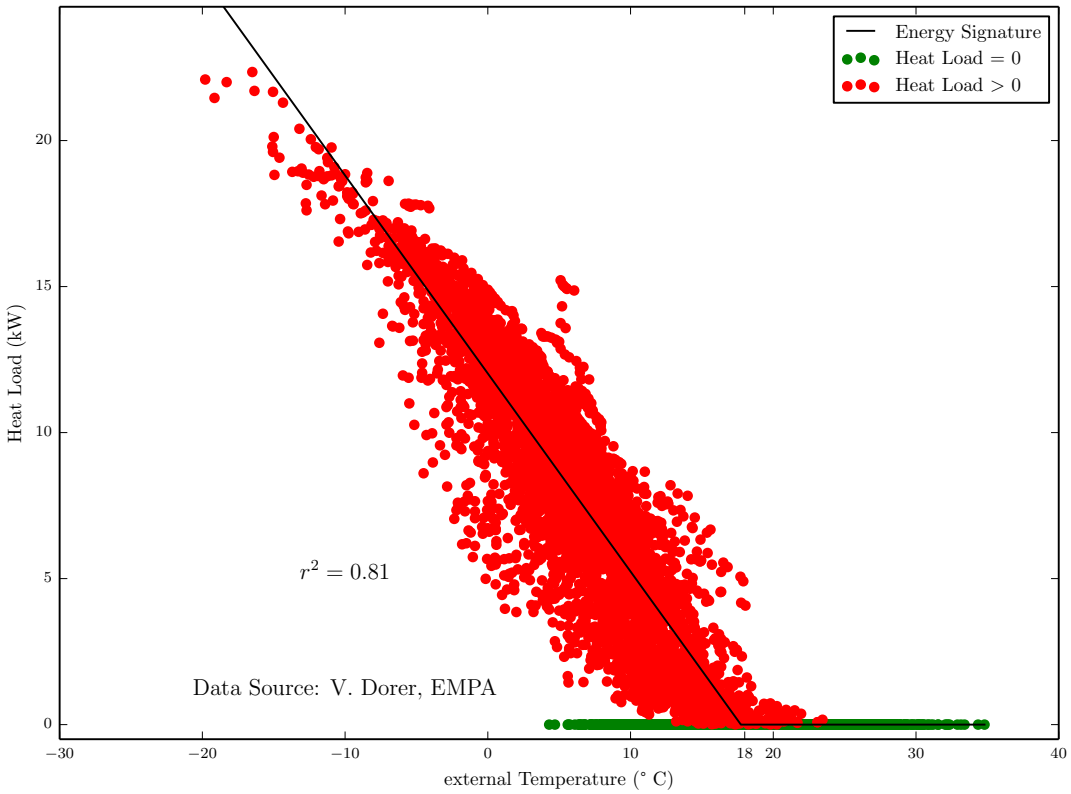


Figure 4.1 – Energy signature of multi-family-house used for the Norm *Swiss Society of Engineers and Architects* [2009a] based on a TRYNSIS-model of *Dorer et al.* [2005]

house buildings, the relation between outside temperature and heat load show are lower value of correlation.

Considering multiple variables, so called multi-dimensional data, often not all variables are meaningful. Generally speaking with more variables the problem tends to get increasingly sparse. The distance or dissimilarity measure such as the Euclidean distance becomes less important as the volume between points in a multi-dimensional space increases a lot more[*Tan et al.*, 2005, p.572]:

> “Consider that as the number of points increases, the volume increases rapidly, and unless the number of (data) points grows exponentially with the number of dimensions, the density tends to 0." ”

Therefore keeping the number of dimensions as low as possible helps limiting this problem and ensures a clearer interpretation of the clustering results.

The data needs to be arranged in rows representing the independent observations and columns representing the variables. 24 hours are chosen to represent the daily variation of both dimensions. It also represents the biological cycle of the building's users. In other applications, a different time interval may be chosen to represent the particularities of the problem. The number of days of a year represent the number of N independent observations. The underlying hypothesis is that the periods chosen are independent from each other. Each hour has two variables, heat production from a solar panel and the heat load from the chosen urban environment. Therefore a day contains 48 variables, 24 for each dimension.

Considering two dimensions, one for the hourly heat load q and one for the hourly solar irradiation i with 8760 values each, the individual values are arranged into the clustering matrix:

$$\begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,24} & i_{1,1} & i_{1,2} & \cdots & i_{1,24} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,24} & i_{2,1} & i_{2,2} & \cdots & i_{2,24} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{365,1} & q_{365,2} & \cdots & q_{365,24} & i_{365,1} & i_{365,2} & \cdots & i_{365,24} \end{pmatrix}.$$

4.2.2 Clustering Algorithm Choice

Generally, the techniques can be divided into supervised and unsupervised (machine) learning. As no result is known before the reduction, only unsupervised learning or so called clustering can be used. Within this approach, the majority of time is therefore spent on the justification of a solution found thanks to clustering techniques.

Everitt et al. [2011c] provides an overview of the field, while *Xu and Wunsch* [2010] review common techniques used in biomedical research and compare them. The field is large, the number of publications in the sciencedirect data base using the term "cluster analysis" has been constantly over 15 000 over the last 20 years. *Xu and Wunsch* [2010] review 35 clustering algorithm families and compares them under different criteria such as scalability, robustness and user-dependent cluster number k as input. Each of these families has numerous different implementations that can lead to different results.

The clustering field can be further divided into partitional or hierarchical clustering. Partitioning divides n objects in k classes. Hierarchical clustering goes a step further: It also considers hierarchically-nested sets of such partitions[*A. D. Gordon*, 1999, p.69] showing the whole tree, e.g. an object belonging to a species, a genus, a family instead of only showing the species.

Other methods exist but are less frequently used and can be found in *Everitt et al.* [2011f] such as density based methods, defining clusters as areas of high density. In this work partitional clustering is used (as the hierarchy between the clusters is of no importance).

Vinod [1969] formulated the clustering algorithm as an optimization problem for example minimizing the within group distance of all clusters. *A. D. Gordon* [1999, p.50] gives a resume over mathematical programming formulations. This approach allows then to systematically compare solutions with a different cluster number k, because for each k the best solution can be calculated. When only considering algorithms that definitely assign a data point to one of the k cluster centers, the three remaining techniques most frequently used distinguish each other mainly in the way they define the cluster's center:

1. k-means clustering [*Lloyd*, 1982] chooses from all data points assigned to one cluster the center by calculating the algorithmic mean. Therefore new data points are created;
2. k-medians clustering takes the median instead of the mean to define the cluster center;
3. k-medoids clustering [*Kaufman and Rousseeuw*, 2005, Chapter 2] chooses data points as the cluster center.

In this work, a specific implementation of the k-medoids is chosen: Partitioning Around Medoids (PAM). k-medoids uses the same approach as k-means, choosing randomly start clusters k. PAM uses a deterministic build phase to choose the starting points. According to *Reynolds et al.* [2006], PAM always provides better results than k-medoids when the build phase has enough time to finish.

The medoid element is an analogy to the centroid, a center element. It is less sensitive to outliers and preserves the original profile as an original point from the chosen data set. The k-medoid is not a heuristic approach, meaning that one execution with the same input data set will always give the same result again.

According to *Kaufman and Rousseeuw* [2005, p.177], "k-medoids is more robust than the error sum of squares in most methods" (such as k-means). *Xu and Wunsch* [2010] confirms this. *R Core Team* [2014] and *Reynolds et al.* [2006] cite, that PAM has the advantage of providing a graphical display, the silhouette plot, as well as numerical value for the silhouette with one dimensional data sets. In *Reynolds et al.* [2006, Section 8.6], the quality of the clustering is accessed showing that in most cases, PAM outperforms other partitional clustering techniques with respect to the silhouette coefficient. The silhouette coefficient combines the measure of cohesion within its cluster and separation to other clusters in 1 coefficient, definition in Section 4.2.6. However, PAM often needs more time to find the result.

For speed considerations, Clustering for Large Applications (CLARA) [*Kaufman and Rousseeuw*, 2005, Chapter 3] can be used when the build phase of PAM [*Kaufman and Rousseeuw*, 2005, Chapter 2] takes too long. This happens for big data sets and high k cluster numbers.

4.2.3 Partitioning around Medoids

Since the k-medoids is not used as frequently as k-means, it is explained briefly.

For this work, the original version of PAM from [*Kaufman and Rousseeuw*, 2005, Chapter 2] in the [*R Core Team*, 2014] package cluster is used. It works within two phases as described in Data Mining Algorithms In R, based on [*R Core Team*, 2014] and [*Kaufman and Rousseeuw*, 2005, Chapter 2]:

> A greedy build phase:
> 1. choose k entities to become the medoids, or in case these entities were provided use them as the medoids;
> 2. calculate the dissimilarity matrix if it was not informed(given);
> 3. assign every entity to its closest medoid;
>
> swap phase:
> 4. for each cluster search if any of the entities of the cluster lower the average dissimilarity coefficient; if it does select the entity that lower the most this coefficient as the medoid for this cluster;
> 5. if at least the medoid from one cluster has changed go to (3), else end the algorithm.

Each data point is a member of one cluster. Each cluster does have at least one member. The run is repeated if the cluster remains empty.

4.2.4 Standardization of a multi-dimensional Data Set

When several data sets are used, the data needs to be standardized. This avoids that one data set determines the cluster choice when its numerical values are bigger than the ones of other sets.

For this purpose, the standardized data can be multiplied with a weight. Depending on the impacts of outliners in the data set, there might be a need to eliminate them. Especially clusters based on the euclidean distances are sensible to this.

For the standardization of data, the data is whitened as proposed by *Steinley* [2004] and *Milligan and Cooper* [1988] over the total range of the variable:

$$Z_1 = \frac{X}{max(X) - min(X)}, \text{or} \tag{4.1}$$

$$Z_2 = \frac{X - min(X)}{max(X) - min(X)}. \tag{4.2}$$

Often the z-score is used, defined as:

$$Z_3 = \frac{X - m}{s} \tag{4.3}$$

assuming that s, the mean absolute deviation of the variable X, is nonzero. m is the mean value of the variable X.

Again, according to *Steinley* [2004] and *Milligan and Cooper* [1988], the methods Z_1 (equation 4.1) and Z_2 (equation 4.2) are in general the most promising one. Compared to the z-score shown in equation 4.3, it avoids depending on a mean value that might be strongly affected by an outlier.

Kaufman and Rousseeuw [2005, p11] continues with the following considerations:

> From a philosophical point of view, standardization does not really solve the problem. Indeed, the choice of measurement units gives rise to relative weights of the variables. Expressing a variable in smaller units will lead to a larger range for that variable, which will then have a large effect on the resulting structure. On the other hand, by standardizing one attempts to give all variables an equal weight, in the hope of achieving objectivity. As such, it may be used by a practitioner who possesses no prior knowledge. However, it may well be that some variables are intrinsically more important than others in a particular application, and then the assignment of weights should be based on subject-matter knowledge (see, e.g., Abrahamowicz, 1985). On the other hand, there have been attempts to devise clustering techniques that are independent of the scale of the variables (Friedman and Rubin, 1967). The proposal of Hardy and Rasson (1982) is to search for a partition that minimizes the total volume of the convex hulls of the clusters. In principle such a method is invariant with respect to linear transformations of the data, but unfortunately no algorithm exists for its implementation (except for an approximation that is restricted to two dimensions). Therefore, the dilemma of standardization appears unavoidable at present and the programs described in this book leave the choice up to the user.

Applying the standardization 4.2 to the same example as in 4.2.1 variables leads to

$$Z_2(q) = \frac{\begin{pmatrix} q_{1,1} \\ q_{1,2} \\ \vdots \\ q_{365,24} \end{pmatrix} - min(q)}{max(q) - min(q)}, \; Z_2(i) = \frac{\begin{pmatrix} i_{1,1} \\ i_{1,2} \\ \vdots \\ i_{365,24} \end{pmatrix} - min(i)}{max(i) - min(i)} :$$

$$\frac{\begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,24} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,24} \\ \vdots & \vdots & \ddots & \vdots \\ q_{365,1} & q_{365,2} & \cdots & q_{365,24} \end{pmatrix} - min(q)}{max(q) - min(q)} \quad \frac{\begin{pmatrix} i_{1,1} & i_{1,2} & \cdots & i_{1,24} \\ i_{2,1} & i_{2,2} & \cdots & i_{2,24} \\ \vdots & \vdots & \ddots & \vdots \\ i_{365,1} & i_{365,2} & \cdots & i_{365,24} \end{pmatrix} - min(i)}{max(i) - min(i)}.$$

4.2.5 Choice of Distance Function

The choice in literature is large [*Everitt et al.*, 2011e, p50] for continuous data:

1. Euclidean Distance (D1)
2. Squared Euclidean Distance
3. City block or Manhattan Distance (D2)
4. Minkowski distance (D3)
5. Maximum Distance (D4)
6. Pearson correlation (D5)
7. Angular separation or cosine similarity (D6)

Other distance functions exist for qualitative (or also called discrete or categorical) data. In this work the squared euclidean distance has been chosen as a dissimilarity measure between clusters. The squared Euclidean distance does not use the physical distance but the squared physical distance.

$$d_{i,j} = [\sum_{k=1}^{p} (x_{ik} - x_{jk})^2] \tag{4.4}$$

It discriminates therefore data points that are further outside more than closer points.

This choice is justified with regard to the sizing of the energy system's equipment: Maximum load needs to be represented. The squared ensures higher penalties within the clustering algorithm as the Euclidean distance does. Therefore, peaks are more likely to be represented as cluster centers compared to the Euclidean distance. When this is not the case, either a sizing hour can be added or the number of clusters k can be increased.

4.2.6 Optimal Cluster Number Choice: Clustering Validity Indicators for the Design of Energy Systems

The goal to design an urban energy system is kept in mind when performing and especially validating the clustering results. The indicators to validate the results are adapted to use a comprehensive choice from of typical clustering quality indicators and a set of indicators relevant to energy system design. Only internal and relative tests are performed and no external validation criteria is used.

The literature yields a lot of general methods that can be applied on data sets in order to define the correct number of clusters to be chosen, [*Milligan and Cooper*, 1985], [*Tibshirani et al.*, 2001], [*Liu et al.*, 2013]. Unfortunately none of them can deliver always the correct result for temporal data sets. Therefore a mix of criteria is applied to identify the best compromise out of the given indicators.

As mentioned, the number of clusters k to be found, needs to be specified before running the algorithm. Therefore, a simple benchmark is implemented, comparing systematically the results for different runs for number of k clusters. The results are compared using the following indicators:

1. Silhouette Coefficient proposed by *Rousseeuw* [1987],
2. mean squared error and the mean absolute error,
3. sequence length,
4. overall and per dimension percentage error,
5. relative minimal and maximum (peak) load percentage per dimension and
6. a qualitative graphical validation.

The Silhouette Coefficient [*Rousseeuw*, 1987] combines the measure of cohesion within its cluster and separation with the other clusters of each individual point i, which is a member of a cluster a. a_i represents the average distance to all other points within its cluster. Then the average distance to all remaining points that are not in the same cluster, b_i is calculated. *Rousseeuw* [1987], *Tan et al.* [2005, p.541] define the silhouette coefficient according to equation 4.5:

$$s_i = \frac{(b_i - a_i)}{max(a_i, b_i)}. \tag{4.5}$$

The range of s_i is between -1 and $+1$. Table 4.1 based on *Kaufman and Rousseeuw* [2005, p.88] proposes an interpretation of the silhouette coefficient. Positive values towards one $+1$ indicate that the average distance to other clusters is bigger than the average distance between the points of the same cluster. Small to negative values for s_i indicates meaningless clusters.

Table 4.1 – Subjective interpretation of the Silhouette Coefficient (SC) proposed by *Kaufman and Rousseeuw* [2005, p.88] for two dimensional data sets

SC	Proposed Interpretation
0.71-1.00	A Strong structure has been found
0.51-0.70	A reasonable structure has been found
0.26-0.50	The structure is weak and could be artificial; please try additional methods on this data set
≤ 0.25	No substantial structure has been found

The mean squared error MSE is the second moment or variance of the error including the bias. It is a key indicator to judge whether the clusters chosen predict well the original data

set. When the chosen clusters represent well the original data, the values for the MSE and the mean absolute error MAE will tend to be close to zero or zero if the fit is perfect. Practically this will never happen, however solutions with values closer to zero represent better solutions. When comparing two evaluations of PAM runs, the one with the lower scatter is preferred. According to *Tan et al.* [2005, p.500], the sum of the squared error SSE is defined as

$$MSE = \frac{SSE}{n-m} \tag{4.6}$$

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} dist(c_i, x)^2 \tag{4.7}$$

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} (c_i - x)^2 \tag{4.8}$$

where
n refers to the sample size,
m refers to the number of parameters in the model,
K refers to the total number of clusters,
C_i refers to the cluster i,
c_i represents the cluster i's center point,
x represents the members of the cluster C_i,
$dist$ represents a distance measure.

In Equation (4.8), the distance measure is the squared euclidean distance instead of the general formulation above. [*Tan et al.*, 2005, p.514-515] shows that the minimum SSE is at least a local minimum:

$$\frac{\partial}{\partial c_k} = \sum_{i=1}^{K} \sum_{x \in C_i} \frac{\partial}{\partial c_k} (c_i - x)^2 = \sum_{x \in C_k} 2 * (c_k - x_k) = 0 \tag{4.9}$$

$$\Rightarrow m_k * c_k = \sum_{x \in C_k} x_k \Rightarrow c_k = \frac{1}{m_k} \sum_{x \in C_k} x_k. \tag{4.10}$$

m_k represents the number of objects in the k cluster. In addition, the mean absolute error (MAE) is calculated to avoid choosing a typical operating period that only in average represents well the original period. Therefore, an indicator measuring the absolute difference between the representation of the cluster C's members x with the medoid c_i is calculated.

$$MAE = \frac{1}{n} \sum_{i=1}^{K} \sum_{x \in C_i} |c_i - x| \tag{4.11}$$

The third indicator measures the sequence length. Considering an example with a typical year of 365 days that are assigned to 2 clusters: the indicator counts how many times the clusters centers are alternating. If the sequence starts with 1,1,1,2,1, ..., the sequence length at this

point is 3. For a daily storage model this is not of importance, however when considering long term storage a way of respecting this series has to be found.

The fourth indicator, the relative minimal and maximal peak load percentage, is a key value to energy systems. It links energy to (maximal) power. The input data should represent not only the average period but also the ones representing the peak load. In addition, the peak load curve shows whether a good representation has been found. When this is not the case, peak loads can still be added to representative clusters manually.

$$PE_k = \frac{k(X_{predicted}) - k(X_{original})}{k(X_{original})} \quad \text{with } k\text{: either } max() \text{ or } min() \tag{4.12}$$

The percentage error as an indicator judges the overall energy balance. For each dimension and overall, the percentage error PE is calculated:

$$PE = \frac{\sum_{i=1} X_{predicted} - \sum_{i=1} X_{original}}{\sum_{i=1} X_{original}} \tag{4.13}$$

$X_{original}$ refers to the original data per dimension,
$X_{predicted}$ refers to the original data represented through the clustering per dimension. Combined with the maximum and minimum percent error per dimension, one can get a quick overview of the fit of the curve. All three values, minimum and maximum power as well as the energy are a key indicator for the design of the energy system. Through the choice of data point within the data set, the PE will almost never be exactly zero. k-means guarantees an overall PE of zero.

[*Everitt et al.*, 2011a, p15] underlines the fact that humans can quickly catch patterns, if the data is well pre-processed. Studying the output as load curve plots and plots over the whole year per dimension allows a graphical representation of the clustering quality compared to the study of pure numbers. This is key point as it is difficult to identify errors only based on their numerical value. It might be possible that the average error is very low: The errors could cancel each other out with values higher and lower than the real data. The graphical representation can capture these cases quicker than the proposed indicators do. Therefore we suggest using them always in combination of each other.

4.2.7 Results: Selection Criteria for Key operating Periods

The data presented here during the application comes from the Verbier case study, presented in more detail in Section 5.9 on page 90. The heat load is based on the dynamic building simulation of Verbier through bSol. The solar irradiation data comes from the same meteorological file that is used for the input of the dynamic building simulation.

Table 4.2 – Proposed indicators judging the quality of the clustering and their use

Indicator	Use
Silhouette Coefficient	cluster meaningfulness
MSE	mean squared difference between observation and prediction
MAE	mean absolute difference between observation and prediction
Sequence Length	alternation of typical days
Overall PE	guarantee energy balance
Dimensional PE	guarantee energy balance of each dimension
min. and max. relative Peak Load	correct power representation
Graphical display	verification of indicator fit, mainly with cumulative energy demand curve and daily time -energy curve

With an increasing number of clusters k, the fit for almost all indicators improves. drops below 2%. Only the sequence length of day alternations increases and the silhouette coefficient decreases. Table 4.3 shows the results for the annual data set.

On the one hand, the silhouette coefficient suggest further tests as the values quickly fall below the threshold of 0.5 but they stay above the structure free 0.25 value. On the other hand, the values are suggested for a two dimensional data set and not for a 48-dimensional one.

The other "Overall Indicators" show that even with few clusters such as two or three, the data set can be represented well. Both the MSE, the MAE and the percent error are very low. Logically, sequence length of different day alternations is increasing the more clusters are introduced.

Studying the errors for the dimensions heat load and solar irradiation reveal however, that the peak loads are not represented well. As the peaks especially for heating are important, the first six results can be discarded. The "maximal Load Percent Error" changes in steps for both dimensions. This is normal behavior, because PAM selects an existing day out of all days within a cluster. Therefore it is possible that this medoid representing the group remains the medoid even though it does not represent all indicators very well. It is still the best compromise out of all points. Especially on the solar side, the peaks are not well represented because they appear during one hour a day and not once during several hours of the same day. Their weight is relatively small compared to the heating for example.

The close fit can be confirmed through the graphical representation of Figure 4.2 showing the annual heat load curve and the annual solar load curve Figure 4.3. Only the solar peak is not

Table 4.3 – Choosing seven clusters as a results of annual clustering

Overall Indicators						Heat Demand			Solar Heat Production		
Clusters k	Silhouette	MSE	MAE	Percent Error	Sequence Length	Max. Load % Error	Min. Load % Error	% Error	Max. Load % Error	Min. Load % Error	% Error
2	0.59	0.02	0.08	-10.30	33	-46.56	0.04	-15.70	-42.73	0.00	0.51
3	0.48	0.01	0.07	-9.23	120	-46.56	0.03	-15.01	-24.97	0.00	2.35
4	0.42	0.01	0.07	-8.32	158	-31.54	0.03	-10.74	-24.97	0.00	-3.48
5	0.39	0.01	0.06	-3.21	164	-31.54	0.00	-4.71	-17.01	0.00	-0.20
6	0.38	0.01	0.06	-1.00	192	-10.97	0.02	-3.07	-16.28	0.00	3.15
7	0.36	0.01	0.06	-2.27	202	-2.79	0.06	-3.30	-16.28	0.00	-0.22
8	0.34	0.01	0.06	0.26	206	-1.35	0.02	-0.84	-16.28	0.00	2.46
9	0.33	0.01	0.06	-2.38	230	-1.35	0.02	-3.70	-16.28	0.00	0.25
10	0.30	0.01	0.05	-2.61	237	-1.35	0.02	-3.84	-16.28	0.00	-0.15
11	0.29	0.01	0.05	-0.26	239	-1.35	0.02	-1.11	-16.28	0.00	1.45
12	0.27	0.01	0.05	-2.89	251	-1.35	0.00	-3.48	-17.28	0.00	-1.71
13	0.26	0.01	0.05	-2.16	251	-1.35	0.00	-2.00	-17.28	0.00	-2.47
14	0.25	0.01	0.05	-1.90	258	-1.35	0.00	-2.01	-17.28	0.00	-1.70
15	0.25	0.01	0.05	-1.60	259	-1.35	0.00	-1.44	-17.28	0.00	-1.92

well represented, which is already evident after studying the results of the results overview table. Looking at Figures 4.3 and 4.4 confirms the statement that the overall weight of the solar peak is low and the fact that the peaks appear within different days during one hour only. They can be neglected for the overall energy balance or a manually added day can represent it.

After the annual clustering, the same clustering procedure can be applied to the hours of each day. The results for day 184 are shown in Table 4.4. Again, the number of hours is chosen that offers the best trade off between the indicators. The threshold, is set to a 10% error per dimension. With $k = 6$ clusters, the day can be represented using 6 typical hours that are in a sequence containing 10 alternations between the typical hours. The figures 4.6 and 4.5 allow to study the fit of the clustering algorithm.

During the second clustering, when the hours of the day are reduced, once again the quality of the results must be compared to the original data set. Table 4.5 shows that these differences are not significant.

4.2.8 Discussion of Results

As the number of clusters k increases, the quality of the yearly representation also increases. The error indicator gives a quick overview about which cluster number to choose and what will be the draw-back of the simplification in terms of the energy balance and the power.

When changing the number of clusters, steps especially in the maximal load error per dimension can be witnessed due to the fact that the cluster medoid representing the extreme values

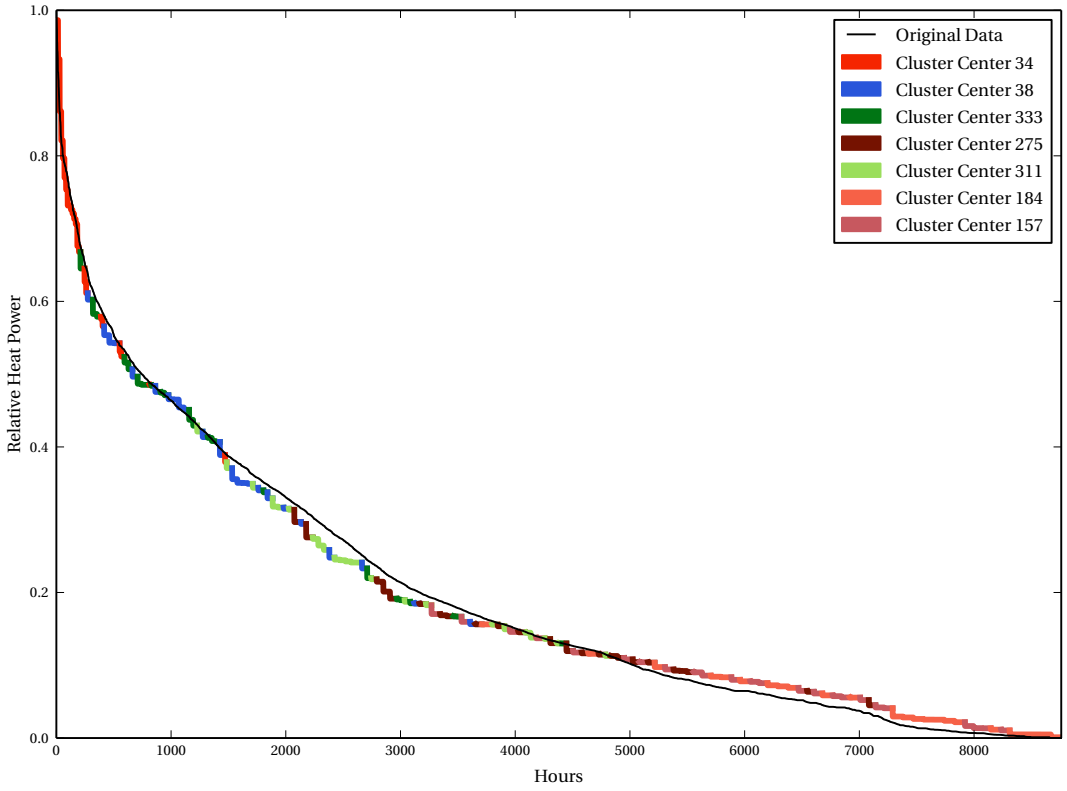


Figure 4.2 – Annual heat load curve of the heating and hot water demand. The black line represents the original data, the colored lines represent the different typical days. For the heat demand, the quality of the fit is very high as both lines are close to each other.

rarely changes. The clusters containing extreme points are rare and do not change frequently due to the distance function's heavy penalty of not using a maximal point. k-means, which takes averages of the data leads to smoother changes when the cluster number is changed. Changing the distance function to the squared euclidean distance has the effect of considering peaks much better.

Choosing seven clusters allows to represent the heat load with a high degree and only very little errors. For solar irradiation, is also well represented with an overall error of around one percent on the annual energy balance. However the peak load value remains 15% lower because the peak for solar has only very little weight compared to the rest. As the peak load for solar appears generally during few hours throughout the year, one hour a day over several different days of the year, the representation of the overall energy balance is more important.

In the case where peak loads are not taken into consideration (as would be shown by the peak load indicator), these can be added manually. When the representation is good, the chosen value comes very close to the original value, small differences can be tolerated: For example, building simulations can be run without a realistic limit on the maximal available power to

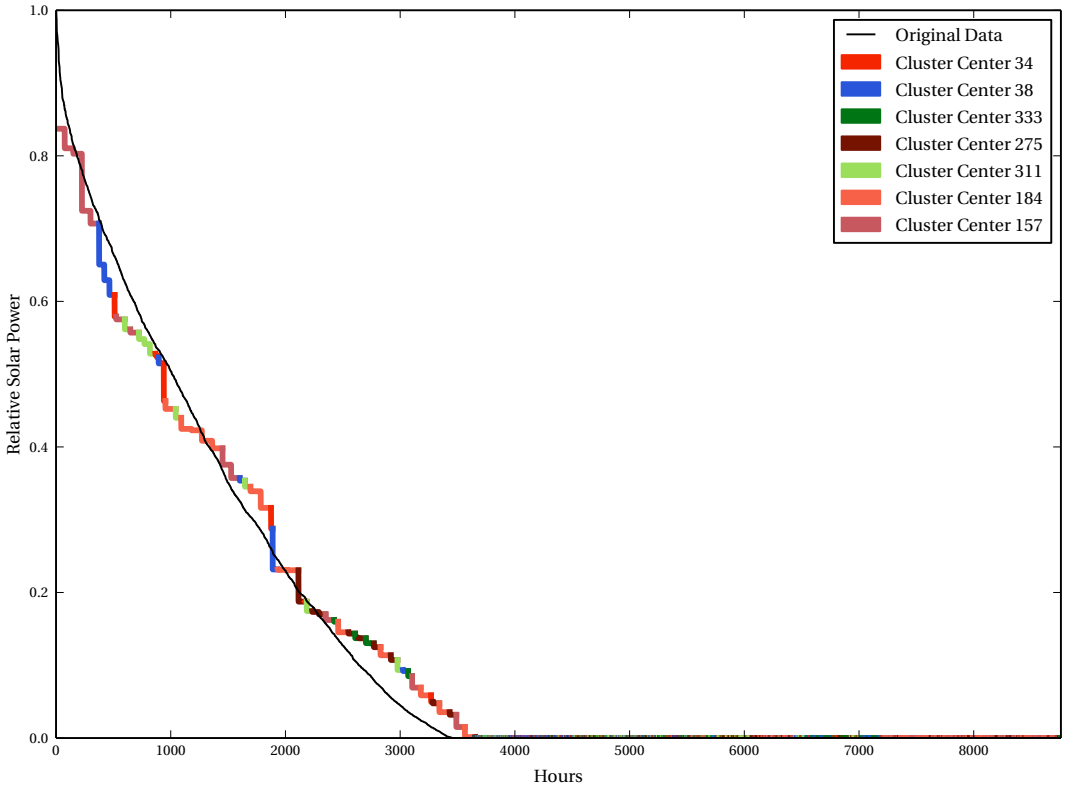


Figure 4.3 – Annual solar irradiation load curve: The black line represents the original data, the colored lines represent the different typical days. The fit for the solar load curve does not reach the maximal load and does not always follow very closely the original data. However the energy balance is still represented without almost no error.

calculate the overall energy demand. Therefore, it is acceptable to choose a lower value within the clustering. As a consequence, the building might not reach the internal temperature set point instantly, but use more time. However, it should be verified, e.g. with the help of a dynamic building simulation testing whether the set point can be reached.

The silhouette plot and coefficient are not more than indicators for multi-dimensional data sets. The silhouette plot is intuitive to use for one dimensional data set, however the results do not support the use of the silhouette coefficient on a multi-dimensional data set. Instead, the silhouette coefficient indicates the quality of a solution, but it needs to be supported by other indicators. According to the additional indicators, a lower value of the silhouette coefficient than shown in Table 4.1 provides feasible solutions. Even relative low values of them give acceptable results for the clustering. This might be due to the fact that the best practice values are for the use of a one dimensional data set.

When creating the days that represent the year, they can be introduced in the optimization problem either with respect of the sequence in which they appear or just counting the number of their appearances disregarding the sequence. On one side for a cyclic daily storage, the

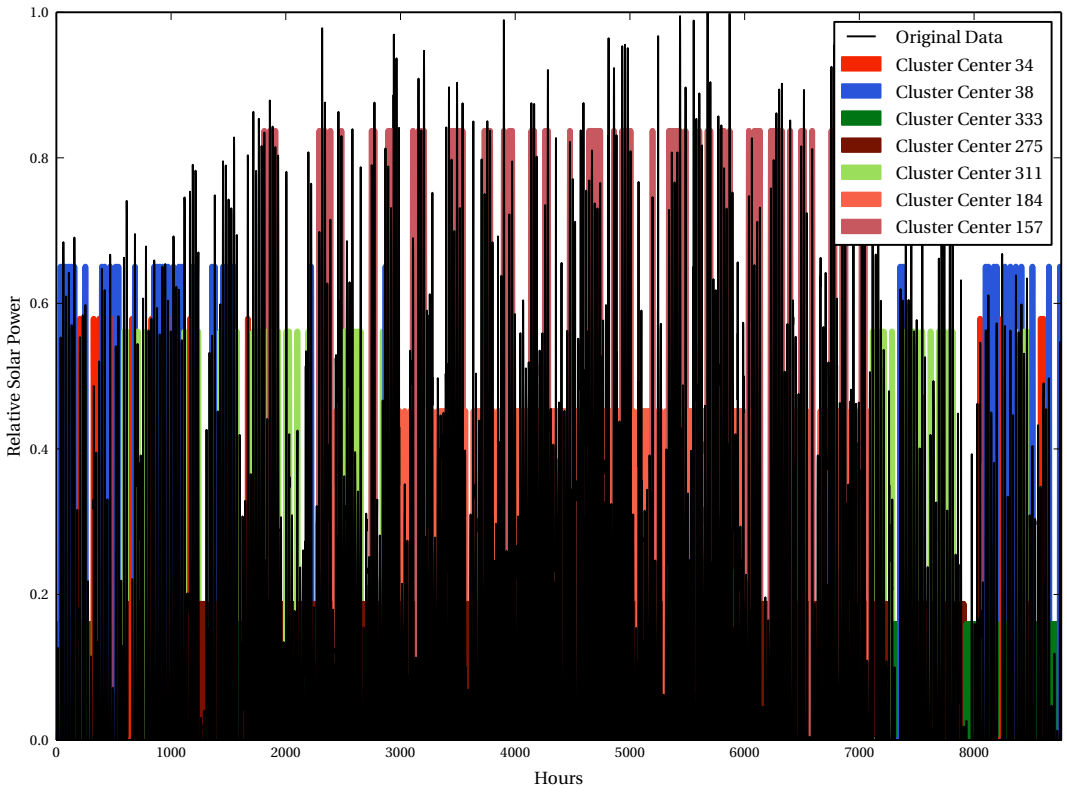


Figure 4.4 – Results of annual clustering for solar thermal heat production: The black line represents the original data, the colored lines represent the different typical days. The variations are well represented, however the peak production of the estimate is about 15 % lower than the original one, while almost perfectly respecting the energy produced.

sequence of days is not important. On the other side for long term storage integration, the sequence directly influences whether or not a storage using the storage makes sense. However the length of the sequence also directly impacts the size of the optimization problem. A sequence of around 200 days is computational heavy. As long as only a daily storage is required, this approach also works fine. When considering the long term storage, the proposed approach is modified using the same clustering strategy over each month with the same indicators to choose a representative day for each month. With this approach, the sequence of the days is representing also the real sequence of their appearance in the year while reaming short. When the indicators show an acceptable result for one day representing a month, the day can be chosen. If not, the month has be split and two days should be chosen. In the case, where the months are represented by similar days, they can be regrouped to one again.

Because this monthly approach has a lower change of well representing the maximum heat or solar load day, the extreme conditions are manually added to the clustering ensuring the correct sizing of the system.

Table 4.4 – Selecting six clusters as the results of clustering hours of day 184

Overall Indicators						Heat Demand			Solar Irradiation		
Clusters k	Silhouette	MSE	MAE	Percent Error	Sequence Length	Max. Load % Error	Min. Load % Error	% Error	Max. Load % Error	Min. Load % Error	% Error
4	0.58	0.00	0.03	7.62	7	0.00	2.30	12.87	-9.67	0.00	2.35
5	0.53	0.00	0.03	4.00	9	0.00	1.51	7.24	-6.07	0.00	0.74
6	0.34	0.00	0.02	3.05	10	0.00	1.51	6.87	-6.07	0.00	-0.80
7	0.36	0.00	0.02	-4.58	12	0.00	1.34	-8.46	-6.07	0.00	-0.69
8	0.29	0.00	0.02	-3.27	14	0.00	1.34	-4.77	-6.50	0.00	-1.76
9	0.18	0.00	0.02	4.03	15	0.00	1.51	9.80	-6.50	0.00	-1.76
10	0.24	0.00	0.02	5.05	15	0.00	1.51	10.88	-6.50	0.00	-0.81
11	0.27	0.00	0.01	3.78	15	0.00	1.51	7.47	-6.50	0.00	0.07
12	0.19	0.00	0.01	3.40	17	0.00	1.51	5.90	0.00	0.00	0.90
13	0.12	0.00	0.01	3.82	17	0.00	1.51	7.36	0.00	0.00	0.26
14	0.14	0.00	0.01	2.26	17	0.00	1.09	4.54	0.00	0.00	-0.03
15	0.06	0.00	0.01	1.82	18	0.00	1.09	3.70	0.00	0.00	-0.07
16	0.10	0.00	0.01	1.28	18	0.00	1.09	2.58	0.00	0.00	-0.01
17	0.11	0.00	0.00	2.78	18	0.00	1.09	5.55	0.00	0.00	0.00
18	-0.03	0.00	0.00	2.05	20	0.00	1.09	4.08	0.00	0.00	0.00
19	0.02	0.00	0.00	-0.41	21	0.00	0.00	-0.82	0.00	0.00	0.00
20	-0.03	0.00	0.00	-0.60	21	0.00	0.00	-1.20	0.00	0.00	0.00
21	-0.02	0.00	0.00	-0.25	21	0.00	0.00	-0.50	0.00	0.00	0.00
22	-0.01	0.00	0.00	0.75	22	0.00	0.00	1.50	0.00	0.00	0.00
23	-0.02	0.00	0.00	0.23	23	0.00	0.00	0.46	0.00	0.00	0.00

Contrary to remarks found in the literature, the deterministic PAM approach performs well in the tested cases concerning computational time. Compared to k-means, where the right number of clusters and the number of restarts has to be chosen, PAM delivers its results with a single iteration.

4.2.9 Conclusion: Choice of k clusters

Choosing the correct number of cluster k remains subjective to the application. The proposed indicators help for the energy system design together with the graphical representation of the results. The acceptable threshold value should be chosen with an overall knowledge of the decision process. Two key question for this are:

1. what is the uncertainty on the input data of the clustering?
2. what are the safety margins that are added at the end of the design process?

The answer to these questions should help to find the compromise between more precision and more reduction.

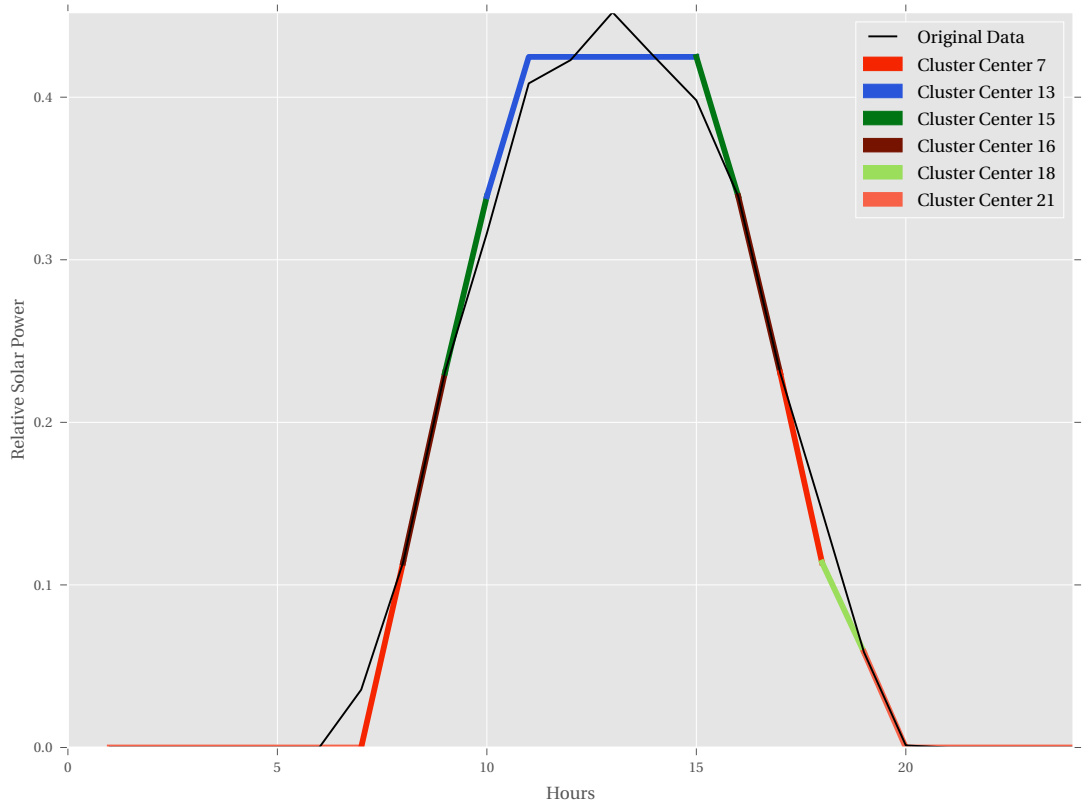


Figure 4.5 – 6 clusters representing the hourly solar irradiation for day 184, that alternate leading to a sequence length of 10

During this work, the acceptable error is chosen to be ≤ 10%, because the uncertainty on the input data is high and the knowledge of current practice in sizing is kept in mind. As stated in the introduction, according to *Porges* [2001, p.119] a safety margin of 10 to 15% that is often added for sizing of equipment after all calculations are made. Furthermore, *Porges* [2001, p.91] states that the margin added to the heat loss calculations are between zero to twenty percent (50 percent for buildings not in daily use) that have exposed north surfaces. However it is important to communicate the chosen hypothesis to avoid that at every step within the decision process, a security margin is added.

The buildings heat load can only be verified based on few annual measurements. As shown in Chapter 3, the variation of the demand can even be bigger. This is a choice in favor of the reduction at the cost of loosing minor precision, especially regarding the solar peak load. For the presented example 7 days are chosen to represent the year. Another clustering on the day, further reduces the data without loosing additional detail. In the example of day 184, 6 clusters are chosen. The same procedure is repeated for every day.

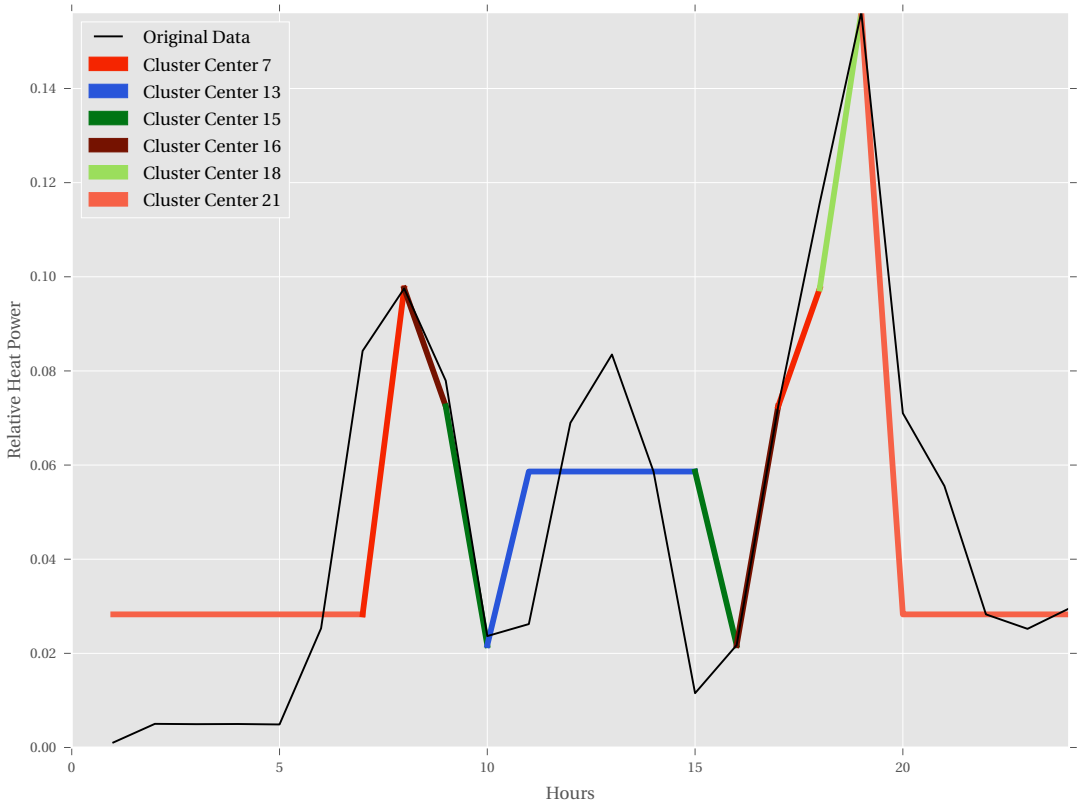


Figure 4.6 – Representation of day 184 when using 6 clusters. The sequence length is also 10.

4.3 Point Reduction of the Composite Curve

When aggregating individual buildings to one zone and therefore one composite curve, a composite curve with a lot of slope changes is created. Especially when using a dynamic building simulation to deduce the heating needs, the buildings almost always have a slightly different temperature level requirement. For large multi-period problems, it is of interest to reduce the number of points within the curve per period to reduce the complexity of the problem.

The composite curves for the heat requirement can be simplified with the help of the Ramer-Douglas-Peucker algorithm [*Ramer*, 1972]. This algorithm is used to simplify coast lines, when zooming out of a map. It tries to reduce the original curve to a new curve with fewer point based on a maximum allowed error by testing systematically how many points can be excluded between the starting and end point of the composite curve. It is of O(n log n) complexity. Figure 4.7 shows the heat demand of Verbier in the winter during hour 8326. The 60 streams from 60 building are reduced to 5 curve segments respecting the start and end points with a predefined error (that is respecting the input data quality data).

Table 4.5 – Compared to Table 4.3, only minor differences after reducing the hours of each typical day to represent the year appear

Overall Indicators					Heat Demand			Solar Irradiation		
Clusters k	Silhouette	MSE	Percent Error	Sequence Length	Max. Load Percent Error	Min. Load Percent Error	Percent Error	Max. Load Percent Error	Min. Load Percent Error	Percent Error
Tab. 4.3	0.41	0.01	-0.07	202	-1.36	0.00	-0.86	-16.28	0.00	1.08
Reduced	0.41	0.01	-0.71	202	-1.36	0.02	-1.24	-19.72	0.00	0.05

4.4 Conclusion

The significant reduction of a two dimensional data set using a deterministic clustering algorithm, PAM, has been demonstrated. The encountered problems and solutions to them were discussed.

One big remaining challenge is the creation of a short sequence with the same high quality output. A year represented by 8760 hourly heat load and solar heat production values can be simplified to a much shorter sequence for the sizing of energy systems. Applying the procedure twice, once for the year and then per day, allows to reduce the sequence length from 360 days to around 200 days when respecting the order of days or between 7 to 10 days, when it is not important. A comprise of monthly medoid days can be used with the same approach, because the reduction respecting the order of days is not sufficiently short enough for the integration of a long term thermal energy storage into an existing MILP framework. However, monthly averages often require the manual addition of extreme periods.

Compared to the k-means approach that is mostly used in literature, this approach is deterministic and thanks to the algorithm development and the increasing computational power, feasible. With the change of the distance measure from the euclidean distance to the squared euclidean distance, the k-medoids captures the peaks well. In addition, PAM chooses days within the data where as k-means creates new mean centroids representing a cluster. If the results should be fed back to a building simulation for example, this approach has an advantage of choosing a real point from the data rather than an average.

The here presented indicators are oriented towards the use through non clustering experts and especially energy system designers while still ensuring a high quality result. Compared to the existing indicators for judging the clustering results such as inter-cluster separation and cluster cohesion or the ones proposed by [*Fazlollahi et al.*, 2014b], they are easier to use. The combination of numerical indicators and graphics allows a quick judgment about the result's quality.

For the energy system design, a recommendation for concerning the threshold or error is made. Nevertheless, it remains subjective to a certain part. It depends on the relation between precision and reduction size, which is linked to the input data quality and expected output

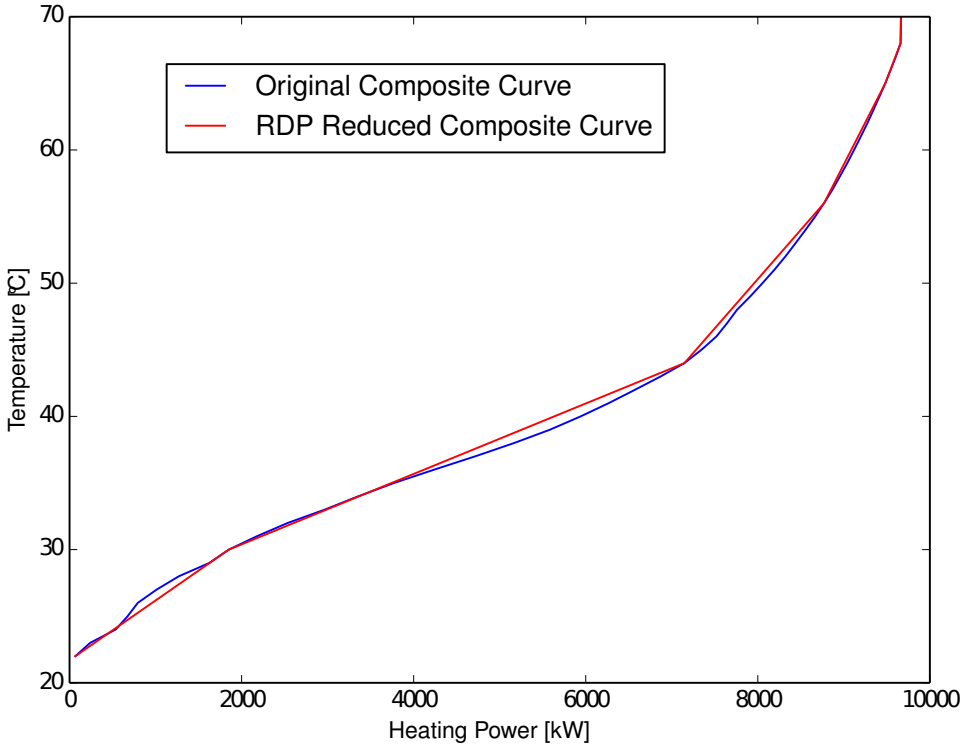


Figure 4.7 – Composite versus reduced composite curve for hour 8326 of the heat demand

quality. When the output quality should be very high, the reduction will be less effective. A high input data uncertainty might allows to accept a higher error.

In this work, the threshold values for an acceptable error is set to smaller than 10% on the energy balance and heat power. For solar, the lowest possible number is chosen, because it does not satisfy this criteria. Compared to a safety margin of 10 to 15% that is often added for sizing of equipment according to [*Porges*, 2001, p.119] after all calculations are made, the here produced quality of output is very high leading to fewer reductions. With this in mind the margin can also be lifted if needed leading to shorter sequences.

This work intensively looks at clustering, especially k-means and PAM. With regard to the vivid scientific interest in the clustering field, further development in clustering techniques to better consider sequences would be of interest.

5 Integration of renewable Energy into Urban Energy Systems with thermal Storage

The integration of renewable energy sources in the urban energy system is an important if not the most important challenge. In addition to storable renewable energy sources like biomass or geothermal that can be accessed when they are needed, it is important to consider renewable sources that are only available stochastically and which present seasonable variations like the sun or the wind.

In order to use renewables which are only stochastically available, the energy system can include a storage system. Here, solar heat energy is used as an example for renewable energy. On the roof surface of a building, the irradiation can be converted into electricity or into heat in form of glycoled hot water. When demand and supply do not match, storage system can fill the gap. On one side, *Marc A. Rosen* [2012, p.6] notes the advantages of energy storage:

> “
> - Increased operational performance, reliability and flexibility.
> - Decreased mismatches between periods of energy supply and demand.
> - Enhanced opportunities for the use of renewable energy resources.
> - Improved opportunities for distributed generation.
> - Improved economics over energy system life time
> - Increased system efficiency and, correspondingly, decreased utilization of energy resources.
> - Reduced space requirements due to smaller size of equipment and overall system (possible).
> - Decreased environmental impacts associated with the provision of energy services.
> - Enhanced energy sustainability.
>
> ”

On the other side, energy storage has the following disadvantages [*Marc A. Rosen*, 2012, p.7]:

> - Loss of efficiency.
> - Increased initial (investment) costs.
> - Economics. The economics of energy storage are not advantageous in all applications.
> - Range of performance. Some energy storages are not able to perform well at all operating conditions encountered in some applications.

The attentive reader notices that the citation of *Marc A. Rosen* [2012, pp.6-7] contains a lot of "ifs". This indicates that a careful case by case decision needs to be performed because it depends on how the storage is implemented into the system, the energy system itself and the (economical) boundary conditions.

5.1 State of the Art

For the urban system design, 3 types of elements are defined: on the one hand the harvesting device, i.e. the solar panel type and its size and its operating conditions, i.e. the temperature at which the heat will be harvested, and on the other hand, the size of the storage tanks, its (their) insulation level that may depend on their location and its operating conditions, i.e. temperatures profiles. It is worth to mention that the storage system will play the role of an intermediate and will therefore be interconnected with the final service delivered, i.e. the heat distribution system for supplying the heat to the building or the hot water supply. The storage device will need in addition a control strategy that will decide when, what and how much is stored, when, what and how much is distributed and when, what and how much is received once it is installed. Such a control system has to be predictive since it has to match the stochastic inputs considering what is known from the future inputs and the future requirements. In this review only design methods are considered. They do not have an implemented controller, they calculate an optimal choice in the definition of the storage equipment.

Pinel et al. [2011] reviews different seasonal storage concepts for storing solar thermal heat in residential applications. They conclude that only sensible heat storage is functioning and available at interesting prices. Representing a sensible heat storage in a one dimensional numerical formulation can be used to accurately describe the performance of it [*Cruickshank and Harrison*, 2010]. Detailed simulation methods for different storage types and models are discussed by *Pinel et al.* [2011], non of them directly includes optimization.

For the design of seasonal storages, especially in the eighties guidelines were published ([*Hadorn*, 1990], [*Drew and Selvage*, 1980], [*Kovarik*, 1981], [*Lund*, 1989]). All of the here cited works use seasonal storage together with solar thermal panels. Normally a backup utility such as a boiler or sometimes also a heat pump exist.

They are a sequential approach to system design, taking one decision after another. However, they are very complete in terms of demand and supply analysis for the time and oriented towards practical usage even though only few realizations were done. The interest in the field started again after 2000: [*Tanaka et al.*, 2000] find that the use of a seasonal storage in a district heating network generally increases the efficiency in their case studies. [*Lindenberger et al.*, 2000] publishes one of the few studies combining seasonal storage, solar heat and optimization. A detailed problem description is missing, the use of the simplex method points towards a linear model that however only optimizes operational strategy within each hour of the year to compare a number of scenarios without sizing the system components.

The first return of experience from constructed thermal seasonal storage systems is available: *Schmidt et al.* [2004] shows the cost functions, while *Bauer et al.* [2010] formulate a first feedback for the cases in Germany. They still need more research on insulating the storage tank in order to reach the planned performance, e.g. recover the amount of energy that was planned. The more recent project of Drake Landing Solar Community on Okotoks, Canada, however reached the planned performance according to *Sibbitt et al.* [2012].

Over the past 30 years the Danish installed more and more so called solar district heating systems: A solar district heating system is a system with solar as the biggest energy resource. Often biomass boilers ensure the rest of the energy demand. Especially smaller cities and communities are installing bigger and bigger collector fields in the order of up to 75 000 m^2 and seasonal storage tanks up to several 100 000 m^3. Compared to the German approach that is integrated into a research program with publications available, the Danish program did not publish their advances. The Danish company PlanEnergi and the Danish District Heating Association offer presentations of their advances in the field.

When looking at optimization approaches in the field of process integration, *Grossmann and Santibanez* [1980] introduce a multi-period formulation, however without storage. *Kemp* [2007] introduced the time slice model, based on the same principles. *Kemp and MacDonald* [1987] conclude that time average approaches give too optimistic results. Working with time slice approaches leads to a realistic integration of energy storage. In the field of batch processes, *Krummenacher* [2002] identifies the minimum number of heat storage units for heat recovery based on a heuristic targeting method.

A different approach is proposed by: *Schütz et al.* [2015] define a constant mass per layer and estimate temperature variations within the storage tank. Such a model can not be integrated into the heat cascade, because the temperature is a variable in their model whereas the heat cascade requires it to be fixed.

Varbanov and Klemeš [2011] propose a framework using a total site approach for the integration of renewables with a heat storage. *Nemet et al.* [2012] use solar thermal energy in combination with storage in a one-day case study showing that it is of interest to store solar thermal energy when the demand is varying over the day. They do not discuss the problem of temperature choice for the solar thermal panel, neither is the tank specified. *Rager et al.*

[2013b] propose a method using a genetic algorithm for the optimal temperature choice of the solar collectors, maximizing their use. This approach allows to compare PV and solar thermal collectors, however the use of the genetic algorithm is time consuming. In [*Rager et al.*, 2013a], the solar temperatures were predefined as done by many, which is also rarely an optimal solution. The buildings do not require heat at the same level except for the hot water demand. This can lead to an inefficient integration of solar energy, when the temperatures are not chosen correctly.

In the work of *Collazos et al.* [2009], a predictive method using a linear programming model has been proposed to size the storage tanks of a cogeneration unit. The predictive control algorithm has then been adapted to pilot the cogeneration unit. In the approach proposed, the authors consider a cyclic constraint that allows to size the system with typical days without depending on the initial and final conditions. It has to be mentioned, that, during the sizing procedure, the future is estimated but is perfectly known which implies that the control algorithm does not need any correction measures. The authors however mention that one of the characteristics using the control algorithm is that the building inertia can also be considered as a storage opportunity which implies for the building system to consider comfort temperature ranges instead of the constant value typically used. The method proposed is simplified since it does not consider temperature variations in the storage tank feed. *Becker* [2012] models a thermal energy storage considering different temperature levels in the storage tanks for the use in industrial processes. *Fazlollahi et al.* [2014a] applied and adapted Becker's model for the design of the energy conversion system of a district heating system. In addition, the method has afterwards been adapted to integrate heat pumps and solar PV systems in microgrids applications [*Menon et al.*, 2013].

The discussed models use water as a storage medium. Especially in the building sector, phase change materials are considered as an interesting alternative ([*Tatsidjodoung et al.*, 2013], [*Sharma et al.*, 2009]) According to *Hadorn* [2008], water is still the medium that is the cheapest and easiest to handle compared to other mediums and chemical or latent heat storage.

Numerous publication exist proposing the optimal energy system design including thermal storage, however they rarely consider the temperature levels. Either the problem is limited to one temperature level or the storage is often modeled using only an energy balance. Losses are considered in a limited context, e.g. only as a function of the storage size ([*Ren and Gao*, 2010], [*Lozano et al.*, 2010], [*Heidari Tari and Söderström*, 2002]). This bears the risk that the potential is not correctly identified, because the heat storage might not be at the temperature any more that is needed to fulfill the requirements.

5.1.1 Resume

The reviewed literature has a wide variety of simulation approaches concerning thermal energy storage to offer. On the optimization side, few approaches exist. In the reviewed literature for optimization, thermal energy storage is often modeled with an energy balance leading to

inaccuracy, because the thermal losses can not be evaluated based on the temperature of the storage.

Seasonal thermal storage appears often together with solar thermal panels in the literature. Sizing of seasonal energy storage in an optimization based approach was not found, only useful guidelines. The return of experience from the first monitored seasonal thermal energy storage shows that the technology can work, but needs a careful planning and execution. The current planning is however rather guided by idealism or political will than by a methodology leading to an optimal result.

The method proposed here uses as a basis the approach developed by *Becker* [2012] that is adapted to study the integration of solar heat in the system. This implies on the one hand to consider the daily variations of the solar heat, on the other to also to consider the possibility of the seasonal storage in order to use the harvested heat later in the year.

Based on the findings in the literature review, a sensible heat storage is modeled. Water is used as a medium in the storage model.

[*Angrisani et al.*, 2014] points out that the consideration of different temperature levels is crucial. Especially, when integrating solar thermal collectors, the result can become very different. [*Angrisani et al.*, 2014] confirms these findings with a dynamic simulation showing −32% in solar energy use and +15% percent in boiler usage, when the stratification is not taken into account.

The following model will be developed based on the shortcomings found in the literature:

- a daily energy storage adapted to urban energy design,
- a seasonal energy storage and
- a solar thermal model maximizing the use of solar energy as an example for the integration of renewable energy.

5.2 Daily thermal Energy Storage Model

First, the daily storage model is explained: it can pass energy from one time slice to another time slice of the same day.

The heat storage model is based on a set of virtual storages with fixed temperatures that correspond to a discrete temperature range in which the storage tank can be operated. Water can be transferred from one temperature level v to another one passing through a heat exchange that is included in the heat cascade model. Heat can be delivered from the virtual tank v of the storage system through a hot stream from the virtual tank's temperature level T_v to the temperature level T_{v-1}. The mass flow $\dot{\boldsymbol{M}}_{\boldsymbol{v,h,p,t}}$ is the decision variable that represents the amount of water transferred from tank v in the time slice t of period p to the tank $v-1$. Similarly a heat supply to level v is defined by a cold stream from level T_{v-1} to T_v with a mass

flow rate of $\dot{\boldsymbol{M}}_{\boldsymbol{v,c,p,t}}$. The corresponding heat loads are calculated, considering a constant c_p for the fluid used in the heat storage with Equations (5.3) and (5.4).

The heat loss in the virtual tank v with the temperature T_v is represented by a mass flow that flows with the temperature T_v to the level $v-1$ with the temperature T_{v-1}. It exchanges its heat with the environment cooling the stream down.

Each energy storage has nv discrete temperature levels. Between each temperature level, a hot stream h and a cold stream c exist totaling to $(nv-1)\cdot 2$ streams for each storage. Both streams c and h belong to the same unit u. These units link the storage tank with their unit multiplication factor $f_{u,p,t}$ to the heat cascade. They exist in every period p with time slice t. The power that every stream can provide, depends on the temperature difference between the two virtual tanks v and $v-1$, the specific heat c_p of the medium and the nominal mass flow rate of the $\dot{M}_v$.

5.2.1 Mass Balance

The total mass M_{max} of liquid contained in the storage is fixed. The sum of all virtual tanks nv within the storage add up to the total mass. It is constant for each period p and time slice t (Equation (5.1)) and only limited by the total mass M_{max}. Additionally, non-negativity constraints are required, ensuring that the mass can not not take negative values.

$$\sum_{v=1}^{nv} \boldsymbol{M_{v,p,t}} = M_{max} \tag{5.1}$$

$$\boldsymbol{M_{v,p,t}} \geq 0 \tag{5.2}$$

For Equations (5.1) and (5.2) : $\{v \in [1, nv]\} \qquad \{p \in [0, np]\} \qquad \{t \in [0, nt]\}$

The neighboring virtual tanks are interconnected via heat exchangers that can either absorb heat for storing it or release heat fulfilling requirements of the system. The mass balance for the time slice t in period p is used to calculate the size of the virtual storage tank v in the next time slice $t+1$. Based on the already stored mass in the virtual tank v in the current period p and time slice t, $\boldsymbol{M_{v,p,t}}$, mass flows $\dot{\boldsymbol{M}}$ from and to the next higher level $v+1$ as well as the ones to and from the next lower level $v-1$ can exist during one time slice. They are calculated to know the mass $\boldsymbol{M_{v,p,t+1}}$ in the next time step $t+1$.

Figure 5.1 shows the energy storage. As illustrated in Figure 5.1, 5 virtual tanks exist however only 4 mass flows linking them, because they are always connecting them. Excess heat received can be stored through a mass flow from a lower to a higher level via an heat exchanger. Or heat can be released from the storage tank, leading to a mass transfer of heat from a higher to a lower level through the heat exchanger. The heat received or given is inserted into the system's heat cascade. The heat losses are calculated as a function of the virtual tank's temperature and

the ambient temperature of the tank's surrounding environment leading to a mass flow from the higher to the next lower level.

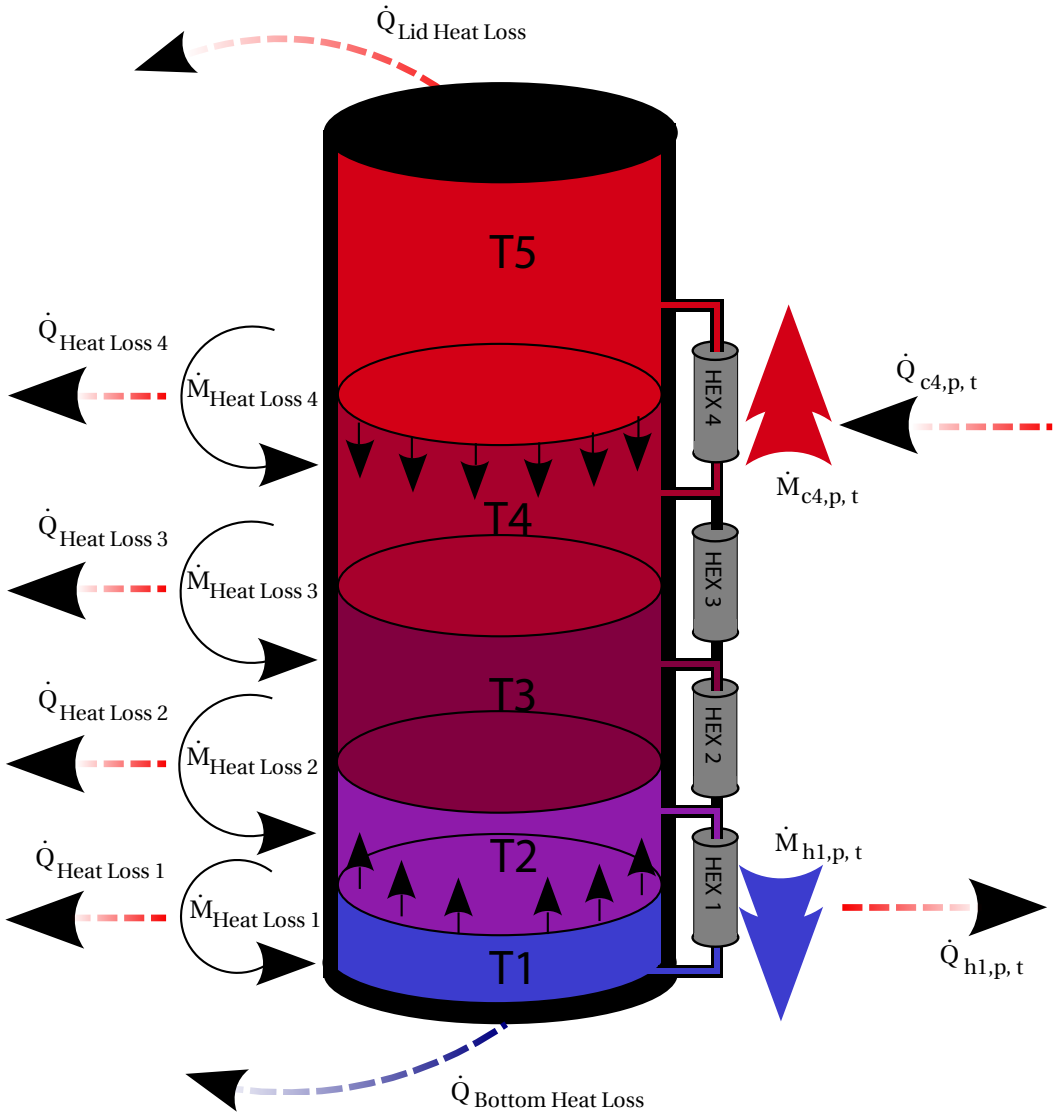


Figure 5.1 – The stratified thermal energy storage model used for the optimization of a daily and seasonal thermal energy storage. Bottom lid losses appear only when the virtual tank at the bottom is empty and therefore a virtual tank with a higher temperature T_v than the ambient temperature T_a is filled.

The following convention has been applied: the subscript $v = 1$ for a mass flow $\dot{\boldsymbol{M}}_{\boldsymbol{v=1}}$ connects virtual tanks 1 to 2. No mass flow with the subscript nv exists, the highest one is $nv-1$ connecting the level nv to $nv-1$.

The mass flows are defined with the multiplication factor $\boldsymbol{f_{u,p,t}}$ linking the nominal mass flow $\dot{M}_v$ to the heat cascade. Both values are always positive. The multiplication factor $\boldsymbol{f_{u,p,t}}$ is limited by the maximal amount of mass that can be moved within one time slice without destroying the stratification of the tank. The specific heat capacity c_p is constant, the temperature difference between 2 virtual tanks is defined as $(T_{v+1} - T_v)$. The heat released $\boldsymbol{\dot{Q}_{h,v,p,t}}$ or received $\boldsymbol{\dot{Q}_{c,v,p,t}}$ from a virtual tank is then integrated in the heat cascade with its power-temperature profile. Equations (5.3) and (5.4) are based on Equation (2.17), the same restrictions apply

Equation (2.17): $\dot{M} = \dfrac{\dot{Q}}{c_p \cdot \Delta T}$

$$\boldsymbol{\dot{Q}_{c,v,p,t}} = \boldsymbol{f_{u,p,t}} \cdot \dot{M}_v \cdot c_p \cdot (T_{v+1} - T_v) \tag{5.3}$$

$$\boldsymbol{\dot{Q}_{h,v,p,t}} = \boldsymbol{f_{u,p,t}} \cdot \dot{M}_v \cdot c_p \cdot (T_v - T_{v-1}) \tag{5.4}$$

$$\{c \in [1, nc]\} \quad \{h \in [1, nh]\} \quad \{u \in [1, nu]\} \quad \{v \in [1, nv-1]\} \quad \{p \in [0, np]\} \quad \{t \in [0, nt]\}\ .$$

The nominal flow rate $\dot{M}$ is the same for all charging and discharging levels of a storage tank. For the rest of the model description, the multiplication of the two terms $\boldsymbol{f_{u,p,t}}$ and $\dot{M}_v$ are regrouped to $\boldsymbol{\dot{M}_{h,v,p,t}}$ referring to the hot stream with the subscript h and $\boldsymbol{\dot{M}_{c,v,p,t}}$ to the cold one with c respectively.

The mass balance for the bottom virtual tank $v = 1$, shown in Equation (5.5), receives the mass flow $\boldsymbol{\dot{M}_{h,1,p,t}}$ with the subscript h for the link to the hot stream $v = 1$ coming from the level $v = 2$ to level $v = 1$ providing heat to the system. The cold stream c is linked to the mass flow with the same subscript, $\boldsymbol{\dot{M}_{c,1,p,t}}$, and has a negative sign while absorbing heat. In addition to the intended possibility of a controlled heat exchange, the heat losses hl are introduced, which are going to the environment. They are represented in form of a mass flow $\boldsymbol{\dot{M}_{hl,v,p,t}}$ reducing the next higher level's mass through a transfer of heat to the environment. The first level is set at ambient temperature (or lowest ambient temperature in the overall problem) and therefore has no losses.

The mass balance of the virtual tanks in between the top and bottom tanks are described in Equation (5.6). They can either receive or give the neighboring levels. Therefore the heat losses appear twice, because the virtual tank receives the losses from the next higher level and

has himself losses to the next lower level. The top level $v = nv$ is described by Equation (5.7) and is only connected to the level $nv-1$.

$$\text{For } \{v = 1\}: \quad \boldsymbol{M_{1,p,t+1}} = \boldsymbol{M_{1,p,t}} + \sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{h,1,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{hl,1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{c,1,p,t}}\right) \tag{5.5}$$

$$\text{For } \{v \in [2, nv-1]\}: \quad \boldsymbol{M_{v,p,t+1}} = \boldsymbol{M_{v,p,t}} + \left[\sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{c,v-1,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{h,v,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{hl,v,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{c,v,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{h,v-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{hl,v-1,p,t}}\right)\right] \tag{5.6}$$

$$\text{For } \{v = nv\}: \quad \boldsymbol{M_{nv,p,t+1}} = \boldsymbol{M_{nv,p,t}} + \sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{c,nv-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{h,nv-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{hl,nv-1,p,t}}\right) \tag{5.7}$$

For Equations (5.5), (5.6) and (5.7) $\quad \{p \in [0, np]\} \quad \{t \in [0, nt]\}$

5.2.2 Heat Loss Calculation

An energy flow from the hotter tank to the colder environment through the storage wall is called heat loss. Energy heat losses are modeled as a mass flow from a higher to the next lower level. They depend on the volume of the virtual tank calculated over the current mass $\boldsymbol{M_{v,p,t}}$ divided by its density ρ and multiplied by the ratio of the diameter d to the surface $(4/d)$ for obtaining the surface area in contact with the environment.

$$\boldsymbol{A_{storage\ wall,v}} = \frac{\boldsymbol{M_{v,p,t}}}{\rho} \cdot \frac{4}{d} \tag{5.8}$$

The lid losses are neglected in this model, because no heat transfer is assumed to happen between the virtual tanks. The top lid losses can be included in κ_{hl}, the tank specific heat loss coefficient.

The heat losses to the environment for each temperature level are calculated by (5.9) with the difference of the current temperature level T_{v+1} to the ambient temperature T_a (which is equal to $T_{v=1}$). The mass flow representing the heat losses v uses the same convention as the mass flows in the mass balance: it comes from the level $v+1$ to the level v. Therefore the temperature of the level $v+1$ is taken as reference.

$\boldsymbol{A_{storage\ wall,v}}$ represents the surface of storage of wall in contact with the ambient temperature T_a. The heat loss is related to a mass flow by Equation (5.10).

$$\dot{\boldsymbol{Q}}_{\boldsymbol{hl,v,p,t}} = \kappa_{hl} \cdot (T_{v+1} - T_a) \cdot \boldsymbol{A_{storage\ wall,v}} \tag{5.9}$$

$$\dot{\boldsymbol{M}}_{\boldsymbol{hl,v,p,t}} = \frac{\dot{\boldsymbol{Q}}_{\boldsymbol{hl,v,p,t}}}{c_p \cdot (T_{v+1} - T_v)} \tag{5.10}$$

$$\{v \in [1, nv-1]\} \qquad \{p \in [1, np]\} \qquad \{t \in [1, nt]\}$$

As shown in Figure 5.1, the losses of tank level 5 are the energy flux $\dot{\boldsymbol{Q}}_{\boldsymbol{hl,4}}$ and equal to the mass flow $\dot{\boldsymbol{M}}_{\boldsymbol{hl,4}}$. The level $v = 1$ is considered to be at ambient temperature and therefore has no losses.

The streams $\dot{\boldsymbol{Q}}_{\boldsymbol{hl,v,p,t}}$ do not appear as process streams within the heat cascade in order to avoid that utilities are required at times of heat scarcity (e.g. during the winter). This is an change compared to the existing model. In contrast to [*Becker*, 2012, pp.131-132]'s storage model, the losses are modeled as a mass flow from a higher temperature level to the next lower one instead of a cold stream into the heat cascade that needs to be heated. For an industrial process, keeping the temperature level might be a requirement. In an urban energy system, uncompensated losses are part of the storage use. Her approach has the drawback of requiring heat power in every period compensating the losses, when losses appear. Losses appear, when the virtual tanks that have a higher temperature than the surrounding environment are filled. For example, an optimization of an only solar powered house with a storage would not converge, because periods exist where no utility or no sun power is available compensating the heat losses.

The heat loss factor κ_{hl} is estimated to be around 1 $W/(m^2 \cdot K)$ by *Becker* [2012, p.131] and 1.3 $W/(m^2 \cdot K)$ *Angrisani et al.* [2014] for daily storage. *Angrisani et al.* [2014] assume a higher heat loss factor κ_{hl}, because they state that the insulation of tank is rarely bigger than 10 cm and often the bottom is not insulated.

Exergy losses such as losses due to de-stratification are not considered. They can be neglected, if charging and discharging of the storage are done avoiding turbulent flows through different temperature levels in the storage tank. *Cruickshank and Harrison* [2010] confirm this finding by stating that an average u-value for loss calculation can be used if it accurately represents the unit as installed.

5.2.3 Cyclic Constraint

The cyclic constraints of the daily storage model optimizes the storage usage because it can choose which level has how much initial mass. It imposes that the conditions in the virtual

tanks at the beginning of the period p are identical to the conditions at the end of the period. Therefore the energy balance is fulfilled. This is expressed by:

$$\boldsymbol{M}_{v,p,t=0} = \boldsymbol{M}_{v,p,t=nt(p)} \qquad \forall v,\ \forall p \tag{5.11}$$

$$\boldsymbol{M}_{\boldsymbol{v,p,t=0}} \geq 0 \tag{5.12}$$

$$\text{for (5.11)} : \{v \in [1, nv]\} \qquad \{p \in [0, np]\} \qquad \{t \in [0, nt]\}$$

The cyclic constraint (5.11) ensures that the storage returns into its initial state at $p, t = 0$ at the end of each respective cycle: The storage's discharge mass flows over time including the heat losses equal the charge rate. The model sets the initial state $\boldsymbol{M}_{\boldsymbol{v,0,0}}$ leading to one more degrees of freedom.

This formulation avoids fixing the initial levels while having an optimal storage usage. Furthermore, the problem is independent of its starting point in time. The daily storage can not pass energy to the next period. It works only over all of the time slices of one period.

5.2.4 Summarizing the daily Storage Model

With the adaptation of the storage to the requirements for the urban energy system design, a new daily storage model has been developed. Compared to the existing model, losses are modeled differently and it is an integer-free model (Becker's implementation contained integers for every stream).

When addressing solar energy integration, the model should have the possibility to store harvested energy longer than only a day to address seasonal load shifts. A second model, based on the daily model, is now developed to fit the needs for estimating a long term storage that can pass energy also from one period to another.

5.3 Long term thermal Energy Storage Model

Based on the clustering approach discussed in Chapter 4, a seasonal storage requires a sequence of typical days to pass energy from one period to another. The development of this model is based on the daily storage model. Therefore in the following, the explanations of the already presented equations is kept short, only differences to the daily model are described.

5.3.1 Mass Balance

A weighting factor ω for each period p is introduced, considering that each period repeats itself a number of times. In order to transfer heat loads from one period to a another, the periods must be in a sequence representing their appearance.

Based on the previously discussed mass balance, a new mass balance for the seasonal storage model is developed with ω_p:

$$\text{For } \{v = 1\}:$$
$$\boldsymbol{M_{1,p+1,t}} = \boldsymbol{M_{1,p,t}} + \sum_{p=1}^{np} \omega_p \left[\sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{h,1,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{hl,1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{c,1,p,t}} \right) \right] \tag{5.13}$$

$$\text{For } \{v \in [2, nv-1]\}:$$
$$\boldsymbol{M_{v,p+1,t}} = \boldsymbol{M_{v,p,t}} + \sum_{p=1}^{np} \omega_p \left[\sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{c,v-1,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{h,v,p,t}} + \dot{\boldsymbol{M}}_{\boldsymbol{hl,v,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{c,v,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{h,v-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{hl,v-1,p,t}} \right) \right] \tag{5.14}$$

$$\text{For } \{v = nv\}:$$
$$\boldsymbol{M_{nv,p+1,t}} = \boldsymbol{M_{nv,p,t}} + \sum_{p=1}^{np} \omega_p \left[\sum_{t=1}^{nt} d_{p,t} \cdot \left(\dot{\boldsymbol{M}}_{\boldsymbol{c,nv-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{h,nv-1,p,t}} - \dot{\boldsymbol{M}}_{\boldsymbol{hl,nv-1,p,t}} \right) \right] \tag{5.15}$$

For Equations (5.5), (5.6) and (5.7) $\quad \{p \in [0, np]\} \quad \{t \in [0, nt]\}$.

With the weighting factor ω_p the weight of all periods, their number of appearances, is taken into consideration to estimate a long term storage. When charging the storage in a given time slice t of a period p, it is charged ω_p times. The losses are calculated directly as a function of the current mass reparation in the storage and storage losses appear. This can be interpreted as a pessimistic estimate, because the losses appear directly. When taking the example of an only solar powered house with a storage, the storage would be charged during the day and loose a little bit over the night. During the harvesting period, the storage would then continue to charge during the day again, however more or less from the level where it stopped the day before. The temperature level rises slowly. With this rise of the maximal storage temperature, the losses during the end of the charging period increases and not from the beginning on.

In this model this sequence of charging events is estimated with the weighting factor ω_p.

5.3.2 Cyclic Constraint

The cyclic constraint imposes that the conditions in the tanks before the beginning of the first period are identical to the conditions at the end of all periods pn. This is expressed by Equation (5.16), which introduces a zeroth period before all other periods to choose the initial conditions. Equation (5.17) ensures that this state cannot contain values smaller than zero.

$$\boldsymbol{M}_{v,p=0,t} = \boldsymbol{M}_{v,p=np,t=nt} \qquad \forall v \tag{5.16}$$

$$\boldsymbol{M_{v,p=0,t}} \geq 0 \tag{5.17}$$

for (5.16) : $\{v \in [1, nv]\} \quad \{p \in [0, np]\} \quad \{t \in [0, nt]\}$

Again, the optimal storage use is estimated through adding the initial state of the storage as a variable to the optimization problem.

5.3.3 Heat Loss Calculation

The heat losses are also based on the approach used for the daily storage. It is however adapted to represent the lid losses correctly.

An energy flow from the hotter tank to the colder environment is called heat loss. Energy heat losses are modeled as a mass flow from a higher to the next lower level. The heat losses to the environment for each temperature level are calculated by (5.18) with the difference of the current temperature level T_v to the ambient temperature T_a. κ_{hl} is a tank specific heat loss coefficient, A_v represents the surface of storage of wall in contact with the ambient temperature T_a. The heat loss is converted into a mass flow by Equation (5.10) assuming a constant specific heat capacity for the given temperature interval.

$$\dot{\boldsymbol{Q}}_{\boldsymbol{hl,v,p,t}} = \kappa_{hl} \cdot (T_{v+1} - T_a) \cdot \boldsymbol{A_v} \qquad \{v \in [1, nv-1]\} \qquad \{p \in [1, np]\} \qquad \{t \in [1, nt]\} \tag{5.18}$$

$$\dot{\boldsymbol{M}}_{\boldsymbol{hl,v,p,t}} = \frac{\dot{\boldsymbol{Q}}_{\boldsymbol{hl,v,p,t}}}{c_p \cdot (T_{v+1} - T_v)} \qquad \{v \in [1, nv-1]\} \qquad \{p \in [1, np]\} \qquad \{t \in [1, nt]\} \tag{5.19}$$

Recent seasonal storage using expanded glass granulate or foam glass gravel for insulation material show heat loss coefficients that are well below the values of 1 $W/(K \cdot m^2)$ from the daily storage. The change in insulation material compared to earlier installations ensures that the thermal properties remain in an acceptable range when the maximal storage temperature is reached or when water moistens the insulation with an humidity up to 30 %. [*Ochs et al.*, 2008] gives the effective thermal conductivity for temperature and humidity ranges, [*Ochs and Müller-Steinhagen*, 2008, p.127, p.184] shows construction of typical seasonal storage. The heat losses only depend on the average storage insulation thickness. With a mean thickness of 0.75 m values between $0.15 - 0.6$ $W/(m^2 \cdot K)$ can be reached, when the insulation's water content is kept below 50 % and can dry off again. Alternatively, the heat loss coefficient κ_{hl} can also be expressed as a time dependent heat loss $\kappa_{hl}(t)$, when for example rain is integrated into the model increasing the heat losses due to humid insulation.

It is important to notice that besides via the wall surface, heat can also be lost at the bottom and the top of the storage depending on the charging level. Often the bottom is not considered neither isolated ([*Ochs*, 2010, p.46] or [*Angrisani et al.*, 2014]) leading to much higher losses than estimated theoretically.

A set of binary variables y per level v activates the additional surface area of the storage at the top $y_{top,v}$ and the bottom $y_{bottom,v}$. This is particularly important for the seasonal storage

due to the storage duration. For every time slice, the surface area of each virtual tank is recalculated:

$$\boldsymbol{A_{storage\ wall,v}} = \frac{\boldsymbol{M_{v,p,t}} \cdot 4}{\rho \cdot d} \tag{5.20}$$

$$A_{lid,v} = A_{top,v} = A_{bottom,v} = \frac{\pi \cdot d^2}{4} \tag{5.21}$$

$$\boldsymbol{A_v} = \boldsymbol{A_{storage\ wall,v}} + \left(\boldsymbol{y_{top,v}} + \boldsymbol{y_{bottom,v}}\right) \cdot A_{lid} \tag{5.22}$$

$$\{v \in [1, nv]\} \qquad \{p \in [1, np]\} \qquad \{t \in [1, nt]\} \qquad y_{top \vee bottom} \in \{0,1\} \forall\ v \geq 2$$

Equation (5.22) adds two binary variables to each virtual tank with a temperature higher than T_a. The additional surface should only considered once for whole storage tank and only if heat is stored in it and the virtual tank's temperature is higher than the ambient temperature T_a. Additional constraints are required to ensure that only the lowest temperature level higher than T_a from the bottom on filled with heat (Equation(5.24)) has active heat losses.. The same is done for the top temperature level (Equation (5.23)).

$$\boldsymbol{y_{top,v,p,t}} \geq \frac{\boldsymbol{M_{v,p,t}}}{M_{max}} - \left(\sum_{i=v+1}^{nv} \boldsymbol{y_{top,i,p,t}}\right) \tag{5.23}$$

$$\boldsymbol{y_{bottom,v,p,t}} \geq \frac{\boldsymbol{M_{v,p,t}}}{M_{max}} - \left(\sum_{i=1}^{v-1} \boldsymbol{y_{bottom,i,p,t}}\right) \tag{5.24}$$

$$\sum \boldsymbol{y_{bottom,v,p,t}} = 1 \qquad \sum \boldsymbol{y_{top,v,p,t}} = 1 \tag{5.25}$$

$$\{v \in [1, nv]\} \qquad \{p \in [1, np]\} \qquad \{t \in [1, nt]\} \qquad y_{top \vee bottom} \in \{0,1\} \forall\ v \geq 2$$

The constraints (5.25) guarantees that only one binary heat loss variable y_{bottom} and y_{top} has a value equal to one in every period and time slice.

The binary lid heat loss variables y_{bottom} and y_{top} are always equal to one in every period and time slice. The equations (5.23) for the top lid and (5.24) for the bottom lid choose the correct higher or lower level according to the current storage filling. Choosing the value M_{max} as divisor ensures that the divisor is a true upper bound of the equation. This division can at maximum become equal to one. For the case where the two highest tanks contain each half of the total mass, the sum of binaries y_{top} containing only the tanks with higher temperatures than themselves ensures that only y_{top} for the highest temperature tank is one. A too high value for the divisor creates numerical problems because the division produces always results close to zero, e.g. when the maximal storage volume is very big and the content of the highest filled virtual tank is close to zero. Then the constraints for all tanks are the same, making it impossible to choose. Depending on the numerical precision used in the solver, this might become a problem. [*Klotz and Newman*, 2013, Section 3.4] and [*Williams*, 2013, page 166] confirm this. They suggest looking whether the modeling language and the solver support

writing indicator variables directly as such instead of using "big M approach". If yes, the advantage is that the branch and bound can use direct branching on the node avoiding the potential numerical trouble from the "big M approach".

Through the modeling of the lid losses, the model is an MILP model. The losses introduce per virtual tank two binary variable per time slice. When keeping the number of periods and time slices low, this formulation does not pose any problem. Increasing the number of time slices leads quickly to much longer computational times due to the increase of integer variables. Or changing the data can have the same effect as described by [*Fischetti and Monaci*, 2014] due to the fact that the solver starts with a heuristic at a different point. When a different point is found as a starting point, the solution time of a large model might significantly change. The model can also be simplified again to have the same loss calculation as the daily storage.

5.4 Storage Costs and Cumulative Exergy Demand LCIA Values

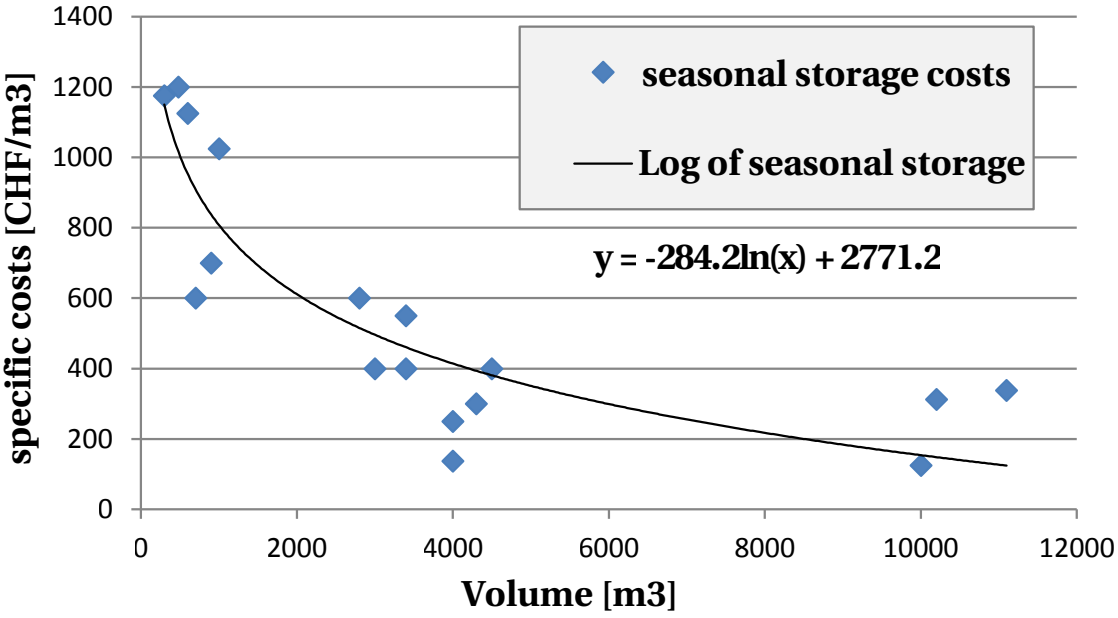


Figure 5.2 – Cost function for seasonal storage based on CostDBCREM [*Poumadère et al.*, 2015]

During this work, cost or Cumulative Exergy Demand (CExD) values are used in the objective function. For the costs, the cost data base of the CREM, CostDBCREM[1] [*Poumadère et al.*, 2015], is used: It provides capital and operational expenditures for technologies based on projects realized in Switzerland. Figure 5.2 shows the exponential cost function for the seasonal storage. With the total mass of the storage M, the specific investment costs can be calculated:

$$inv_{storage} = -284.2 \cdot log(M) + 2771.2\,. \tag{5.26}$$

[1] A data base established by the CREM, a research institution on urban energy research, containing cost data sets with various technologies of project realized in Switzerland.

As a common hypothesis, the capital expenditure value has a yearly annuity of 0.078 based on 6 % interest rate and 25 equipment lifetime.

$$\boldsymbol{Inv_{storage}} = \boldsymbol{f_{storage}} \cdot inv_{storage} \tag{5.27}$$

Table 5.1 – Storage units added to the optimization problem and CExD of their construction ([1] from [*Bösch et al.*, 2007], [2] extrapolation according to [*Gerber et al.*, 2011])

Type	Capacity	CExD $[kWh-Eq./m^3]$	Inv. Costs $[CHF/m^3]$
domestic hot water storage, steel tank[1]	0.6 m^3	2 004	2 300
hot water storage, steel tank[2]	200 m^3	859	985
seasonal heat storage, pebble/water tank[2]	10 000 m^3	289	332

For the optimization, a piece-wise linear function estimates the installation costs or CExD, extrapolating from the closest data point. *Hahne* [1999] compares different large scale storage technologies in terms of cumulative energy demand: he shows that in terms of accumulative energy demand (predecessor of the cumulative energy demand) for large storage systems with the size around 10 000 m^3, the production impact for concrete compared to pebble/water or vertical duct storage is up to a factor of 3 higher. The operational energy consumption is estimated well below 1%, which means neglectable compared to the incertitude of the input data.

5.5 Parameter Overview of the two Storage Models

Two storage formulations are proposed: the so-called daily storage is a LP model, well adapted for storage integration over shorter time periods or for the use in big scale problems. The seasonal storage is an MILP model using a more precise heat loss calculation at the price of introducing integers into the formulation. When the storage size is known and well bounded or the problem is of smaller scale, this formulation can be used. The values of Table 5.2 are based on [*Ochs*, 2010, Appendix A11 and Chapter 3], when not indicated otherwise.

5.6 Conclusion of Storage Model Development

Based on [*Becker*, 2012, pp126]'s time slice model, a new storage model is implemented. The same restrictions apply as cited by [*Becker*, 2012, pp128]:

- Only a sensible heat storage is considered. The temperature difference, the mass flow and the specific heat capacity define therefore the heat load.

Table 5.2 – Overview of key parameters for the storage model

Parameter	Storage Type		Si-Unit
	Daily	Seasonal	
Storage Medium	Water	Water	-
Medium Density	997.5		kg/m^3
Heat Capacity of Medium	4.18		$kJ/(kg \cdot K)$
Thermal Conductivity of Storage Medium	0.63		$W/(m \cdot K)$
Specific Heat Loss Rate	1.3 [*Angrisani et al.*, 2014]	0.65	$W/(m^2 \cdot K)$
Insulation Thickness	0.1	≥ 0.3	m
Size: Order of Magnitude	from 1 up to 10 000		m^3
Height Diameter Ratio	3	1	-
Annual Cycles	$260-280$	$1.2-1.6$	-
Geometry	Cylindrical	Cylindrical to inverse pyramids	-
Discrete Temperature Levels	9	9	-
ΔT per Temperature Level	10	10	K
Temperature Range	$10-95$	$10-95$	$°C$
Investment Costs, Table 5.1	≥ 1 m^3: 2300	10 000 m^3: 332	CHF/m^3
Construction CExD, Table 5.1	2004	289	$kWh-Eq./m^3$

- Each temperature in the storage tank is discretized into a pre-defined number of temperature levels or virtual storage tanks at a fixed temperature. These temperature levels correspond to a virtual tank, that has its own mass balance and based on its temperature level and the reference temperature level that is used to calculate its own losses.
- A maximal storage volume is defined as the sum over all virtual storage tanks.
- An initial level of the storage tank is considered and is treated using a cyclic constraints in the model.
- The cost of the storage tank is considered using an exponential based formula based on the size of the storage tank.
- The storage geometry of the storage are chosen to assure neglectable destratification losses, the height is much bigger than the diameter [*Ochs*, 2010, p.25-28].

The differences to the Becker's model are:

- In case of a daily storage, the difference between day p end of day storage filling level and day $p+1$ starting level is neglected as the days are considered to be independent and cyclic repetition of each day is high and therefore more important. (Transition periods could be inserted to deal with this problem.)
- Heat losses are estimated as a function of the reference temperature. The storage lid losses based on the current storage temperature are calculated.
- Heat losses are not integrated into the heat cascade.
- The daily storage is written as an linear problem with only one integer variable deciding over the storage existence. Becker's version has two integer variables per unit and time slice.

In contrast to [*Becker*, 2012, pp.131-132]'s storage model, the losses are modeled as a mass flow from a higher temperature level to the next lower one instead of a cold stream into the heat cascade that needs to be heated up from the utilities. The lowest level is at atmospheric temperature leading to no further losses. For an industrial process, keeping the temperature level is a requirement. In an urban energy system, uncompensated losses are part of the storage use. Therefore her approach is modified to consider uncompensated losses.

The model could also integrate the heat losses in case they could be used for example to heat up a building. The lowest temperature level of storage is always at ambient or reference temperature level. It therefore has no losses anymore as the temperature difference is zero. Furthermore, instead of using a simple factor for the lid losses, they are calculated as a function of the currently highest temperature level for the upper and lower lid in case of the seasonal storage. For a daily storage, these losses are simplified with a heat loss factor as in Becker's work.

Limiting the maximal storage mass $\boldsymbol{M_{max}}$ bounds the variable and brings often important performance improvements. Here, the ground surface of the storage is limited ensuring a favorable geometry for neglecting the stratification losses. If the maximal height is given, the ground surface of the storage tank can be calculated with an additional constraint ensuring a favorable height to ground surface ratio. [*Ochs*, 2010, p.27] suggests a height diameter ratio for short term (buffer) storage around 3.0 with a cylindrical geometry and for a seasonal storage around 1.0 with geometries of cuboids, cylinders, inverse (and truncated) pyramids or cones.

5.7 Solar Thermal Collector Model

Renewable energy integration is a key challenge for today's energy system design. With the help of the previously discussed storage models, a stochastic resource such as solar energy can be integrated into a system, because the storage allows to decouple demand and supply by dephasing it.

For the integration of stochastically available renewable energy such as solar energy, a mixed integer linear programming (MILP) model with different temperature levels is developed. The literature review revealed that solar energy often used at an already fixed operating temperature, e.g. for hot water production and sized by best practice rules. In commercial software, only a simulation of different operational modes with different installation sizes is possible: the user needs to guess the best use of the solar energy and how much to install. The installation size and choice of an operating temperature have a major impact on the total solar yield.

Based on the standard approach by [*Duffie and Beckman*, 2013, Chapter 6], the Institut für Solartechnik (SPF) created a database with available and tested collectors based on the methodology of [*Brunold et al.*, 1994]. A simplified solar thermal collector is modeled considering the following points:

- Localization with horizon if given and therefore the local irradiation,
- the available roof area,
- the roof (or surface) orientation,
- optical and thermal efficiency of a given solar collector.

Estimating solar Heat Power

Based on Equation (6.17.7) in [*Duffie and Beckman*, 2013, Chapter 6.17], the efficiency of a collector can be estimated with a small error as presented in Equation (5.28):

$$\eta = \eta_0 - a_1 \cdot \frac{(T_m - T_a)}{G_n} - a_2 \cdot \frac{(T_m - T_a)^2}{G_n} = \frac{\dot{m} \cdot c_p \cdot (T_{outflow} - T_{inflow})}{A_c \cdot G_n} \tag{5.28}$$

The error comes from the fact that the temperature difference in the fluid is calculated as an arithmetic mean, even though a logarithmic temperature difference should have been used. According to [*Duffie and Beckman*, 2013, Chapter 6.17], Cooper and Dunkle (1975) state that this approximation yields very small errors for the practical design of the collector. An increasing temperature difference between the collector's mean fluid temperature T_m and the outside temperature T_a leads to higher losses. A higher mass flow leads to a lower temperature increase of the fluid. The measurements of the SPF use the same approximation for measuring the parameters η_0, a_1 and a_2 that can be found in the SPF's database together with an estimate of the incident angle modifier IAM. The incident angle modifier IAM regroups all optical losses and is usually given at an angle $\Theta = 50°$. a represents the measured coefficient for the given collector.

$$IAM = 1 - tan^a(\frac{\Theta}{2}) \tag{5.29}$$

G_n in Equation (5.28) is the global normal radiation on a surface, a sum of the beam, diffuse and reflected radiation.

$$G_n = I_{beam,n} + I_{diff,n} + I_{refl,n} \tag{5.30}$$

With the Equations (5.28), (5.29) and (5.30), the efficiency η_{eff} for each time interval and radiation type can be calculated:

$$\eta_{eff} = \eta \cdot IAM \tag{5.31}$$

Based on η_{eff}, the solar thermal heat production for different temperature levels can be estimated:

$$q_{avail.\ solar} = I_{beam,n} \cdot \eta_{eff,beam} + I_{diff,n} \cdot \eta_{eff,diff} + I_{refl,n} \cdot \eta_{eff,refl} \tag{5.32}$$

The result $q_{avail.\ solar}$ of this model depends on the mass flow and the temperature difference of the collector's fluid on given meteorological condition with a given outside ambient temperature T_a and a global irradiation G_n on a roof top (or surface).

Integration of the solar Panel in the Energy System

The solar thermal panel introduces a hot stream to be considered to supply the heat to the energy system. The solar collector is calculated considering a discretized temperature at which the solar collector is operated. The temperature defines the initial and target temperature for the operation of the solar collector. For these temperatures, the mean temperature T_M is calculated and the maximum production of the solar panel is calculated with respect to the area of the panel. In our model, we allow for a by-pass of the solar panel if the heat available is higher then the one that is needed by the system. This model chooses the optimal overall solar thermal collector installation size and its operating temperature for each time slice using mathematical programming.

Per time slice p, t, the solar power $\dot{\boldsymbol{Q}}_{\boldsymbol{solar,p,t}}$ is calculated through the available specific solar power $\dot{q}_{avail.\ solar,p,t}$ per square meter for each given temperature interval of collector in- and out-flow temperatures using Equation (5.28) and the surface factor $\boldsymbol{f}_{\boldsymbol{solar,p,t}}$ limited between zero and the maximal available surface. The total solar yield represents the sum over all active solar streams $1, \ldots, nsolar$ (5.33), each representing a couple of one in- and one out-flow temperature.

$$\dot{\boldsymbol{Q}}_{\boldsymbol{solar,p,t}} = \sum_{solar=1}^{nsolar} \boldsymbol{f}_{\boldsymbol{solar,p,t}} \cdot \dot{q}_{avail.\ solar,p,t} \tag{5.33}$$

Two implementation are proposed: In the first one, only one temperature level can be activated per time slice, constraining the number of factors $\boldsymbol{f}_{\boldsymbol{solar,p,t}}$ with a value bigger than zero to 1. $\boldsymbol{y}_{\boldsymbol{solar,p,t}}$ is an integer variables that activates the temperature level (solar) is used ($\boldsymbol{y}_{\boldsymbol{solar,p,t}} = 1$) in the solar panel during the time slice t and the period p or not ($\boldsymbol{y}_{\boldsymbol{solar,p,t}} = 0$), where $\boldsymbol{y}_{\boldsymbol{solar,p,t}}$ is the fraction of the solar heat available that is used during the time slice t of period p:

$$\sum_{solar=1}^{nsolar} \boldsymbol{y}_{\boldsymbol{solar,p,t}} \leq 1 \tag{5.34}$$
$$\forall p = 1.., np \quad \forall t = 1.., nt$$

$$\boldsymbol{f}_{\boldsymbol{solar,p,t}} \leq A_{roof} \cdot \boldsymbol{y}_{\boldsymbol{solar,p,t}} \tag{5.35}$$
$$\forall solar = 1.., nsolar \quad \forall p = 1.., np \quad \forall t = 1.., nt$$

The Equation (5.34) allows to activate at maximum one (or any other number of) temperature level for the solar panel. (5.35) limits the production to the maximal available roof size.

The second formulation does not use Equation (5.34) and omits $\boldsymbol{y_{solar,p,t}}$ in Equation (5.35), leading to reduction of binary variables by a factor of number of $solar$ temperature levels times total time slices p, t. If in addition, the solar module is available in all time steps, the Equation (5.35) can be omitted by limiting the multiplication factor $\boldsymbol{f_{solar,p,t}}$ directly and limiting the overall solar sizing factor $\boldsymbol{s_{solar}}$ to the maximal available roof size:

$$\boldsymbol{s_{solar}} \geq \sum_{solar=1}^{nsolar} \boldsymbol{f_{solar,p,t}} \tag{5.36}$$

with: $0 \leq \boldsymbol{s_{solar}} \leq A_{roof}$.

The model produces heat at different temperatures during one time slice. The described model simplifications can be made from an operational point of view of a solar collector field when:

- different independent loops exist allowing to produce heat at different temperature levels at the same time or
- the streams at different temperature level are interpreted as streams appearing at different time intervals but within the same time slice.

From the mathematical programming point of view, the model becomes an LP model instead of an MILP model.

Both formulations are linked to cost constraints by the overall sizing factor $\boldsymbol{s_{solar}}$ which is the maximum of the time dependent sizing factors $\boldsymbol{f_{solar,p,t}}$(Equation (5.37)):

$$\boldsymbol{s_{solar}} \geq \boldsymbol{f_{solar,p,t}} \tag{5.37}$$

$$\boldsymbol{INV_{solar,total}} = \boldsymbol{s_{solar}} \cdot inv_{solar} \tag{5.38}$$

$$\boldsymbol{OP_{solar,total}} = \sum_{solar=1}^{nsolar} \sum_{p=1}^{np} \sum_{t=1}^{nt} \boldsymbol{f_{solar,p,t}} \cdot op_{solar} \cdot \dot{q}_{avail.\ solar,p,t} \ . \tag{5.39}$$

The investment for the solar panels are considered as being proportional linearly to the area of solar panels installed. The operating and installation CExD values of Table 5.3 are used in Equations (5.38) and (5.39). The costing data in the same Table 5.3 is extracted from CostD-BCREM[2], a data base, containing costing data sets from projects realized and constructed in Switzerland. The economic model developed enables to calculate capital expenditures (CAPEX) and operational expenditures (OPEX) associated with each technology. It was developed based on the feedback from various industry partners such as heating technology manufacturers, district heating network operators or engineering firms in Switzerland. As much as possible, parameters are used to be able to have a flexible cost database.

[2]A data base established by the CREM, a research institution on urban energy research, containing cost data sets with various technologies of project realized in Switzerland.

Table 5.3 – Non renewable CExD *Dones et al.* [2007] and cost data based on CostDBCREM[1] [*Poumadère et al.*, 2015] for solar thermal Collectors.

Mode	CExD	Costs
Operating	0.03 $\frac{kWh-Eq.}{kWh}$	0 $\frac{CHF}{kWh}$
Investment ($\geq 200m^2$)	290 $\frac{kWh-Eq.}{m^2}$	280 $\frac{CHF}{m^2}$

The presented models for solar thermal panels can be used to size the solar thermal installation together with storage. It finds the optimal operational strategy in terms of temperature level and heat capture rate for thermal needs.

5.8 Connecting Storage to Utilities: Stratification and Choice of discrete Temperature Levels

The developed models address the shortcoming found in the literature by taking the stratification of thermal energy storage into account as well as the heat losses. The solar model offers power at different discrete temperature levels. Therefore the temperature levels of the storage and the solar thermal collectors are chosen together, ensuring that the solar energy can be stored at the corresponding temperature level maximizing the solar yield. From the computational point of view, the number of discrete levels should be kept as low as possible to keep calculation time lower.

The solar model uses steps of 10 K over the range of 20 °C to 100 °C. All combinations are introduced as well, e.g. $20-40$, $20-60$, $40-60$ and so on. A heat pump uses then the same intervals, but is limited to 80 K. The storage uses the same intervals of 10 K.

This definition avoids that one unit is chosen over another because a finer choice is available that can adapt better to the demand. When a compromise is made between the number of discrete temperature levels, besides the solar collector model, the uncertainty of the energy demand temperature should be taken into consideration. Only very few measurements of the distribution temperatures within heating systems exist.

5.9 Case Study

For this demonstration case study, the models previously discussed can help local municipalities, engineers and utility providers in finding optimal solutions addressing the new nation-wide strategies. At a (pre-)design level, they proposes appropriate technologies and energy infrastructure while showing the impacts on the existing infrastructure (if there is any).

The community of Verbier is constructing a new district heating system. In mountain resorts such as Verbier, Switzerland, the population of people can raise up to 40 000 during the skiing

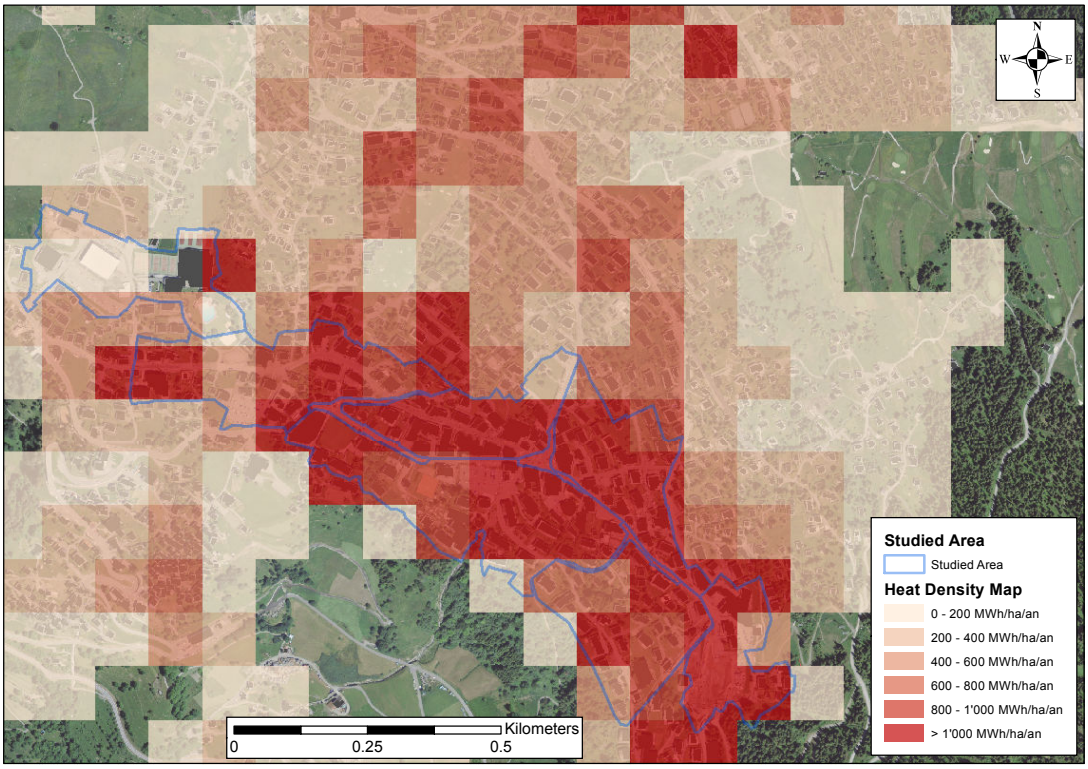


Figure 5.3 – Heat density per block of 100 ∗ 100 m (1 hectare) for Verbier, Switzerland

holiday season around Christmas, in February break or during spring break. However, only 3 000 persons live in Verbier all year long [*Commune de Bagne*, 2014]. Buildings with only seasonal occupation still may use a direct electrical heater, increasing the need for electricity import during the winter.

The tool Planeter [*Blanc et al.*, 2013] uses all available information about the buildings in the area to give an overview about the current situation. Figure 5.3 is one of Planeter's results. It is based on [*Girardin et al.*, 2010] extended with the feedback from the everyday usage with clients. Building registry information is combined with consumption data from the local utility. Together with the Swiss building regulation a complete picture of the each building can be drawn. When no daily or monthly measurements are available, which so far has almost always been the case; the tool provides the necessary input data to run dynamic simulation software. Simulating building dynamically helps understanding the problem of the seasonality. In addition to the building's current energy demand, a refurbished state is estimated based on current Swiss standards. The local available (renewable) energy sources are mapped. For heat pumps the available sources are evaluated and mapped. Available surfaces for the installation of solar thermal collectors are sorted with respect to their orientation to integrate the solar potential of the area.

A dynamic building simulation software, bSol [*Page et al.*, 2014], fed with parameters from the previous step, evaluates the heat load over a given meteorological year for each building with an hourly resolution. Furthermore, retrofitting actions can be simulated to assess a scenario for each building with decreased heat demand. bSol helps to evaluate the impact of hypothesis such as temperature set-points or building envelope's u-values that completed the data sets. Thanks to increasing computational power and the simplicity of bSol, a lot of evaluations can be run quickly allowing input parameters to be tested systematically, ensuring an appropriate set of input parameters.

Verbier does not have a district heating system yet, a main branch for the district heating system is being installed during 2015 as well as several wood fired or oil boilers. About 300 buildings, which are according to the community the most likely to be linked to the district heating system, consume according to a dynamic building simulation about 24.4 GWh of heat annually. The peak load is 13.3 MW in the winter.

The dynamic building simulation calculates the heat demand throughout the year. This input is reduced to 12 key days representing the year, one for every month. A temporal k-medoids clustering, see Chapter 4, is used to identify them, respecting the annual energy and power balance. January and February are regrouped to one month, because they have a very similar profile. Then an extreme period is inserted at the place of January.

For the operating conditions, the hydronic model from [*Girardin et al.*, 2010] calculates the supply and return temperature within the heating system using the design temperatures of the building's heating system provided by Planeter. It assumes a constant mass flow. The nominal mass flow $\dot{m}_0$ multiplied by the specific heat capacity c_p equals the nominal heating power $\dot{Q}_0$ divided by the nominal temperature difference between supply and return temperatures $T_{supply,0}$ and $T_{return,0}$. The supply temperature $T_{supply,p,t}$ in the period p and time slice t is then calculated as a function of the internal temperature T_{int} and the ratio of the current heat load $\dot{Q}_{p,t}$ over the nominal one, multiplied with the nominal temperature difference $T_{supply,0} - T_{return,0}$ and a factor representing the heat exchange. Subtracting the ration out of current heat load $\dot{Q}_{p,t}$ and product of nominal mass flow $\dot{m}_0$ and specific heat capacity c_p, the supply temperature $T_{supply,p,t}$ is obtained.

$$\dot{m}_0 c_p = \frac{\dot{Q}_0}{T_{supply,0} - T_{return,0}} \tag{5.40}$$

$$T_{supply,p,t} = T_{int} - \frac{\dot{Q}_{p,t}}{\dot{Q}_0} \cdot (T_{supply,0} - T_{return,0}) \cdot \frac{\alpha}{1-\alpha} \tag{5.41}$$

$$\text{with } \alpha = \frac{T_{supply,0} - T_{int}}{T_{return,0} - T_{int}} = exp(\frac{U_0 A_0}{\dot{m}_0 c_p})$$

$$T_{return,p,t} = T_{supply,p,t} - \frac{\dot{Q}_{p,t}}{\dot{m}_0 c_p} \tag{5.42}$$

The solar thermal heat production is calculated for discrete collector temperatures with the approach of 5.7 using data for a typical collector used in Switzerland. Figure 5.4 shows the annual solar thermal potential of Verbier's roofs.

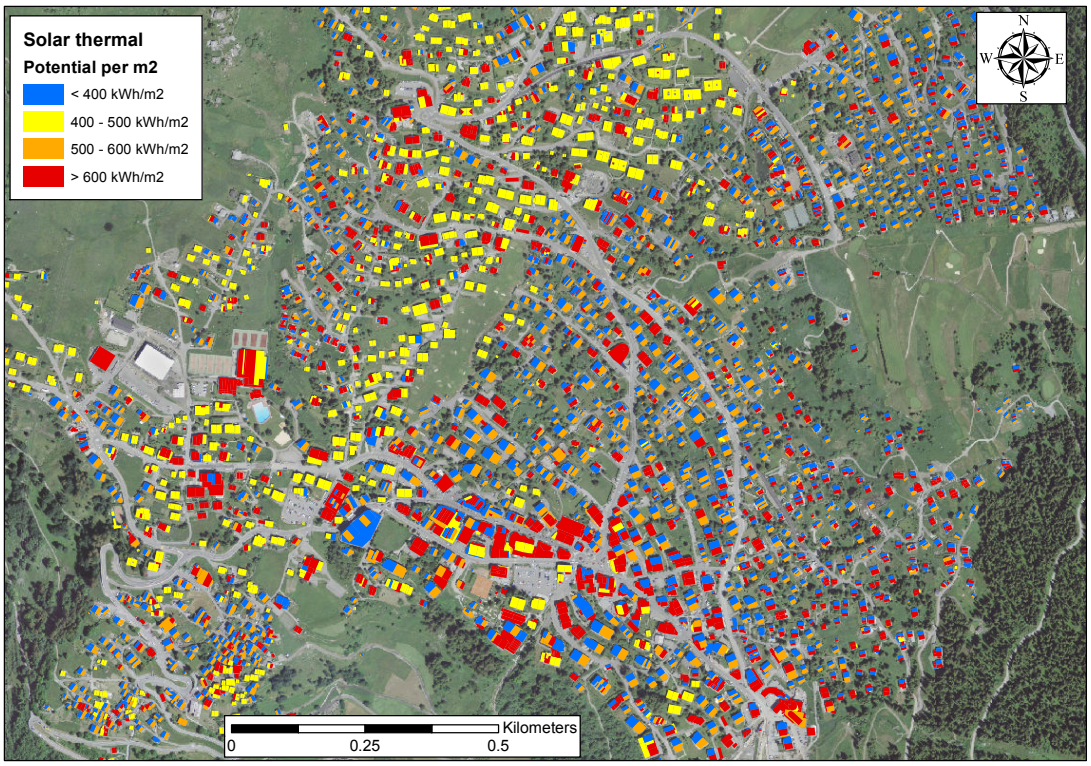


Figure 5.4 – Annual solar thermal production potential per m^2 of roof.

The examples shown do not cover all superstructure combinations of technologies available for the heat supply neither all possible integration possibilities. Here, the model is used a screening tool on community level, that can quickly give an overview pointing towards the interesting solutions. Energy integration can integrate more technologies and also integrate with existing heat sources during a project while knowledge of the area is increasing. For the superstructure choice existing algorithms such as [*Voll*, 2013] can be chosen.

For this case study, it is important to show the comparison of the different storage options for the energy system design.

5.9.1 District Heating System

The piping of an district heating system connects the energy demand, (renewable) energy resources and the large thermal heat storage. When using energy integration, either the network streams can be introduced in between demand and supply with heat exchange restriction [*Becker and Maréchal*, 2012] or a higher $\Delta T_{min}/2$ can be chosen.

This work tries to show and integrate the potential of storage and a stochastically available resource into the energy system. Considering the limited knowledge about potential additional restrictions such as already existing pipes, natural obstacles or property related restrictions, decisions in practice are often not rational [*Aringhieri and Malucelli*, 2003]. Therefore the network is not modeled.

In addition, a network model was tested using the different discrete temperature levels to connect demand and supply. It has the advantage of providing network temperatures and deciding between central and decentralized technology installation. The chosen formulation contains a big-M approach similar to the storage, creating numerical instabilities in the optimization. For small scale problems, such as around 20 demand streams in less than 50 periods results can be obtained. This module allows to further compare whether to work with one or two supply temperature pipes. Studies of [*Dalla Rosa et al.*, 2011, 2012] show the interest of studying a two supply temperature piping system.

Taking the hypothesis of a large $\Delta T_{min}/2 = 5K$ avoids using this model, because the temperature difference between hot and cold streams are big enough to integrate a network with its heat exchangers. Not modeling the network allows to keep the time steps of the storage model at a higher resolution.

5.9.2 Results of Storage Option Comparison

The impact on the system design of the three storage options are compared under the two objective functions: either a daily, a seasonal or both storage types can be chosen under one of the objective functions, minimizing annual costs or the annual CExD.

It is assumed that the total size of the daily storage is 1 $m^3/building$ totaling to 300 m^3, because this size can be installed without any constructional building modification. It is also the maximum steel tank size available, if bought as one single storage tank at the central heating plant. For the seasonal storage, the limit is set to 12 000 m^3, because one of the biggest realized projects in Friedrichshafen, Germany, is of this size.

Table 5.4 – Storage comparison: sizing for the case study with a maximal load of 13.3 MW

Objective Function	Storage types	Boiler Size [MW]	Solar Panels [10^3 m^2]	Daily Storage [m^3]	Seasonal Storage [m^3]
CExD	both	10.5	29.2	254	10 500
	daily	12.5	23.7	300	0
	seasonal	12.5	22.8	0	8 200
Cost	both	10.6	15	169	2 900
	daily	12.5	15.1	168	0
	seasonal	12.5	14.7	0	4 600

Table 5.4 shows the results for the comparison of storage options. The model chooses in all cases to use storage, because it improves the objective function's values by at least 7 to 15% compared to a solution with no storage. With different objective functions, either CExD or total annual costs, the results vary. The CExD objective lets the model choose more solar panels and a bigger storage size in every scenario. When both storages are active, the peak load can be significantly reduced, because the storages are also used during peak load. In addition, the solar thermal panel installation size reaches almost 30 000 m^2 when both storages are active for the CExD case.

All values are converted into CExD in Table 5.5 and annualized as for costs with the factor 0.0782, equivalent of an annuity for 25 years at 6%. This is done to compare results on equal basis. In the cases with CExD as minimization objective, the two storages are difficult to differentiate because the objective values are very close. While the big size of the seasonal storage has an impact on the construction value, the annualization used reduces this value. The operating CExD are the same for both cases. In all scenarios, the seasonal storage scenario is the most expensive one in terms of total annual CExD.

When the cost minimization are converted to CExD, they have a higher CExD for the cases with only one storage type. Especially the operating CExD is a lot higher. For the case with both storages, the operating value is similar to the one of the single storages in the CExD minimization. Due to the lower construction value, it is closer to the CExD minimization with both storages than to the individual storage cases.

Table 5.5 – Storage comparison: all costs are transformed to CExD values. Construction costs are annualized with the factor 0.0782.

Objective Function	Storage types	non-renewable Construction CExD $[GWh/a]$	Annual Construction Value $[GWh/a]$	Operating non-renewable CExD $[GWh/a]$	Total annual non-renewable CExD $[GWh/a]$
non-renewable CExD	both	38.0	3.0	14.6	17.6
	daily	28.6	2.2	17.2	19.4
	seasonal	29.5	2.3	17.2	19.5
Cost	both	19.1	1.5	16.9	18.3
	daily	18.4	1.4	18.7	20.2
	seasonal	19.0	1.5	18.9	20.4

The results converted to costs are in Table 5.6. As before, the investment costs are annualized and added to the annual operating costs to the total annual costs. The cost optimizations are significantly less expensive than the CExD ones. The relative difference between the solution with both storages is increasing, when they are converted to costs. For the daily storage, it is exactly the opposite: the cost solutions are closer to each other than the CExD solutions.

Table 5.6 – Storage comparison: all values in cost comparison, the investment costs are annualized with the factor 0.0782 to calculate total annual costs.

Objective Function	Storage types	Investment Cost [$MCHF/a$]	Annual Investment Costs [$MCHF/a$]	Operating Costs [$MCHF/a$]	Total annual Costs [$MCHF/a$]
non-renewable CExD	both	12.11	0.95	1.46	2.41
	daily	7.51	0.59	1.74	2.33
	seasonal	9.03	0.71	1.74	2.44
Cost	both	5.63	0.44	1.71	2.15
	daily	4.78	0.37	1.90	2.28
	seasonal	5.66	0.44	1.93	2.37

The overall CExD balance of Table 5.7 shows that the storage losses are higher for the seasonal storage. The daily storage has a daily cycle which is repeated for each typical day the number of times it appears in the given month compared to the seasonal one which is multiplied by the weighting factor of each period.

Dividing the production of solar panels Q_{solar} by the total heat demand Q_{Demand} is a standard indicator of the solar fraction. The indicator does not consider the temperature level at which the energy is provided.

$$fraction_{solar} = \frac{Q_{solar}}{Q_{Demand}} \tag{5.43}$$

For the given case study with high temperature buildings and large consumers, the highest solar fraction possible, defined as solar production over with the given choices, is almost 50% in the combined storage CExD case. A second group of cases is around 35% of solar fraction.

This shows that the storages are used in the same way for both cases: the storage is charged during the day when sun is available to replace if possible or delay the use of the boiler. The gas boiler ensures the rest and uses the storage during the peak load day so that it does not have to install a full power. Given the high temperature of the demand and the additional technologies available, even it is a gas boiler, the model never chooses to install a fully solar and storage powered system: only solar and storage based solutions a not economical neither from a pure cost based approach nor from a CExD based approach in this case study. Nevertheless, integrating storages in the existing system under the given conditions saves at least 15% of the CExD or 7% of the costs.

5.9.3 Discussion

When the thermal storage is introduced, the size of the peak load equipment is decreasing. This is an option that can be influenced through the maximum available power that the storage

Table 5.7 – Storage comparison: the total energy demand is 24.4 GWh

Objective Function	Storage types	Boiler Energy $[GWh/a]$	Solar Energy $[GWh/a]$	Solar Fraction [%]	Storage Losses $[MWh/a]$
non-renewable CExD	both	13.8	11.7	47.95	460
	daily	15.8	8.7	35.66	13
	seasonal	15.8	9.0	36.89	110
Cost	both	15.5	8.9	36.48	277
	daily	17.5	7.0	28.69	27
	seasonal	17.3	7.1	29.10	147

can use for charging or discharging. With the here chosen approach, the peak load period is perfectly known in advance: the model uses a perfect horizon. The model charges the storage with the peak load utility in the hours before the peak. During the peak hours, the peak load utility and storage discharging ensure the heat requirement. From the functional side, this behavior is correct. When installing such a system, this requires a predictive control otherwise the available power is not sufficient. This bears a certain risk of not being able to fulfill the heating requirements in all future situations and gives rise to include uncertainty into the system design.

The solar fraction of 50% is relatively high for the case with both storages under a CExD minimization when taking the relatively high heat losses into consideration. They are the same number of magnitude than in Friedrichshafen, where the lower third of the tank is not insulated and get compensated through the installation of more solar thermal collectors.

For the CExD approach, obtaining a coherent data set of CExD values remains the biggest part of the work and the key challenge: the data sets available such as [*Dones et al.*, 2007] are gathered under a life cycle perspective and not for energy system design. The data sets itself seems to be studied with great accuracy. However, the case studies used for defining the CExD of a given technology do not always use the technology in the same way as in this work, neither does a choice of different equipment sizes exist.

In addition, often the documentation already leaves a key question unanswered: which exact technology is studied from which year and how did updated values get calculated? When using the data, ideally the production site of technology should be known to estimate the transportation impact as well. If these questions can not be answered precisely, the proposed CExD values should be considered as an order of magnitude value rather than exact number. In the case where the solutions are close to each other, the answers to these questions can help to distinguish between them. CExD has the advantage of being a useful supplement to energy integration considering the whole life cycle as shown for the refurbishment example.

Mass production of a given technology significantly reduces its price. In a cost based optimization, these technologies will be chosen primarily. The CExD suffers less from this effect even though a mass production might also be more effective in terms of CExD as it might be able to reduce the CExD as well. Comparing the solutions of CExD minimization to the cost ones generally shows towards more infrastructure expenses than operating ones.

With this problem structure, additional indicators such as greenhouse gas emissions, environmental impact points or primary (non-) renewable energy use can be integrated quickly, which can help gaining a larger view on the solution. Compared to the CExD, the mentioned indicators are in common use and more values can be more easily obtained.

5.9.4 Comparison of seasonal Storage Formulations

Figure 5.5 compares the different execution times of MILP and LP formulation of the seasonal storage. For the LP model, the top and bottom losses are ignored. The MILP model uses the top and bottom losses as described in 5.3.3. The difference in term of storage losses is around 10 % . The difference in terms of resolution however needs to be plotted on a logarithmic scale. Especially when the time slices are increased, the MILP model depends more and more on heuristics within the solver to find a solution quickly. Depending on how fast the heuristic finds a feasible solution, the solution time varies between on and the same run. Measuring the resolution time when the input data is changed, even if the change is only minor, also offers the whole range of resolution time variability with at least the same order of magnitude shown in the Figure 5.5.

In order to increase the solution time, the MILP gap can be lowered. The solution time decreases, however the system configurations that are found with a low gap value of 0.01 % are sometimes not found anymore. Instead a similar solution in terms of objective value is found, that differs substantially in terms of system sizes. Reducing the gap leads therefore to inconsistent solutions and is therefore not studied any further.

5.10 Conclusion

The developed LP and MILP models allow to propose energy system design while integrating storage with a stochastic renewables. Compared to the previous works of [*Fazlollahi et al.*, 2012] using a MILP daily storage model with independent typical days in the slave optimization, in this work it is possible to link the days to each other. A seasonal energy storage is possible at the cost of creating a bigger model. Most projects with seasonal storage have a buffer storage and a seasonal long term storage. The optimal solution proposed through the model also propose a two storage type solution under two different objective functions.

Further, the comparison of the solutions obtained through two different objective functions show that under a CExD objective more investment into storage and solar thermal panels is

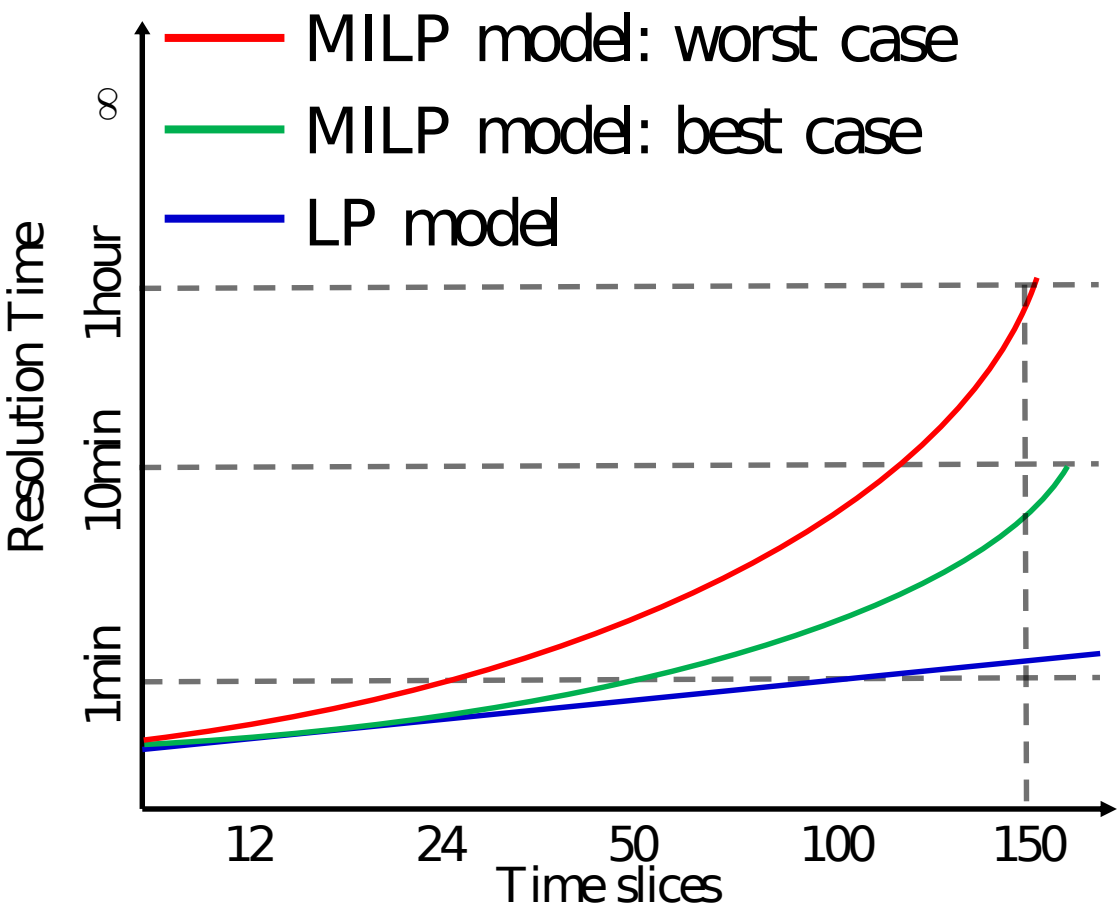


Figure 5.5 – Computation time comparison of LP and MILP models for the seasonal storage model

optimal. If a lower total CExD value is defined as a more sustainable solution, this indicator points clearly towards more storage usage during this case study. However, the CExD values have to be considered as an order of magnitude number rather than exact number because it is very difficult to ensure that for all systems the same hypothesis are considered.

Compared to classical approach with cost functions, renewable penetration is favored, because it is available as a quasi-free source within the system after the installation costs or CExD. The CExD indicator is therefore a suitable combination for energy integration to gain a more holistic overview of the system. Of course, the exergy indicator has known drawn backs such as ignoring toxicity but neither does the cost indicator. Combining both approaches in an MILP identifies the best solution under each objective, costs and CExD, directly with two separate runs.

About 40 years after the first formulation of a multi-period problem for energy system design, writing the problem remains fairly easy compared to solving it, when integrating storage. Reducing the number of time slice dependent integer variables or reducing the time slice

number remains a key aspect for obtaining a solution. The problem is easy to formulate, but hard to solve.

5.11 Future Work

Introducing two detailed temporal scales with storage into an MILP model leads to large problems. For that reason, the problem needs either to be simplified leading to approximations. Or a decomposition approach as presented in [*You et al.*, 2011] should be used as it seems to be the most promising one.

6 Buildings as thermal Energy Storage

In this chapter, CitySim, a dynamic building simulation software is used to estimate EPFL's heating demand. For the chosen buildings of EFPL, the building's active heat storage capacity is estimated respecting the comfort temperature range. This capacity is integrated into the energy system to show that the proposed models of Chapters 4 and 5 can also be used on much more detailed level. The results show the potential savings, when using the building as a thermal heat storage.

6.1 Introduction

The energy demand of EPFL is taken from *Coccolo et al.* [2015]. The model of the EPFL campus is realized with the software CitySim, calculating the heating and cooling demand of the campus and the electricity produced by the PV on the rooftop of buildings. The model was validated with on-site monitoring, showing a good correlation factor between the measurements and the model.

The optimization considers the impact of using the heat capacity of building's envelope within the indoor comfort temperature range; the best optimal operational mode for the building's capacity and the solar panels is identified, showing the variation of the internal temperature using the building mass as energy storage. This approach allows shifting peak loads by several hours, reducing the heating power demand.

This is called demand side management (DSM). The term DSM was introduced in the 1980's by the Electric Power Research Institute[*Arteconi et al.*, 2012, *Warren*, 2014] and refers to the ability of changing electric energy demand patters behind the meter that are mostly out of (direct) control of the network operator, historically also known as load management. In most cases, this is done through education and (financial) incentives. The importance of DSM is in the increase of the network utilization rate: more consumers can be connected to the current network without further investment, increasing the network's cost-effectiveness. [*Warren*, 2014] defines it as reverse thinking: DSM tries to match demand with the available supply.

Three categories of DSM exist: energy efficiency measures such as changing appliances, on-site backup with additional generation and storage capacity and demand-side response using incentive based approaches. Further, [*Warren*, 2014] points out that no unique definition exists on whether to go further than only electricity. With the introduction of so-called smart grids the definition is extended to all services (heating, cooling, electricity). only the use of the building mass as a thermal energy storage is considered for the purpose of demand side management.

The available operating temperature range, which is equivalent to the internal temperature available to the DSM, is limited: It needs to be adapted to the comfort levels of people living or working the building (or to the industrial requirements).

In the area of building physics, accessing the thermal mass or capacity impacting the heating or cooling load is a widely discussed field. Experimental [*Childs et al.*, 1983] and analytic [*Ma and Wang*, 2012] studies demonstrate that most estimations of the buildings thermal capacity are wrong due to the fact that heat convection is neglected or not correctly calculated.

[*Reynders et al.*, 2013] uses a detailed approach for the modeling of the different heating distribution systems. Then they evaluate different options concluding that DSM shows strong peak shaving potential of up 94% for the heat pump's electricity consumption. [*Xue et al.*, 2014] proposes a management system enabling the use of DSM. Through the use of HVAC utilities the peak shaving and load shifting in the electric grid is demonstrated.

According to [*Xue et al.*, 2014] an interactive building power demand management strategy for the energy management of a building point is key driver for demand alternation potential. On the one side [*Reynders et al.*, 2013] use a detailed approach for the modeling of the different heating distribution system than [*Xue et al.*, 2014] to show the impact of different heating systems and building structures. On the other side, [*Xue et al.*, 2014] use a more detailed approach to model the thermal mass of the building. The lumped method used by [*Reynders et al.*, 2013] has to be treated with caution as they did not show how they treat the heat convection within the building's structure. Both show promising results with varying reduction potentials up 94% of the original demand. These results vary as they depend on the building type, the state of the energy conversion system together with the application case.

The Danish Heat Atlas [*Möller*, 2012] identifies DSM as a key technology for networks in the context of heat supply. But it does not provide any further ideas on how to reach this potential.

6.2 Method

For the purpose of using a building as an energy storage, first the thermal mass is calculated with a simplified model considering heat convection based on [*Ma and Wang*, 2012] findings. This result is implemented in an existing MILP model minimizing overall costs or costs while fulfilling the heating requirements. Combining simulation to estimate the heat capacity and

optimization its use allows to decide whether using the building as an energy storage is of economic interest. Compared to an only simulation based approach requiring frequent runs to find the optimum, the result can be presented after directly .

The gains through the sun, persons in the building or appliances are integrated into this approach by the building simulation. When the gains increase, the heating demand decreases. This means that they are not taken twice into consideration.

6.2.1 Effective thermal Capacity of a given Wall

Depending on the length of the storage period, different thermal masses should be considered. Figure 6.1 represents a simplified thermal mass model estimating the heat capacity of the wall. It is based on the hypothesis of one-dimensional heat conduction without internal heat generation and a constant thermal conductivity. In our case, the internal air temperature acts as an excitation source following a sinusoidal function and has a period of 24 hours as in [*Ma and Wang*, 2012].

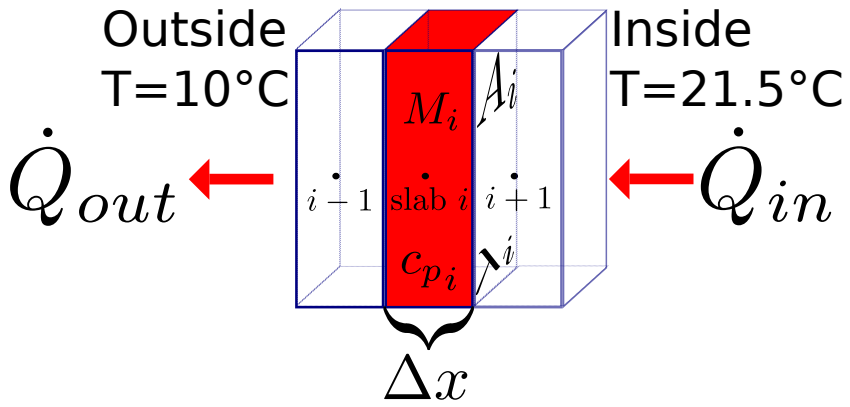


Figure 6.1 – Simplified illustration of the one dimensional heat flow $\dot{Q}$ across i elements with a width of Δx and a cross section surface A_i

No heat energy is created in the wall, nor do mass transfers happen. Equation 6.1 applies an energy balance to a given wall section i with a cross section A. The change of internal energy U over time t equals the difference of the incoming heat load $\dot{Q}_{in}$ and the outgoing heat load $\dot{Q}_{out}$ as the mass M and the specific heat capacity c_p as well as the thermal conductivity λ remain constant. This implies that the temperature change ∂T over time ∂t is proportional to the temperature change over the distance ∂x.

$$\frac{dU}{dt} = \dot{Q}_{in} - \dot{Q}_{out} \implies Mc_p \cdot \frac{\partial T}{\partial t} = A\lambda \frac{\partial T}{\partial x} \tag{6.1}$$

$$\frac{\partial T}{\partial t} - \frac{A\lambda}{Mc_p}\frac{\partial T}{\partial x} = 0 \text{ with: } \frac{\partial T}{\partial t} = \frac{T_i^t - T_i^{t-1}}{\Delta t} \text{ and } \frac{\partial T}{\partial x} = \frac{T_i^t - T_{i-1}^t}{2\Delta x} + \frac{T_i^t - T_{i+1}^t}{2\Delta x} \tag{6.2}$$

Equation (6.2) shows the discrete time derivative of $\frac{\partial T}{\partial t}$ using a first order implicit scheme and the spatial derivative of $\frac{\partial T}{\partial x}$ with a centered finite difference scheme of second order along the direction of x. Δt represents the discrete time t and Δx the discrete width of a slab. Mc_p depends on the mass M and the specific heat coefficient c_p of the considered wall section. Equation (6.2) can be transformed to Equation (6.3) that calculates the temperature T of the i-th wall section in the time $t-1$ as function of the surrounding wall sections' temperatures ($i+1$ and $i-1$) and the temperature of the wall section i in the next time step t. λ represents the heat conductivity of the material in wall section i.

$$T_i^{t-1} = T_i^t - \frac{A_i \lambda_i}{M_i c_{p_i}} \frac{\Delta t}{\Delta x} \left(T_{i-1}^t - 2 \cdot T_i^t + T_{i+1}^t \right) \tag{6.3}$$

For the boundary conditions, the first element of the wall in contact with the air in the room is calculated as follows:

$$U_i = \frac{1}{\frac{1}{h_{int}} + \frac{x}{2\lambda_i}} \tag{6.4}$$

$$\frac{T_i^n - T_i^{t-1}}{\Delta t} = \frac{U_i A_i}{M_i c_{p_i}} \frac{(T_{int}^t - T_i^t)}{\Delta x} - \frac{(T_i^t - T_{i+1}^t)}{\Delta x} \frac{A_i \lambda_i}{M_i c_{p_i}} \tag{6.5}$$

$$T_i^{t-1} = T_i^n - \underbrace{\frac{\Delta t}{\Delta x} \frac{1}{M_i c_{p_i}} \left[U_i A_i (T_{int}^t - T_i^t) - A_i \lambda_i (T_i^t - T_{i+1}^t) \right]}_{\text{excitation vector } \vec{c}} . \tag{6.6}$$

The indoor convection h_{int} is fixed at $10 \frac{W}{kg\,K}$. Replacing the indoor temperature T_{int} and the convection h_{int} with the external ones, allows to calculate the wall section temperature on the external side. In order to obtain the thermal capacity of the wall a sinusoidal temperature variation is applied:

$$T_{int}(t) = 21.5 + 1.5 \cdot sin(2\pi \cdot \frac{t}{24h}) . \tag{6.7}$$

The indoor temperature varies within a defined comfort band of 3° Celsius over the duration of a day based on [*Swiss Society of Engineers and Architects*, 2006].

All wall section temperatures over the whole 24 hours can be calculated at once using this implicit linear Equation:

$$B\vec{x}_{(t)} + \vec{c} = \vec{x}_{(t-1)} \leftrightarrow \vec{x}_{(t)} = B^{-1} B \vec{x}_{(t)} = B^{-1} \vec{x}_{(t-1)} - B^{-1} \vec{c} . \tag{6.8}$$

With Equation 6.8 based on the temperature vector of the previous time step $\vec{x}_{(t-1)}$, the temperature of each wall section at the current time step $\vec{x}_{(t)}$ is expressed with the excitation vector $\vec{c}$ and the coefficient matrix B. This approach is unconditionally stable and thanks to the matrix with constant coefficients only a single matrix inversion is needed leading to computationally inexpensive estimation of the wall's behavior. Figure 6.2 shows the wall's

temperature after going through a cycle raising the internal temperature from 20°C to 23°C and going back to 20°C.

The storage capacity can then be estimated as follows: Over the discharging cycle, when the internal temperature T_{int} is falling, the effective wall thickness x_{eff} used for heat storage can be estimated as a function of the first wall slab $i = 1$ in contact with the internal air:

$$x_{eff} = \frac{\sum_{T_{max}}^{T_{min}} \dot{Q}_1 \, \Delta t}{A_1 \, \rho_1 \, c_{p_1} \, (T_{int}^{t=T_{max}} - T_{int}^{t=T_{min}})} \, . \tag{6.9}$$

For concrete, this approach yield 10.5 cm, which is the same value as reported by [*Ma and Wang*, 2012] with 10.5 cm. The total effective heat capacity for a building is calculated using the total surface A_{tot} in contact with the heated air within the building:

$$C_{eff} = x_{eff} \, A_{tot} \, \rho_1 \, c_{p_1} \, . \tag{6.10}$$

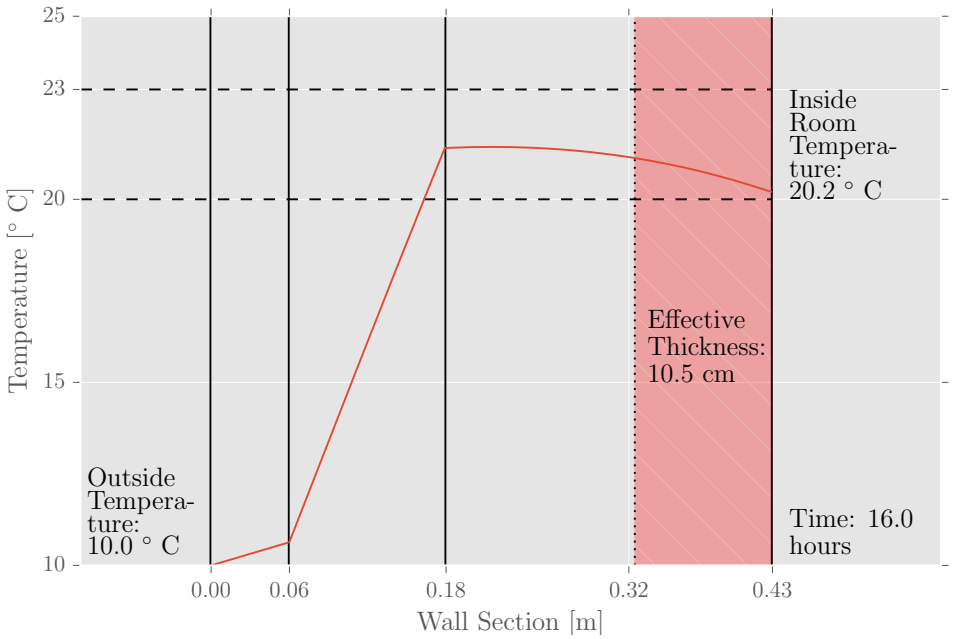


Figure 6.2 – Temperature evolution within a type 2a wall of EPFL while in cooling down phase of sinus with a duration of 24 hours going from 23° to 20°C. The layers from left to right are: brick ($0-0.06m$), insulation ($0.06-0.18m$) and concrete ($0.18-0.43m$)

6.2.2 Linear Optimization Model

The objective function of the MILP model minimizes the overall cumulative exergy demand, the same formulation can be used to minimize the overall costs. The model uses the heat cascade to size the equipment. For this work, two equipment types are available: a gas boiler and a solar thermal collector. The solar thermal collector has the choice to deliver heat at different discrete temperature levels according to the available irradiation calculated by CitySim. Higher temperatures yield lower panel efficiencies. The boiler delivers heat at constant high temperature.

The DSM model can be activated by two mechanisms: either different tariffs for resources are introduced such as day and night tariffs or a stochastic free resource is available during the certain hours of the day. The second option is chosen: the solar thermal panels operate for free during the hours where sufficient irradiation is available. They can be combined with a seasonal storage. The detailed model for the seasonal storage and the solar thermal collector can be found in 5.

In simulation approaches, such as used in *Coccolo et al.* [2015], the internal temperature is fixed to a specific value. In this work, the energy demand is introduced as a bounded variable to show the potential of energy storage in a building's wall. The bounds are the comfort temperature range. More heat is used at a certain moment than the current heating requirement to store it in the building's wall.

All **variables** within the MILP are printed in **bold**. Equation (6.11) guarantees the thermal comfort of the building: the room temperature $\Delta\mathbf{T_b} + T_{min}$ should not be higher than the maximal room temperature T_{max} defined as being within the thermal comfort range by the occupants neither lower than the minimum temperature T_{min}. The temperature bandwidth variable $\Delta\mathbf{T_b}$ can be set for each time step, according to occupation patterns, day and night shifts or seasonal preferences for each building.

$$T_{min} \leq \Delta\,\mathbf{T_b} + T_{ref} \leq T_{max} \tag{6.11}$$

$$\Delta\mathbf{Q_b} = \Delta\mathbf{T_b}\; C_{eff} \tag{6.12}$$

The heat stored in the building's structure ΔQ is calculated via this temperature difference ΔT and the effective heat capacity of the building C_{eff} with Equation (6.12). The multiplication of the overall heat transfer coefficient U_{tot}, the building's surface A_{tot} and the temperature difference provide the additional heat losses due to the temperature raise in Equation (6.13).

$$\Delta\mathbf{Q_b} = \Delta\mathbf{Q_{b,0}} + \sum_{t=1}^{t=nt} d_t \cdot \left(\dot{\mathbf{Q}}_{\mathbf{heating:t}} - U_{tot}\; A_{tot}\; \Delta\mathbf{T_b}\right) \tag{6.13}$$

The sum is only over the time slices and not the periods because the heat can not be stored long enough in the building's structures. (The subscript p is therefore omitted.)

$$\dot{\mathbf{Q}}_{\mathbf{heating:t}} = \left(\mathbf{f}_{\mathbf{heating:t}} - 1\right) \cdot \dot{Q}_t \tag{6.14}$$

They are linked to the reference heat load $Q_{heating:t}$ via the multiplication factor $\mathbf{f_t}$ in Equation (6.14): during charging, f is bigger than 1, during discharging it will be smaller than 1. A value of 1 for $\boldsymbol{f}$ delivers only the current heating demand. The heat cascade links the energy conversion technologies with the varying heat demand.

When heat should be stored in the building, the heat distribution temperature level T_l needs to choose a stream with a higher temperature. In total, three levels are proposed, two with a higher temperature. The level 0 is the lowest level that can only satisfy the demand without lifting the temperature of the building's wall. All levels have predefined temperature levels but can be activated continuously through their multiplication factor $\boldsymbol{f_t}$. A combination out of all three streams can be used to first supply the required heat and then heat the buildings' walls leading to a continuously floating temperature of the wall. Equation (6.15) ensures that the building is always heated with the lowest temperature level 0 plus the current building temperature.

$$\forall\, t \text{ and temperature levels } l \in 0,1,2: \quad \sum_{l=1}^{nl} \mathbf{f_{l,t}} \cdot T_{l,t} \geq \sum_{l=1}^{nl} \mathbf{f_{l,t}} \cdot T_{min,t} + \Delta \mathbf{T_b} \tag{6.15}$$

If only the heating requirement needs to be fulfilled, the current level is sufficient. When the building should be heated to a higher internal temperature, also a higher heating level needs to be used. A higher heating requirement leads to a higher demand of utilities that a certain (CExD) price. Therefore the heating requirement will only be increased, if the overall costs can be increased.

In order to run the optimization, the hourly input data, the heat load and the solar irradiation calculated by CitySim, are clustered/reduced to 10 representative days with 13 time steps per day. The method explained in Chapter 4 is applied to the data set. Compared to the data of Verbier, more days are needed to represent the data, because EPFL has a different usage pattern with university holidays and weekend versus week day load. The reduction of input data makes running optimization model possible while respecting the power and energy balance. Compared to 8760 hours only 113 time steps need to be taken into consideration. Two buildings were analyzed, as they represent a light mass building (Polydome), and a massive one (AAB). Their physical envelope's characteristics are summarized in Table 6.1. The values of the effective depth are a function of the first internal layer. The heat loss coefficient represents the total weighted average U-value of the building.

Table 6.1 – Effective thermal capacity of studying buildings: Polydom, the lightest building on campus based on a wood structure and AAB, the massive building with a concrete structure.

Building	Total Thermal Capacity $[\frac{MJ}{K}]$	Effective Thermal Capacity $[\frac{MJ}{K}]$	Capacity Percentage [%]	Effective Depth $[m]$	Material [-]	Total Building Surface $[m^2]$	Total U-Value $[\frac{W}{m^2 \cdot K}]$
Polydom	1583	490	25.6	0.12	Wood	3249	0.22
AAB	2450	627	31.0	0.09	Concrete	2564	0.38

6.3 Results

The Table 6.2 shows 7 calculations:

1. reference case with gas boiler and 201 m^2 of solar thermal panels,
2. building heat storage is available with the same solar panel surface,
3. building heat storage is available with the optimal solar thermal panel size chosen by the model,
4. a thermal heat storage of 100 m^3 is added to the reference case with the gas boiler and 201 m^2 of solar thermal panels,
5. the building heat storage, 100 m^3 of thermal heat storage and 201 m^2 of solar thermal panels are available,
6. 100 m^3 of thermal heat storage and optimal solar thermal panel size chosen by the model,
7. building heat storage and 100 m^3 of thermal heat storage are available and optimal solar thermal panel size is chosen by the model.

When solar thermal energy is available during the day, generally a slight load shift can be achieved starting the boiler later. The boiler also uses the building heat storage because its peak power decreases slightly by about 3% or 4.5 kW.

Figure 6.3 shows the temperature variation of the indoor temperature over time for scenario 3. In between the time slices 1 and 55, the building heat capacity is used when sun is available during the day and a heat load at night is required. During the summer, represented by time slice 55 to 90, no heat demand is present therefore this storage is not activated. Table 6.2 summarizes the different scenarios: using the building's heat capacity plays an important role on the overall cumulative exergy consumption and reduces the heating power for all scenarios. With the same system configuration, but using the building's heat capacity, the objective value improves by at least 6 %. Introducing a seasonal energy storage of 100 m^3 is only of significant advantage when the building's heat capacity is used as well. Without it, there is only a minimal improvement.

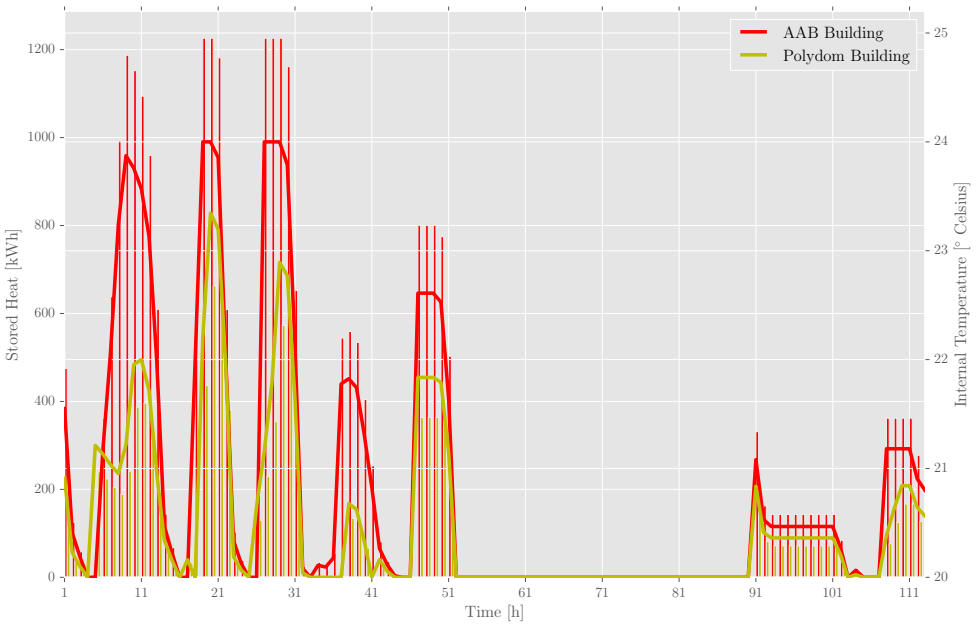


Figure 6.3 – The red line represents the AAB building, the yellow one the Polydom. Their internal temperature changes over time are shown through the colored lines and the stored heat per building in the bar plots for scenario 3

6.4 Discussion

The results show the interest in using the building as thermal energy storage. The simulation model can be used to compare different wall compositions. For the estimating the capacity, the knowledge of the wall's composition is important. Otherwise big errors can be made, for example if the building is isolated in the inside due to the fact that it is a historical building.

On one side, the information can not be accessed easily, because they are not found within the available data bases nor do all the owner know how they walls are composed. During energetic refurbishment, one might have the chance to get a detailed picture of the wall's composition. On the other side, with this approach different compositions can be tested and their capacity can be compared so that the most realistic (or pessimistic) one can be chosen.

The shown behavior of the model is based on perfect knowledge of the future heat demand (and therefore the occupation and the weather forecast). The boiler also uses the buildings mass as a thermal storage, decreasing its size. Comparing the reference case and the cases with no building heat storage to the ones where the building storage is additionally included, shows savings up to 22 % .

The optimization results depend strongly on the configuration of the energy system: the storage with twice as much solar thermal panels can save as much as the active building heat

Table 6.2 – Results of all scenarios: The storage is always at indicated size, optimal solar panel size indicates that the model chose the solar panel size to be installed.

#	Scenario Description	Building Heat Storage	Improvement to 1 [%]	Boiler Size [kW]	Solar Panels [m^2]	Long term Storage [m^3]
1	Reference Case	No	-	127.8	201	0
2	Solar as in 1	Yes	6.28%	123.3	201	0
3	Optimal Solar Panel Size	Yes	16.90%	123.3	831	0
4	Solar as in 1	No	4.52%	127.8	201	100
5	Solar as in 1	Yes	8.58%	123.3	201	100
6	Optimal Solar Panel Size	No	6.64%	127.8	449	100
7	Optimal Solar Panel Size	Yes	22.01%	123.3	1098	100

storage. The highest savings result from a combination with a long term storage that is also heated up during the time where no heat demand is present realizing the about 6% of CExD economy shown in scenario 7. The results of scenario 3 and 6 add up to it.

6.5 Conclusion

The two-step approach allows to first estimate the effective thermal capacity of a building and second optimize the overall energy system with this additional storage. Instead of searching for an optimum internal building temperature through running simulations, the optimization model uses the effective heat capacity to optimize the indoor temperature minimizing non-renewable energy use. Through the optimization, the presented approach indicates operating strategy for the energy system. A linear optimization model will not give the same result as detailed building model, but it helps to identify which operating conditions should be studying in more detail with a low computational effort.

Using the building structure as an additional heat storage within a comfort temperature bandwidth shifts peak loads, reducing slightly the equipment size. Realizing the load shift can be done using command predictive control. A final decision of which equipment to deploy at which size requires a detailed study of the energy conversion system already available on site.

6.6 Future Work

A rolling horizon could be used instead of the perfect horizon to consider the uncertainty in weather predictions and people's presence. Comparing the two solutions, one with the design model and the other one with the rolling horizon, would allow to get an idea on how well in advance the weather and the demand need to be known to activate this potential.

Additionally, a Photovoltaic module with a heat pump should be added. The heat pump could use the electricity of the PV module to either supply heat to satisfy the demand, store it in

a thermal energy storage or in the building. This allows to compare two competing system configurations and would allow to decide between solar thermal or PV installation for a given roof top.

7 Retrofitting as Decision Variable

The building stock in Switzerland is responsible for the biggest part in the final energy consumption. Only using more renewable energy sources on the supply side is a step towards reaching the actual political goals, however it will not be sufficient. The reducing the demand is key challenge to a more sustainable overall energy system.

Therefore a trade-off exists between either investing into a more renewable energy supply or reducing the energy requirements. The presented approach based on the CExD can help finding a solution, because it considers a holistic view of the energy system including the supply chain.

Jennings et al. [2014] present a similar concept using a MILP framework for retrofitting residential energy systems at urban scale. They use a spatial scale of 19 districts and an annual time resolution to propose measures for the future design. In their case studies the centralized solutions can be very interesting.

The here presented approach extends their framework by considering a more detailed energy storage model and demand profiles in an hourly time resolution based on the typical days.

7.1 Model Extension for Demand Side Technology: energetic Building Renovation

When considering different supply options for the integration into the energy conversion system, it is important to consider at the same time the options that allow buildings to be refurbished.

In the model, we consider that the buildings can be in either a refurbished state or left at original state. As in Chapter 5, Verbier is taken as a demonstration test case and the heat load is modeled with a dynamic building simulation ([*Page et al.*, 2014]). The heat load for both cases is estimated with and without insulation.

When the building is refurbished, the heat load is reduced as well as its distribution temperature: the new distribution temperatures are estimated using the model proposed in [*Girardin et al.*, 2010]. It assumes a constant mass flow in the heat distribution system. The nominal mass flow $\dot{m}_0$ multiplied by the specific heat capacity c_p equals the nominal heating power $\dot{Q}_0$ divided by the nominal temperature difference between supply and return temperatures $T_{supply,0}$ and $T_{return,0}$. The supply temperature $T_{supply,p,t}$ in the period p and time slice t is then calculated as a function of the internal temperature T_{int} and the ratio of the current heat load $\dot{Q}_{p,t}$ over the nominal one, multiplied with the nominal temperature difference $T_{supply,0} - T_{return,0}$ and a factor representing the heat exchange. Subtracting the ratio of the current heat load $\dot{Q}_{p,t}$ and the product of nominal mass flow $\dot{m}_0$ and the specific heat capacity c_p, the supply temperature $T_{supply,p,t}$ is obtained.

$$\dot{m}_0 c_p = \frac{\dot{Q}_0}{T_{supply,0} - T_{return,0}} \tag{7.1}$$

$$T_{supply,p,t} = T_{int} - \frac{\dot{Q}_{p,t}}{\dot{Q}_0} \cdot (T_{supply,0} - T_{return,0}) \cdot \frac{\alpha}{1-\alpha} \tag{7.2}$$

$$\text{with } \alpha = \frac{T_{supply,0} - T_{int}}{T_{return,0} - T_{int}} = exp(\frac{U_0 A_0}{\dot{m}_0 c_p})$$

$$T_{return,p,t} = T_{supply,p,t} - \frac{\dot{Q}_{p,t}}{\dot{m}_0 c_p} \tag{7.3}$$

The objective function is modified to take into account the additional CExD needed for refurbishment. The new terms are:
Ref_b represents the CExD for refurbishment of the building b
$y_{refurbished}(b)$ is the binary variable activating refurbishment of building b

$$F_{obj,cexd} = min\Bigg(f_A \cdot \Bigg(\sum_{u=1}^{nu} invCExD_u \cdot \boldsymbol{s_u} + \sum_b Ref_b * y_{refurbished}(b)\Bigg) + \sum_{p=1}^{np} w_p \cdot \sum_{t=1}^{nt} \sum_{u=1}^{nu} \Big(cexd_{f,u} \cdot \boldsymbol{f_{u,p,t}} \cdot \dot{Q}_u^+ + c_{el,p,t}^+ \cdot \dot{\boldsymbol{E}}_{el,p,t}^+ - c_{el,p,t}^- \cdot \dot{\boldsymbol{E}}_{el,p,t}^-\Big) \cdot d_{p,t}\Bigg) \tag{7.4}$$

Now the objective function includes the demand side: when it is optimal to refurbish a building, the additional CExD is added to the objective function.

For the integration of the demand side technology energetic refurbishment, the model is extended to ensure that it either contains the current state of the building $y_{current}$ or the retrofitted state $y_{refurbished}$. The model modification introduces two indicator variables that

activate either the current heat demand in the heat cascade or the refurbished one. Equation (7.5) ensures that either the current heat demand or the refurbished one is used.

$$y_{current} + y_{refurbished} = 1 \qquad \forall b\,. \tag{7.5}$$

The CExD to retrofit a building is not available. Therefore the values shown in Table 7.1 are used to represent specific retrofitting CExD values. Triple glazed windows and 20 cm of

Table 7.1 – CExD values used for energetic refurbishment

Type	CExD
thermal insulation	106 $kWh-Eq/kg$
triple glazed windows	460 $kWh-Eq/m^2$

thermal insulation are added to each building. The corresponding specific CExD values can be found in Table 7.1. The catalog of building parts [*Marti*, 2002] and [*KBOB*, 2014] regroup more Life Cycle Assessment data for additional materials related information.

7.2 Case Study

24 houses from the community of Verbier are selected. The dynamic building simulation estimates the energy demand after the retrofit and the current one. With the building's data available the material CExD use is estimated for the energetic refurbishment. For each building, the CExD is estimated through multiplication of the building's wall and window surface with the specific CExD shown in Table 7.1. The result for each building can be found in Table 7.2: the part of the isolation is much higher than the one for the windows. The 24 buildings are all of different size with different heat loads.

The results of the two building simulations, the current state and the refurbished state, are shown for 24 buildings in Table 7.3. For most buildings, the energy consumption is reduced by a factor of around 4 and the power divided by a value of around 2. The new maximal distribution temperature of each building is also reduced by up to 26° C. Building 1 for example is exactly in this range. In contrast, building 11 is already a low energy building, the additional measures have nearly no impact.

With the new input data, a holistic recommendation can be given regarding the best actions on the demand and supply side. Two scenarios are calculated, both have a maximal solar panel surface of 30 000 m^2 and include both storages types with the limitation of 300 m^3 for the daily and 12 000 m^3 for seasonal storage available. In scenario 1, only an additional gas boiler is available. In scenario 2, a heat pump and also the gas boiler are available. The heat pump uses the Swiss electricity mix. Therefore the heat production costs are $2.00/COP$ $[kWh-Eq./kWh]$.

Table 7.2 – Overview of the energetic refurbishment CExD values for the 24 buildings

#	ID	Isolation CExD $[MWh-Eq.]$	Windows CExD $[MWh-Eq.]$	Total CExD $[MWh]$
1	356	83.6	4.9	88.6
2	559	820.1	73.0	893.1
3	602	86.6	3.6	90.2
4	1104	120.4	6.0	126.4
5	1341	431.1	30.0	461.1
6	1525	3622.8	381.7	4004.6
7	1710	327.6	24.6	352.1
8	1722	114.4	5.8	120.1
9	1819	147.8	8.1	156.0
10	1901	213.5	10.9	224.3
11	2079	230.4	16.0	246.4
12	2082	165.1	9.2	174.3
13	2084	387.0	19.9	406.9
14	2454	108.1	7.2	115.3
15	2458	437.3	29.9	467.1
16	3424	692.9	43.1	735.9
17	3429	251.9	16.7	268.6
18	4323	290.9	17.0	307.9
19	5030	330.6	16.8	347.4
20	5357	231.2	10.6	241.9
21	5375	446.8	24.7	471.5
22	5769	110.2	5.2	115.4
23	6133	140.6	10.5	151.1
24	6867	632.5	42.9	675.4

7.3 Results

For scenario 1, 19 out of 24 building are refurbished. In scenario 2, 11 out of 24 are chosen. The Table 7.4 shows the results. From the 11 buildings, 9 are refurbished in both scenarios. The remaining two, number 3 and 14 are small buildings that are only refurbished in the retrofitted scenario. They both have a high distribution temperature.

The buildings 1, 2, 5, 6, 7, 15-17, 20, 22, 24 are only refurbished in scenario 1. Almost all buildings that have a maximal distribution temperature at 70 °C are retrofitted. The 14 buildings with the highest energy consumption out of the 24 are chosen to be retrofitted in this scenario. For the heat pump, no strategy clear strategy is found: The refurbished buildings are rather placed in the middle in terms of annual energy consumption.

Building 1 is refurbished, when the energy supply mix is a combination of solar and gas, compared to scenario two with the heat pump where it is not retrofitted. The operational CExD

Table 7.3 – The energy statistics for the 24 buildings

#	ID	Energy Demand $[MWh]$	Ref. Energy Demand $[MWh]$	Power $[kW]$	Ref. Power $[kW]$	T^{max}_{supply}	$T^{max}_{ref,supply}$
1	356	21.5	5.4	5.8	3.3	70	45.9
2	559	362.4	93.6	149.7	78.2	70	44.4
3	602	37.0	9.5	6.3	3.6	70	45.9
4	1104	33.7	8.2	10.6	6.0	70	45.8
5	1341	49.7	32.1	60.5	40.3	54.6	40.6
6	1525	1744.0	1147.2	673.1	384.6	60.1	40.4
7	1710	91.4	23.0	43.3	24.8	70	45.8
8	1722	17.2	7.8	10.2	5.8	70	45.8
9	1819	41.7	10.1	14.3	8.2	70	45.9
10	1901	100.6	24.1	25.6	13.9	70	44.6
11	2079	29.4	29.4	17.1	17.1	40	38.3
12	2082	76.1	17.9	16.2	9.3	70	45.9
13	2084	115.2	29.2	58.6	33.5	70	45.8
14	2454	44.4	10.9	8.5	4.8	70	47.9
15	2458	69.5	32.5	70.2	40.1	59.7	40.6
16	3424	231.5	57.5	126.5	72.3	69.8	46.3
17	3429	62.2	27.2	26.2	16.8	55	41.1
18	4323	84.7	21.2	40.0	22.9	70	45.8
19	5030	156.8	37.8	49.4	25.4	70	43.3
20	5357	55.7	25.8	22.7	14.3	70	49.5
21	5375	132.5	34.0	72.7	41.6	69.8	45.8
22	5769	11.7	7.5	7.8	5.2	59.8	43.7
23	6133	24.4	14.6	10.6	7.1	55	40.8
24	6867	184.3	49.9	126.1	72.0	69.8	45.7
	$\sum$	3777.6	1756.4	1652	951.1		

for building 1 is too small to be refurbished for scenario 2 with respect to the refurbishment CExD expenses required. Building 11 is never refurbished, because it is already a new building with a low heat demand. The expenses for the refurbishment would be higher than the economies over the estimated lifetime. Building 4's heat losses are important enough to be a candidate for refurbishment in any supply case.

Both scenarios use storage for peak shaving. And they also use the retrofitting option to reduce the power and energy requirement. The heat pump based system needs a small supplementary boiler whereas for scenario 1 the gas boiler ensures the highest share of energy to be delivered. Through the efficient way of producing heat when the heat pump is available, less buildings get refurbished. The energy demand is reduced by around 15% through retrofits. The solar thermal energy plays a key role in the gas boiler scenario with a solar fraction of over 70%. In the heat pump scenario, the share of solar is reduced to around 12%. The energy storage

Table 7.4 – Overview of which building is refurbished (=1) based on two scenarios: in the first solar thermal panels and a gas boiler can be used, in the second a central heat pump is available in addition to the previously mentioned technologies. Both use a daily and seasonal storage.

#	ID	Scenario 1	Scenario 2
1	356	1	0
2	559	1	0
3	602	0	1
4	1104	1	1
5	1341	1	0
6	1525	1	0
7	1710	1	0
8	1722	0	0
9	1819	1	1
10	1901	1	1
11	2079	0	0
12	2082	1	1
13	2084	1	1
14	2454	0	1
15	2458	1	0
16	3424	1	0
17	3429	1	0
18	4323	1	1
19	5030	1	1
20	5357	1	0
21	5375	1	1
22	5769	1	0
23	6133	0	0
24	6867	1	1
	Sum	19	11

follows the solar thermal collector installation size: when many panels are installed, also the storage is big and vice versa.

7.4 Discussion and Conclusion

This model completes the urban energy design options. Introducing refurbishment is a key point of bringing demand and supply together: as a function of the heat supply and the buildings demand, buildings are energetically refurbished. Different ways are shown on how to design an urban energy system with thermal storage and a stochastic renewable heat source: the most complete one is considering demand and supply as a variable at the same time, because they mutually influence each other. This approach shows directly which building

Table 7.5 – Sizing and operating of the 24 example buildings

Technologies	Scenario			
	Heat Pump Sizing	Gas and Solar Sizing	Heat Pump Energy	Gas and Solar Energy
gas boiler	31 kW	417 kW	26 MWh	681 MWh
heat pump	446 kW_{el}	-	2 900 MWh	- MWh
solar thermal panels	671 m^2	5 100 m^2	377 MWh	1 720 MWh
daily storage	15 m^3	100 m^3		
seasonal storage	33 m^3	6 000 m^3		
		Energy Sum:	3 303 MWh	2 401 MWh

should be retrofitted. It is a very practical solution compared to the standard in other studies that provide district level results such as in *Jennings et al.* [2014].

The model chooses the optimum to find the best way: for the heat pump case, rather the medium energy consumption buildings are being refurbished, whereas in the other case a clear selection from the most energy intensive buildings are chosen.

The long term thermal energy storage, however, is only interesting for very few scenarios: the demand temperatures and overall heat power requirement must be low. As shown, existing buildings, even after refurbishment do not have a low enough demand for an only renewable powered system as an optimal solution.

Practically this model might run into the limitation of property rights, because it uses a technocrat's approach. However for a given community or/and a big property owner are stakeholder of a district heating network project, the concerned buildings can be integrated in the solution for finding the optimal energy system design. Alternatively, only buildings that might eventually be retrofitted, for example the community owned ones, could be included.

7.5 Future Works

The current MILP formulation of this model could be changed to an LP formulation assuming that the building could be retrofitted also in between the current and the completely retrofitted state. One alternative can be to only change the windows. Then, an LP formulation could replace the MILP one to reduce the number of binary variables and accelerate the search.

It would be also interesting to see whether an annual formulation instead of the typical day approach lead to a similar result. If yes, a multi-stage problem could be written that can take every year a new decision on refurbishing or not.

8 Integrating Uncertainty into System Design

"Uncertainty is not an accident of the scientific method, but its substance." [*Saltelli et al.*, 2008, p.3]

8.1 Introduction

The lifetime of the equipment used in district heating systems such as a boiler can be estimated to be 15 or 20 years. The network delivering the heat to the consumer has a longer lifetime, when used correctly. The pipes can survive the equipment by a factor of three or more times more before the pipes themselves need to be exchanged. This gives a typical lifespan of a district heating network between a minimum of 20 years to as long as 60-80 years[1].

Giving such long planning horizons, normally a forecast is done to justify a decision of installing a district heating system. Often, only few scenarios are considered, each of them assuming that the future is known. Instead of choosing some scenarios (manually), [*de Neufville*, 2003] suggests integrating uncertainty into the system design. In our case, we want to evaluate how we can protect ourselves against wort cases.

8.2 Defining Uncertainty

Two main uncertainties exist:

- model uncertainty or the fact of simplifying a phenomenon and
- parameter uncertainty or the fact that the phenomenon varies.

Each model represents a simplification of a real system. Each simplification can lead to systematic or random errors in the model's output. Model uncertainty can be identified through model changes, model cross checking or through expert assessment. Further reading

[1] based on verbal communication with our industry partners

on model uncertainty and how to cope with it can be found in [*Cullen and Frey*, 1999, Chapter 3]. Within this work, it is not further considered.

Models contain parameters. The actual value of a parameter might not be known or simply differ from the one used in the model. Parameter uncertainty is therefore inherent to the model. The effect of parameter uncertainty is addressed in this work.

8.2.1 Addressing Uncertainty

[*Dubuis and Maréchal*, 2012] propose sorting the model parameters into two groups: certain parameters θ_c and uncertain parameters θ_u. Uncertain parameters are further separated into short term $\theta_{u,s}$ and long term $\theta_{u,l}$ parameters. $\theta_{u,s}$ can be fixed during the detailed planning or operation of the system (e.g. unit performance). The building occupancy is an example for $\theta_{u,l}$, because it will remain uncertain over the lifetime of the system.

[*Saltelli et al.*, 2002] propose a scientific approach from the uncertainty classification to a global sensitivity analysis. They explain the classical one parameter at a time sensitivity analysis (OAT), which has limited explanatory power over to systematic OATs to variance based methods.

During the factor fixing, where parameters are classified into important and non relevant (that is: can be fixed at any value within the range without impacting the output), the method of [*Morris*, 1991] estimates the output's variation based on systematic input variation. The number of run increase in as a linear function of the number of parameters. [*Saltelli et al.*, 2010] proposes a variance based method calculating the exact variation at higher computational cost.

8.2.2 Applications in the Field of Urban Energy System Design

For the field of urban energy system design, [*Keirstead and Shah*, 2013] apply Saltelli's approach of uncertainty and sensitivity analysis. On city wide studies, few examples exist. [*Kavgic et al.*, 2013] identifies the parameters with the highest influence on Belgrad's energy demand. In the building simulation community, frequently only One-At-a-Time sensitivity analysis methods are used. Few works use different approaches such as the one proposed by Saltelli: two examples are [*Hopfe and Hensen*, 2011] or [*Jaraminiene and Juodis*, 2006] estimating the building's performance and its heat demand.

[*Keirstead and Calderon*, 2012] uses Saltelli's approach with a quasi Sobol's quasi random sampling strategy to challenge current policy making. [*Fazlollahi*, 2014, Chapter 8] applies a variance based sensitivity analysis to the Pareto frontier resulting from a multi-objective optimization. Points behind the frontier are not considered and only economic parameters are varied. [*Moret et al.*, 2014a] use a simplified energy system model to address the lack of concep-

tual frameworks integrating uncertainty into strategic planning. They classify uncertainties and apply a global sensitivity analysis showing the total cost of a robust energy system.

8.3 Model Definition

The community of Verbier, a mountain ski resort in Switzerland with a fixed population of about 8000 people is chosen for the case study, because they would like to install a new district heating system as part of their energy strategy. The energy system is modeled in an MILP framework with the following equipment:

- solar thermal panels,
- a seasonal thermal storage,
- gas boiler,
- heat demand from a planned neighborhood of Verbier, that is to be built and
- a central district heating system linking demand and supply.

The objective function is cost minimization, the annual operating costs are calculated as well as the annualized investment costs. The annual investment costs are multiplied by the annualization factor of 0.0782 considering a payback time of 25 years and an interest rate of 6%.

The hourly heat demand is modeled with a building simulation software [*Page et al.*, 2014] based on annual measurements, meteorological and physical building data. The production of solar thermal panels is calculated with the same meteorological data for different temperature levels (detailed description in Chapter 5). The seasonal storage is a detailed model with nine discrete temperature levels of 10 Kelvin each between 10 and 95 °C. A k-medoids clustering, partitioning around medoids according to[*Kaufman and Rousseeuw*, 2005, Chapter 2], reduces the input data, heat load and solar irradiation, from 8760 hours to a data set of 12 representative days, each day with 24 hours and one peak load period.

The result of the clustering is shown in Figure 8.1 for the heat load: the red line shows the relative heat load, the blue one tries to fit this line with 12 typical day choices. Each typical day is repeated according to the number of days in this month. An extreme day is introduced ensuring that the maximum power appears. January and February are grouped together, they are very similar in this data set. Each month is represented by the most representative day repeated by the number of days in this month. This approach ensures that energy and the power balance are fulfilled at the same time compared to normal averages. Not all trends can be covered with this approach, however the input size reduction from 8760 to 288 allows the optimization model to run. The same approach is used for the solar energy in parallel, so that a typical day is chosen, when both data sets fit best.

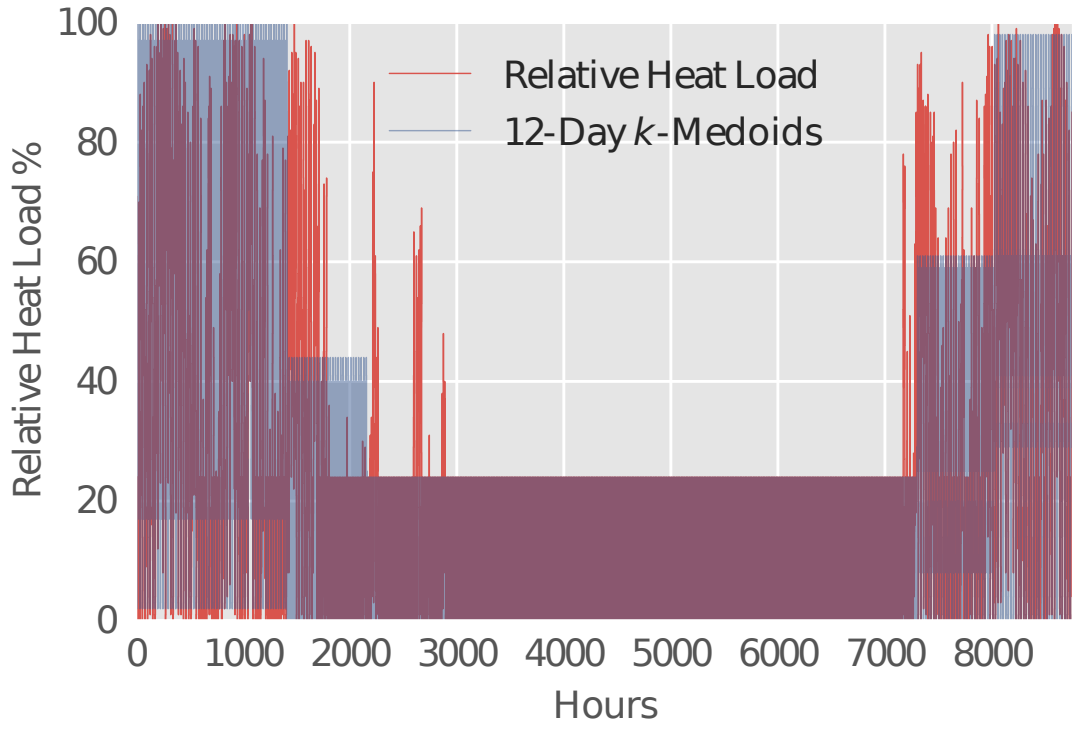


Figure 8.1 – The original heat load and the clustered heat load using 12 typical days to represent the year

8.4 Methodology

[*Moret et al.*, 2014a] proposes the following four steps:

1. All parameters are identified and if possible grouped.
2. The uncertainty classification step classifies and assigns an uncertainty range to the parameters or parameter groups.
3. A global sensitivity analysis [*Saltelli et al.*, 2008, p.11] allows to classify the uncertain parameters into influential and non-influential with regard to the output.
4. A robust optimization with influential parameters is run showing how to protect against the worst case(s).

Each section is presented individually with a short methodology and its results.

8.4.1 Parameter Identification and Grouping

The complete list of parameters is studied. Parameters such as duration of a time slice or number of repetition of key day within a year are assumed to have no uncertainty. Cost values for an equipment need to be studied further, because their values vary. Parameters that exist in every time slice are grouped to one factor, such as the hourly heat demand

$\dot{Q}_{demand,t}$ to the heat demand $\dot{Q}_{demand}$. This significantly reduces the number of parameters by the number of time slices (288) listed in Table 8.1 to be studied in detail. Their values

Table 8.1 – Model parameter list after identification and grouping

#	Parameters	Units	Description
1	C_{inv}(Boiler)	CHF/kW	Linear CAPEX
2	C_{inv}(Solar)	CHF/m^2	Linear CAPEX
3	C_{inv}(Storage)	CHF/m^3	Linear CAPEX
4	$c_{natural\ gas}$	CHF/kWh	current Swiss Natural Gas Price
5	$T_{nom\ demand\ heating}$	$°C$	Nominal Space Heating Demand Temperature
6	$\dot{Q}_{demand}$	kW	Heat Load of Consumers
7	$\dot{Q}_{storage\ heat\ loss}$	$W/(m^2 \cdot K)$	Heat Losses of thermal Energy Storage
8	ε_{Boiler}	-	Thermal Efficiency of a Boiler
9	ε_{Solar}	-	Thermal Efficiency of Solar Thermal Panels
10	Time Slice Duration	h	Time Slice Duration
11	c_p	$J/(kg \cdot K)$	Specific Heat Capacity of Storage Medium
12	$T_{Storage\ Surrounding}$	$°C$	Temperature surrounding the Storage
13	$M_{maximal\ Storage}$	kg	Maximal Storage Mass

can be found in Table 8.2. The values "operating time" and specific heat capacity of water c_p are considered to be certain, fixed and eliminated from further analysis. The storage is assumed to be placed in a construction adapted to its size with a constant $T_{Storage\ Surrounding}$ and maximal storage size of $M_{maximal\ Storage}$. Both parameters are also not considered any further. The nominal heating supply and return temperatures $T_{nom\ demand\ heating}$ are taken from current Swiss norms and validated or changed through expert's experience. Supply and return are considered together to be a set of temperatures, thus one of the two needs to be included in the analysis.

The cost data base CostDBCREM[2] [*Poumadère et al.*, 2015] contains prices for realized Swiss installations for different unit sizes. CAPEX includes equipment costs and installation, OPEX operation and maintenance.

8.4.2 Parameter Classification

The remaining 9 parameters are : the investment costs of a boiler, solar thermal installations and storage, the operating cost of the gas boiler with the gas price, the efficiency of the boiler and solar thermal panels together with the storage losses and the heat demand and temperature levels. They are classified according to the criteria defined in [*Moret et al.*, 2014a] (where additional information on the classification can be found):

C1: Does the parameter depend only on a choice made by the Decision Maker (DM)?
C2: Is it a here-and-now parameter?

[2]The cost data base CostDBCREM [*Poumadère et al.*, 2015] is provided by the CREM and contains prices for current Swiss installations for different unit sizes.

Table 8.2 – Model parameter values after identification and grouping

#	Parameters	Value	Comment/Source
1	C_{inv}(Boiler)	139 $kCHF/MW$	$\geq$ 10 MW Price from CostDBCREM[2]
2	C_{inv}(Solar)	282 CHF/m^2	$\geq$ 200 m^2 Price from CostDBCREM[2]
3	C_{inv}(Storage)	500 CHF/m^3	$\geq$ 10^3 m^3 Price from CostDBCREM[2] and [*Ochs and Müller-Steinhagen*, 2008]
4	$c_{natural\ gas}$	0.11 CHF/kWh	Swiss Natural Gas Price
5	$T_{nom\ demand\ heating}$	$35-70°C$	Expert Estimation as a Function of Building Age
6	$\dot{Q}_{demand}$	kW	Estimation via Simulation
7	$\dot{Q}_{storage\ heat\ loss}$	0.6 $W/(m^2 \cdot K)$	[*Ochs*, 2010, Chapter 3]
8	ε_{Boiler}	0.9	Thermal Efficiency
9	ε_{Solar}	$variable$	Panel data from SPF (Institute for Solar Technology)
10	Time Slice Duration	1 h	Time Slice Duration
11	c_p	4.18 $kJ/(kg \cdot K)$	Fixed for Water as Medium
12	$T_{Storage\ Surrounding}$	$10°C$	Constant Temperature assumed
13	$M_{maximal\ Storage}$	10^3 m^3	Maximum available Storage Size

C3: Does the parameter depend on other parameters?
C4: Can forecasts be made based on historical data?

The result of this exercise is shown in Table 8.3.

The investment cost variation for boilers and solar thermal panel installations is calculated on real data sets from realized projects in Switzerland. The data available in the cost data base CostDBCREM [*Poumadère et al.*, 2015] is fitted to a regression function(Figure 8.2). The point with the maximal distance to the regression defines the range for the parameter. The storage prize variation for tanks are based on realized studies in Germany. The data range for natural gas is estimated using the past's price variation[3].

8.4.3 Global Sensitivity Analysis

The sensitivity analysis draws systematically parameters out of the set of all the possible parameter values and runs the model with them. This allows closing the loop on the parameter uncertainty, because various different input parameters allow to study the output of the model. However, running the model with all input parameters often enough to quantify the output uncertainty can be computational heavy or infeasible due to the model size and or computational time. The current state of the art is presented by [*Saltelli et al.*, 2008] in detail.

A sensitivity analysis can generally be divided into 4 steps:

[3] An exact range will be specified once the data for Switzerland has been delivered

Figure 8.2 – The deviation is determined using the point the furthest away from the cost regression created on realized projects to identify the variation of investment costs for a given technology i

1. Classification: select k uncertain parameters from all parameters, discussed in Sections 8.4.1 and 8.4.2.
2. Sampling: create a sample set containing a number N of input vectors containing different values for the k uncertain parameters.
3. Computation: execute model for each sample to obtain Y^1, ..., Y^N outputs.
4. Evaluation: calculate sensitivity of outputs.

For each step different approaches exist.

$$S_{Ti} = \frac{E(V(Y|X_{\sim i}))}{V(Y)} \tag{8.1}$$

The total sensitivity S_{Ti} of the i-th input is defined as the ratio of the conditional expectation E of the output variance $V(Y)$ while varying only X_i keeping all other parameters fixed and $V(Y)$. $S_{Ti} = 0$ is a necessary and sufficient condition to consider that X_i is a non-influential parameter [*Saltelli et al.*, 2008, p.34]. During application this rarely happens. Therefore the total sensitivity index is normalized and so that the impacts can be compared to each other. Even for simple models with low values of k, the computation of the sensitivity effect can take a long time.

Table 8.3 – Application of [*Moret et al.*, 2014a]'s uncertainty classification to MILP

#	Parameters	C1	C2	C3	C4	Range	Comment
1	C_{inv}(Boiler)	✓				±3.00%	Variation in the data based on realized projects in Switzerland, Here-and-now parameter
2	C_{inv}(Solar)	✓				±3.00%	Variation in the data based on realized projects in Switzerland, Here-and-now parameter
3	C_{inv}(Storage)	✓				±10.00%	Variation in the data from catalogs, Here-and-now parameter
4	$c_{natural\ gas}$				✓	±50.00%	Forecast unreliable
5	$T_{demand\ heating}$	✓			✓	±10.00%	Norms and heating equipment used
6	$\dot{Q}_{demand}$			✓	✓	±10.00%	Only annual measurements available, hourly values based on simulation
7	$\dot{Q}_{storage\ heat\ loss}$			✓		±20.00%	Thermodynamics sets boundaries, long term studies exist
8	ε_{Boiler}			✓		±5.00%	
9	ε_{Solar}			✓	✓	±5.00%	

MILP Models and Sensitivity

For MILP models, [*Williams*, 2013, p.132] points out the importance of a sensitivity analysis: "In many practical situations stable solutions are more valuable than optimal solutions." In MILP models, such as the ones discussed in this work, a sensitivity analysis can lead to long calculation times, because a variation of a parameter might lead to a different combination of integer variable values being in an optimal solution. If only few constraints are leading to changing integer values, a reformulation towards an LP problem is an option [*Williams*, 2013, page 239]. A reformulation of the constraint with the changing integer variable by adding a continuous surplus to it and also to the objective function, leads to a (semi-)continuous formulation with an objective value that does not change suddenly anymore.

Method Choice

With Morris's Elementary Effect (EE) method[*Morris*, 1991], S_{Ti} can be estimated as a linear function of the number of parameters to be studied. Elementary Effect is a individually randomized One-At-a-Time Global Sensitivity Analysis [*Saltelli et al.*, 2002, p.94] with a discrete sampling strategy: For each parameter k, r random trajectories are defined consisting of $(k+1)$ steps. The model is executed with only one input element at a time changing by the amount of $\pm\Delta$. $\pm\Delta$ is calculated as a function of p-levels with $\Delta = \frac{p}{2} \cdot (p-1)$. Fixing p is an open problem. [*Saltelli et al.*, 2002, p.102] suggests fixing it according to its distribution. r and p are to be chosen together to guarantee that all dimensions and their interactions are explored in sufficient depth. [*Saltelli et al.*, 2002] suggests $p = 4$ and $r = 10$, [*Morris*, 1991] sets $r = 4$.

For $j-th$ trajectory, the Elementary Effect of the $i-th$ input is calculated with (8.2):

$$EE_i^j = \frac{\left[Y(X_1,\ ...,X_{i-1},\ X_i+\Delta,X_{i+1},...,X_k) - Y(X_1,...,X_k)\right]}{\Delta}\ . \tag{8.2}$$

After the calculation of the Elementary Effects for each parameter, the following statistics can be calculated:

$$\mu_i = \frac{1}{r}\sum_{j=1}^{r} EE_i^j \tag{8.3}$$

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r} |EE_i^j| \tag{8.4}$$

$$\sigma_i^2 = \frac{1}{r-1}\sum_{j=1}^{r} (EE_i^j - \mu)^2 \tag{8.5}$$

μ_i^* contains the sum of all absolute values. Negative and positive EE_is are taken into consideration without compensating each other. It is a suitable proxy for S_{Ti}. Here, the optimized trajectory implementation of [*Campolongo et al.*, 2007] ensuring a uniform distribution of the trajectories is used within R [*R Core Team*, 2014].

In order to compare the effect of each inputs variation on different model outputs, the statistics μ, σ and μ^* are normalized using sigma-scaling of the EE (SEE) as proposed in [*Sin et al.*, 2009]:

$$SEE_i = EE_i \frac{\sigma_{x_i}}{\sigma_y}\ . \tag{8.6}$$

Saltelli ([*Saltelli et al.*, 2008, Chapter 4] and [*Saltelli et al.*, 2010]) proposes an enhanced variance method based method, from here on called Saltelli's method. It is based on Monte Carlo stimulation of the Sobol's indices. The implementation of it by [*Herman*, 2015] is used to verify the quality of the proxy found through the EE. (The R-package Sensitivity offers additional implementation of similar methods.)

Results of the Elementary Effects

The EE method is applied to the $k = 9$ uncertain parameters (hereafter "inputs") of the MILP model with $p = 10$ and $r = 100$ leading to 1000 model runs. For this example, a finer sampling interval has been chosen to cover the interval better compared to the proposed values. The monitored model outputs are:

1. Normalized Objective Value, the total annual cost,
2. Normalized Solar Panel Sizing,
3. Normalized Boiler Size,

4. Normalized Storage Size.

The bars height gives the normalized value of μ for each input-output couple. The computed statistics μ and σ are shown in Figure 8.3. The bars height gives the normalized value of μ for

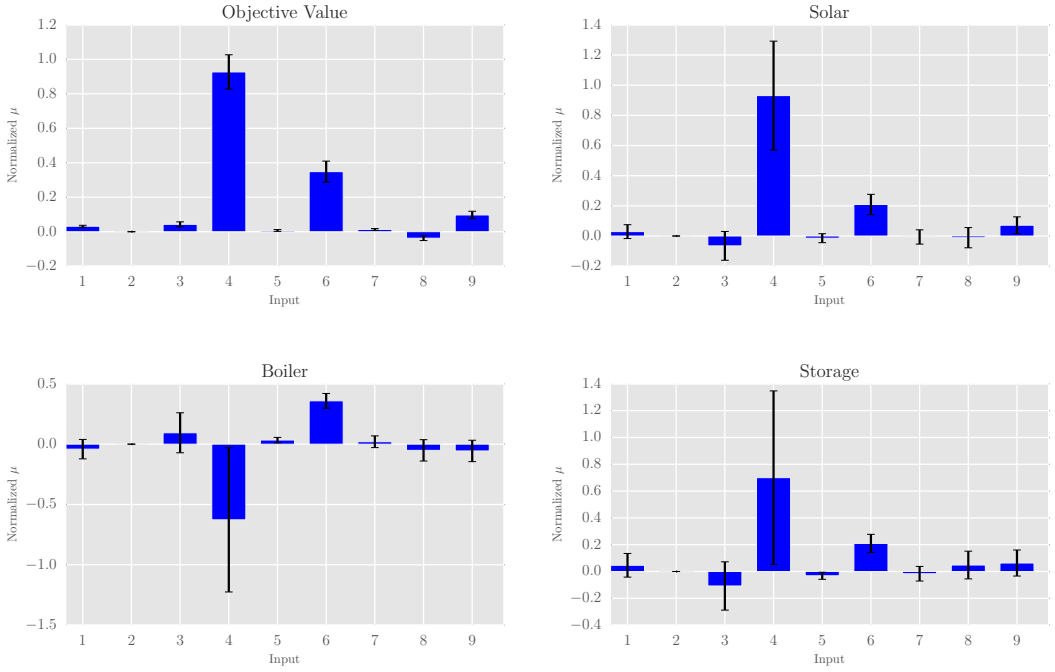


Figure 8.3 – Normalized μ for each input-output couple, inputs according to Table 8.3

each input-output couple. The brackets at the end of the bars gives the normalized statistic σ.

The normalized value of μ gives an idea of the effect of each input on each output in terms of their respective standard deviation. Considering the set of all input vectors $\boldsymbol{X} = (\bar{X}^1, ..., \bar{X}^m, ..., \bar{X}^{r(k+1)})$ and an arbitrary input vector composed by k input parameters $\bar{X}^m = (x_1^m, ..., x_k^m)$, σ_{x_i} is the standard deviation of x_i^m for all m in $[1, r(k+1)]$. Similarly, considering the set of corresponding output vectors $\boldsymbol{Y} = (\bar{Y}^1, ..., \bar{Y}^m, ..., \bar{Y}^{r(k+1)}) = (Y(\bar{X}^1), ..., Y(\bar{X}^m), ..., Y(\bar{X}^{r(k+1)}))$ and an arbitrary output vector $\bar{Y}^m = (y_1^m, ..., y_l^m)$ composed by the l monitored outputs, σ_{y_i} is the standard deviation of y_i^m for all m in $[1, r(k+1)]$. For a given input output couple, for instance (x_1, y_1), a μ_{11} value of 1 means that if input x_1 is increased by σ_{x_1} then output y_1 will, in average, increase by σ_{y_1}.

The computed statistic μ^* are shown in Figure 8.4. The bars height gives the normalized value of μ^* for each input-output couple. The bracket at the end of the bars give the 95 % confidence interval of the population.

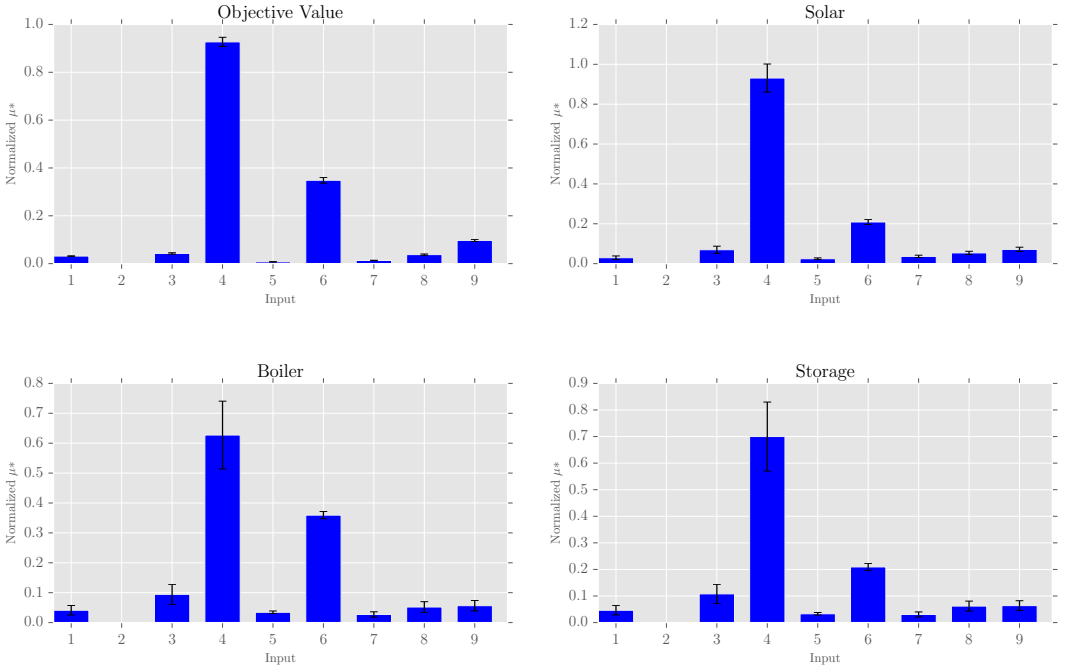


Figure 8.4 – The normalized μ^* statistics with its confidence intervals for all 9 parameter and four outputs, inputs according to Table 8.3

Results of Saltelli's Method

In order to confirm the results of EE, Saltelli's global sensitivity analysis [*Saltelli et al.*, 2010] method is applied to $k = 8$ uncertain inputs with $p = 10$ and $N = 2\,000$, i.e. $N(2k+2) = 36\,000$ model runs in order to compute second order effects. Input 5, the demand temperature level uncertainty, was removed because the uncertainty range is too narrow to provoke a change in the solution's configuration and in order to reduce computational effort. The monitored model outputs are the same. The computed statistics are the first order sensitivity index (S_{1i}) the second order sensitivity index (S_{2ij}) and the total sensitivity index (S_{Ti}). They express the share of the considered output's variance that can be attributed to the considered input i taken individually (S_{1i}), in conjunction with another input j (S_{2ij}) or considering all possible combination (S_{Ti}). They are shown in Figures 8.5, 8.6 and 8.7. The brackets at the end of the bars give the 95 % confidence intervals.

8.4.4 Discussion of the Sensitivity Analysis

Figure 8.3 shows the general behavior of the model. As an example, the effect of the boiler efficiency (input 8) on the objective value is negative which means that the higher the efficiency of the boiler, the lower the total cost of the solution, whereas for the solar panel's efficiency the opposite is observed. The influence of the natural gas cost on the boiler's size is negative

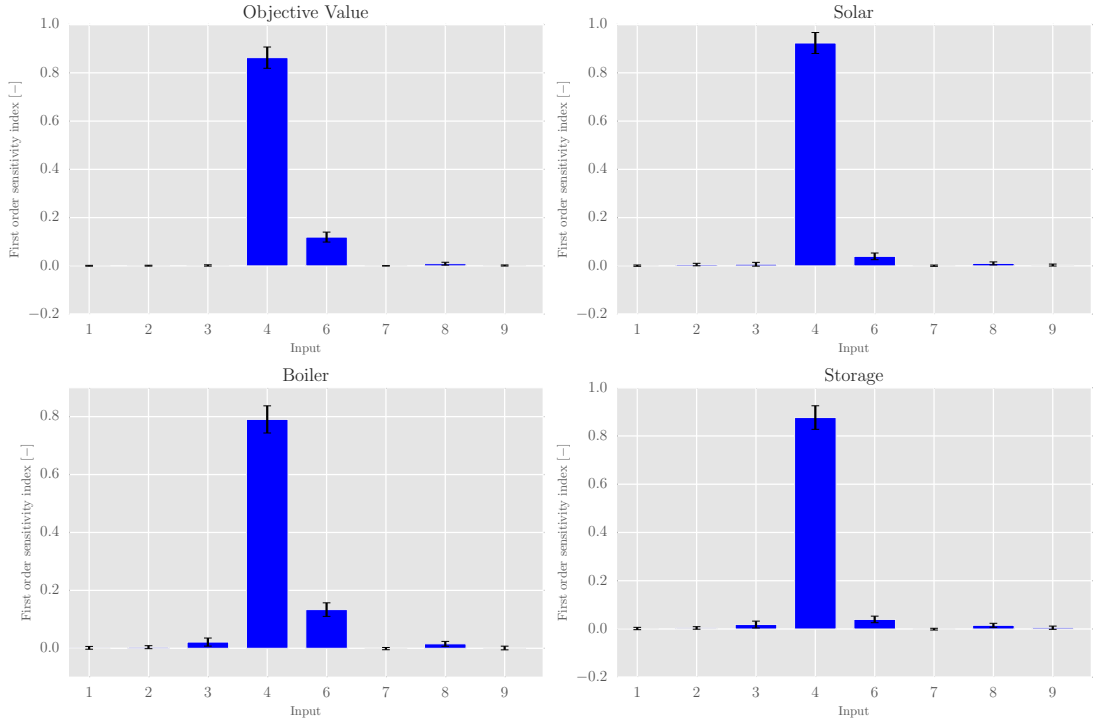


Figure 8.5 – Saltelli's first order sensitivity index for all 8 inputs and 4 outputs, inputs according to Table 8.3

whereas it is positive on all three other outputs. With an increasing price of gas, the boiler size decreases. The total costs also increase, together with the solar and storage sizes and that replace the part of the boiler's energy. All these observation are consistent with the problem physics and economics. Figure 8.3 and 8.4 show an important disparity in the input respective influences. The cost of natural gas (input 4) has the strongest influence on the total cost (highest μ and μ^* values), and on the solar and storage sizing as well. This behavior is consistent with both the high uncertainty range (±50 %) and the weight of this input in the objective function. The natural gas price is likely to be a key factor and a good candidate for robust optimization. On the contrary, solar investment costs uncertainty (input 2) has little to no influence on the outputs (low μ and μ^* values). This too, is consistent considering these inputs' narrower uncertainty range and reduced impact on the objective function due to annualization factor.

Factor fixing consists in identifying non-influential input that can be fixed anywhere in their uncertainty range without affecting the model's outputs. According to [*Saltelli et al.*, 2002, p.108-109], EE method is suitable for identifying non-influential parameters and $\mu^* = 0$ is a sufficient condition for a parameter to be considered non-influential. However, this condition is hardly ever met in practice [*Saltelli et al.*, 2002, p.108-109]. Figure 8.4 shows that only input 2 meets this condition for all outputs. Non-influential parameters can also be pointed out on

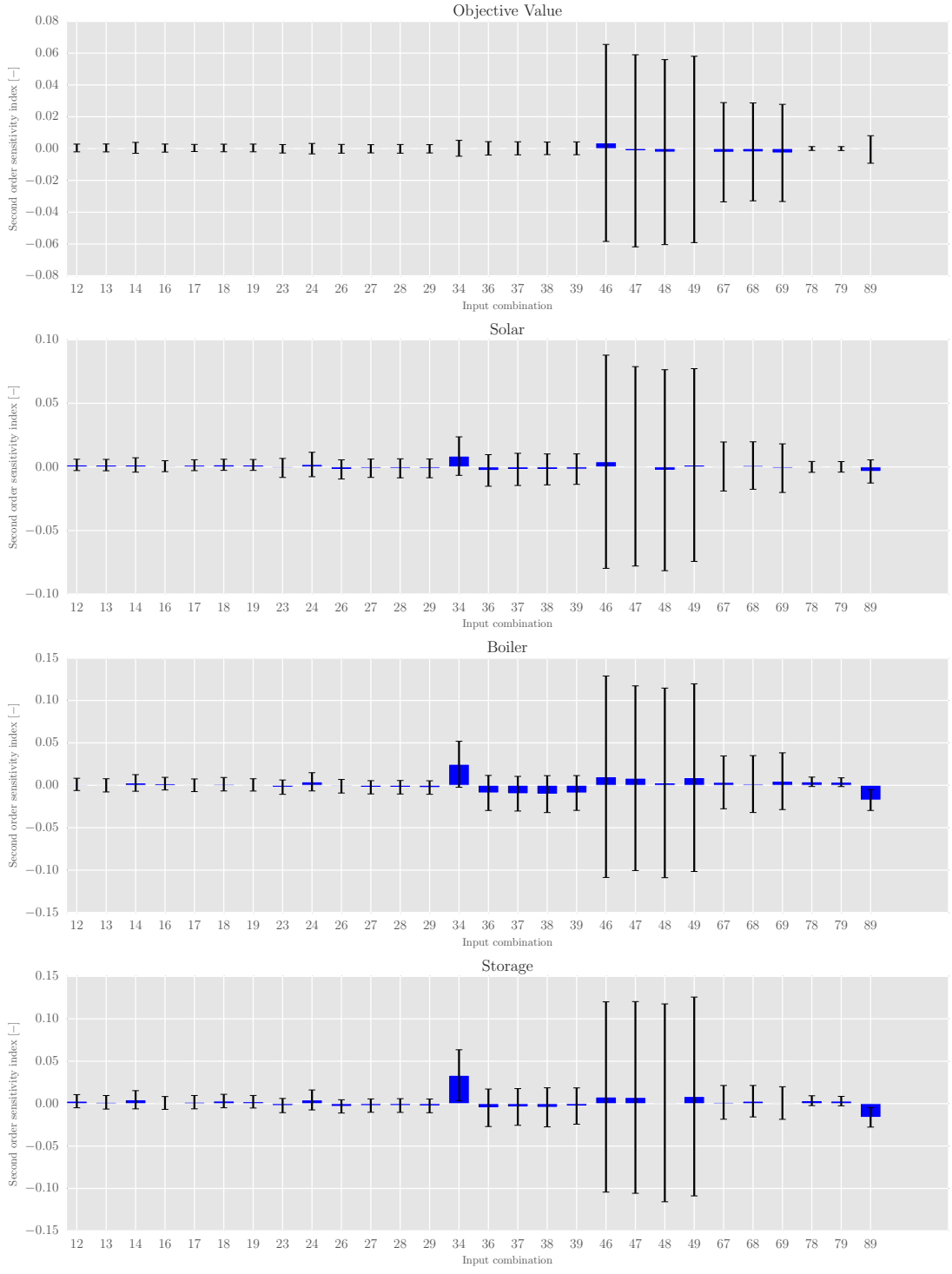


Figure 8.6 – Second order sensitivity index for all 8 inputs and 4 outputs, inputs according to Table 8.3: Two inputs are grouped together to show their combined effect on the output, such 3 and 4 together to 34 increase the storage size, while 4 and 7 together 47 have an impact on all outputs with a large bandwidth in both directions

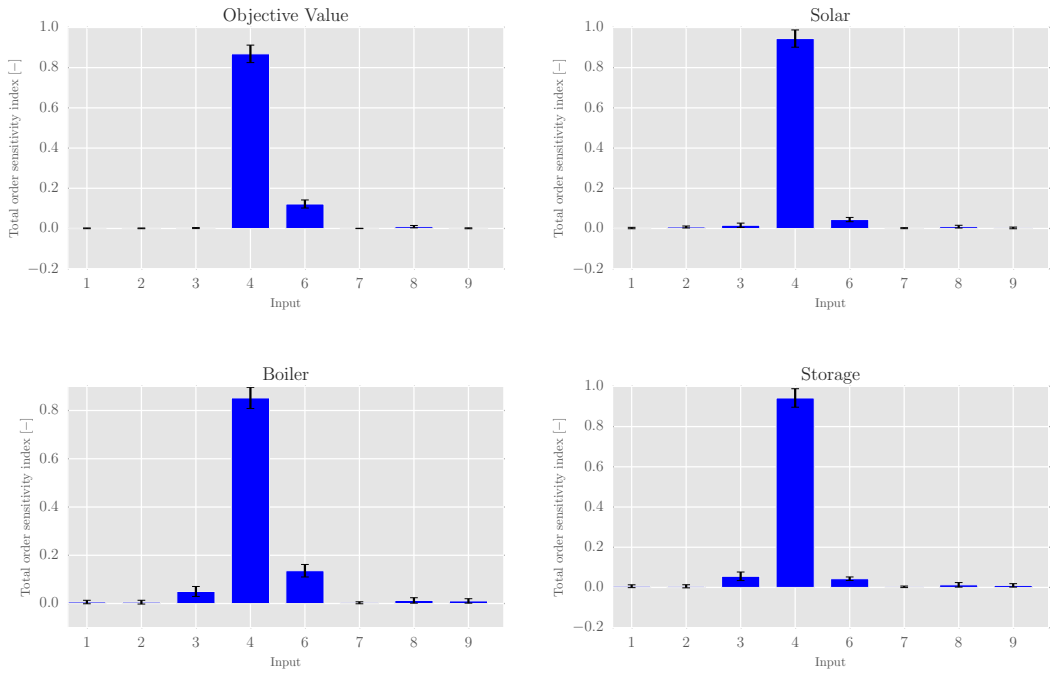


Figure 8.7 – Saltelli's total order sensitivity index of 4 outputs for the 8 inputs according to Table 8.3: parameter 4 and 6 are clearly the most influential ones, followed by 3.

a plot by identifying the parameters that sits within the standard error of the mean (SEM) as proprosed by [*Morris*, 1991] shown in the μ-σ plot Figure 8.8:

$$SEM_i = \frac{\sigma_{x_i}}{\sqrt{r}} \implies \bar{d}_i = \pm 2SEM_i \tag{8.7}$$

to determine whether the change of the $i-th$ input has a significant impact on the considered output. Figure 8.8 shows that no input except input 2 can be considered non-influential for all four outputs based on this approach. This graphical plot confirms the previously discussed results, however is a lot easier to read for the practitioner.

The results from Saltelli's method confirm that the cost of gas uncertainty has the most influence on all outputs, it accounts for more than 86 % of the total cost variance (Figure 8.7). The demand uncertainty (input 6) comes in second for all outputs. The storage investment cost uncertainty does not affect the total cost but it affects boiler, solar and storage size. Figure 8.9 shows the violin plot of the three most influential inputs (3, 4 and 6). The violin plots give detailed information about the average for each interval with the dashed line, the quantiles are indicated with by the dots. The width of each violin at a given y-axis value is proportional to the frequency of this y-value. The representation are useful to grasp the type of dependency between input and output. Compared to the boxplot, it is easier to see the frequency of all y-values. Figure 8.9 confirms the effect of input 4, natural gas price, decreasing the boiler size

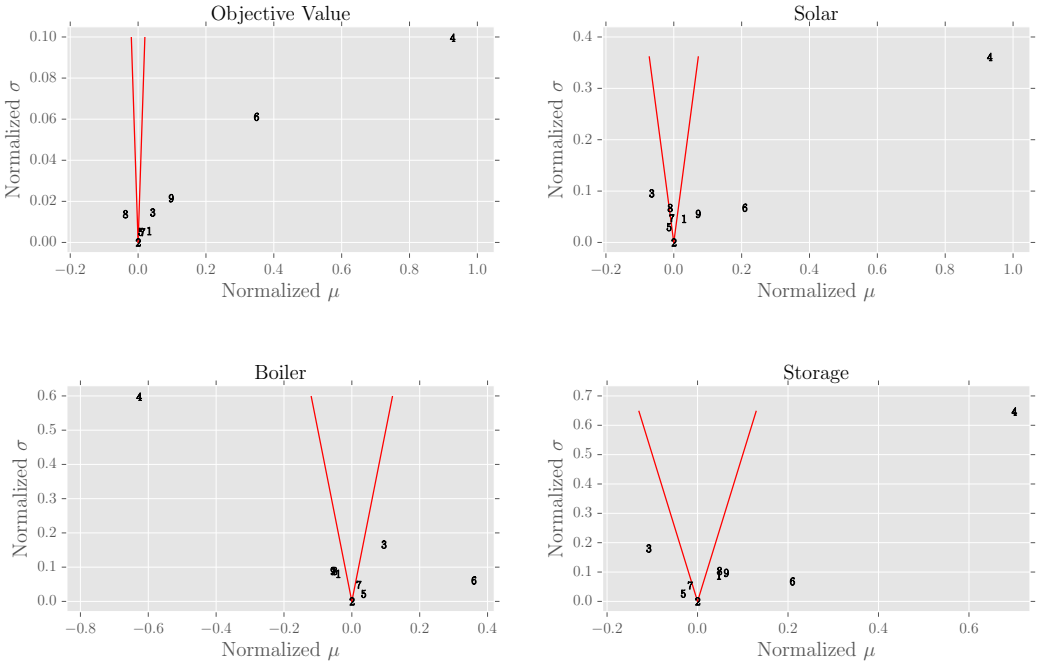


Figure 8.8 – Morris EE: normalized μ and σ with the range of the standard error represented by the red lines , inputs according to Table 8.3: parameter 4, 6 and 3 are the most influential parameters, whereas 2 can be discarded.

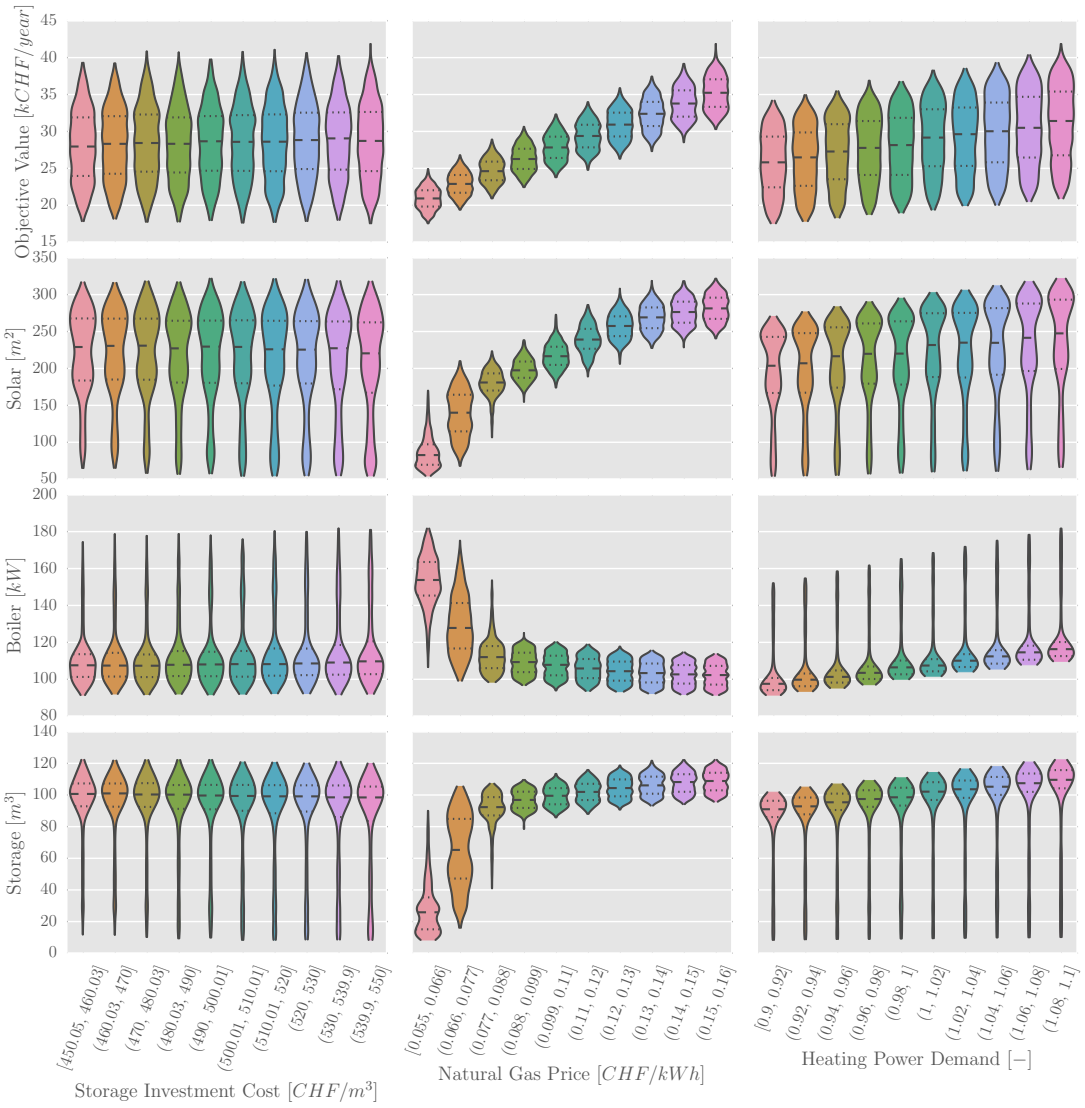


Figure 8.9 – Violin scatter plots show the same information as a box plot with the additional information for inputs 3, 4 and 6 and all outputs on the frequency of each input-output couple, inputs according to Table 8.3

while increasing solar and storage size. Based on the previous observation, one can infer that a trade-off exists between a solution involving more solar and storage or more boiler depending in a non-linear way on gas prices. The storage investment costs are on average non influential on the 4 studied outputs. Increasing the heat demand leads to an almost linear increase in the 4 outputs.

8.4.5 Conclusion of the Global Sensitivity Analysis

The statistics computed with the elementary effects method are based on a limited number of model runs and lead to qualitative assessment. They allow to identify tendencies, troubleshoot or corroborate the model, tell whether an input's uncertainty affects an output or not but they do not give absolute measurement of the sensitivity with a high level of confidence. Furthermore, due to the discrete levels, the input space is not covered entirely. However the comparison between figures 8.4 and 8.7 shows a high degree of resemblance. This proves that μ^* is indeed a good proxy for S_{Ti} in this case. Saltelli's method requires a higher number of model runs but with Sobol's low discrepancy sampling method, it allows to cover the input space more thoroughly. Computed sensitivity indices allow to quantify the sensitivity of the model to each inputs in terms of output variance with a high level of confidence. Depending on the context in which the model is used this quantification can be necessary.

For instance if the variance of the objective value has to be reduced to match a level that is acceptable for the stakeholders, robust optimization can be applied to the input with the highest total sensitivity index. The robust optimization is done in the next step.

8.5 Robust Optimization

Firstly, Soyster [*Soyster*, 1973] proposed a robust formulation where all uncertain parameters are at worst case. Comparing this result of Soyster's approach to the deterministic case, shows a highly suboptimal value (that is to say a much higher costs in a cost minimization problem). Based on Soyster's idea, *Bertsimas and Sim* [2004] proposed to put a varying number of uncertain parameters at worst case taking into consideration that rarely all parameters will be at their worst value at the same time. For this, Bertsimas and Sim introduced the protection parameter Γ_i that oscillates between 0, the deterministic value, and 1, the worst case value. When Γ_i is 1, the corresponding uncertain parameter i is put at worst case. The number of uncertain parameters i at worst case is equal to sum of all Γ_i. A sum of Γ equal to 0 represents the deterministic case, a sum of Γ equal to the maximum number of uncertain parameters is the worst case of Soyster. In between, the impact of uncertainty can be explored in a probabilistic way, because calculating all possible permutations is too demanding. Each time, the sum of Γ_i is fixed to a specific value and the uncertain parameters at worst case are shuffled.

After the identification of the key parameters, the model is reformulated to integrate a protection parameter Γ that ensures that a chosen number of uncertain parameters are at the worst case. Γ is integrated into the objective function or in the constraints to choose the most suboptimal combination out of all uncertain parameters. For parameters that are directly linked to the objective function such as costs in an cost minimization problem, this approach allows to calculate scenarios with a different amount of parameters at worst case as shown in [*Moret et al.*, 2014a].

Applying the approach of [*Bertsimas and Sim*, 2004] to an already existing model might however lead to an significant amount of work, especially if the parameters do not enter directly into the objective function but in several constraints at the same time. For natural gas this is a trivial exercise, the parameter enters directly into the objective function. Efficiency of a given utility is introduced as a change in the power of the utility. Through the heat cascade, it is present in multiple constraints making it hard to implement the parameter. Usually one Γ_i is introduced per uncertain parameter into one constraint. As we have here one parameter entering into multiple constraints, we propose a random sampling that is shuffling the Γ_is equal to 1 among all Γs while keeping the sum of all Γ_is equal to a value in between 0 and the worst case of all protection parameters equal to 1. For each $\sum \Gamma_i$ value, this leads to uniform distribution of samples.

For this approach, the parameters 4, the natural gas price, and 6, the heat demand, are separated into a daily values allowing to change their value in finer steps. In total, this leads to 27 parameters: 12 natural gas prices, 12 heat demands, the storage investment price, the boiler efficiency and the solar collector's efficiency.

For each parameter, the results between the robust formulation and manually changing the parameter are the same. The permutation of picking a Γ, setting a specific number of the 27 parameters to their worst value, is large: For $\sum_i \Gamma = 13$ or $\sum_i \Gamma = 14$, 20 058 300 different permutation without repetition exist. Therefore two sampling strategies are combined: Firstly, 100 permutations for each Γ are created, when possible e.g. for $\Gamma = 1$, only 27 permutations are possible. This ensures that low or high Γ values are represented in the sample. Secondly, 20 000 random integer samples based on a normal distribution are drawn over the whole range of all $\sum_i \Gamma_i$ at once, leading to more samples in the area where a high number of permutation exists. The combination of both ensures that a minimum number of permutations per $\sum_i \Gamma_i$ exist for the simulation of robust optimization.

However, the sampling bears the risk of not covering the whole range of parameter variation for all given Γ. When this happens, additional sample need to generated. Because the model is not changed, the computational costs for an individual model execution stays within the same range. This is not guaranteed for the robust reformulation because additional variables are introduced leading in the best case to an equivalent model execution time. In our tests, the execution time increased significantly.

8.5.1 Results of simulated robust Optimization

With a sample size of as low as a maximum of 100 runs per $\sum_i \Gamma_i$ already shows the trend for the robust optimization. When we ran the model more often, more information on the distribution and the extreme values for each output are found. The objective function, the

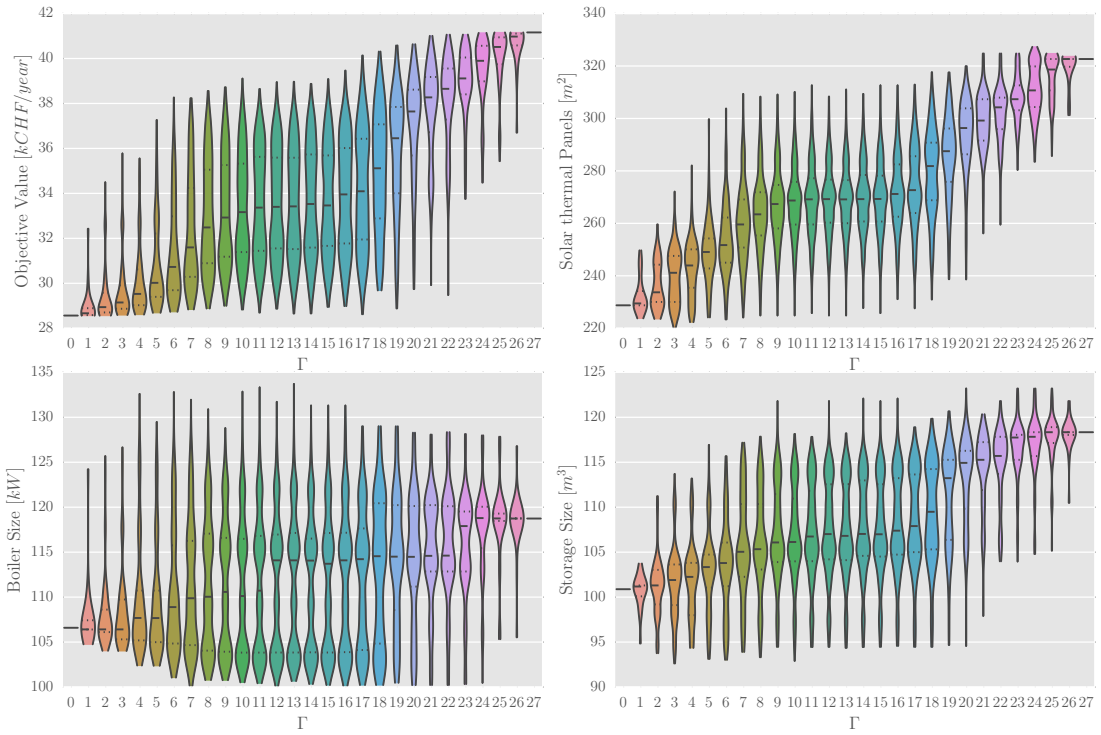


Figure 8.10 – Violin plot for the 4 outputs the annual costs and boiler, storage and solar thermal panel size with increasing Γ values after 25 000 runs

annual total costs, increase on average, when Γ increases. Compared to the increase in solar thermal collectors, the increasing boiler size is not significant. The boiler delivers almost the same amount of energy throughout all scenarios according to Figure 8.11. The increasing energy demand is compensated by the solar collectors. The storage is used for peak shaving during the winter and as a monthly energy storage for solar energy in the summer months.

8.5.2 Discussion of Results

When uncertainty is increased, e.g. the natural gas price and the heat load increases, more and more solar thermal panels combined with an increasing thermal energy storage are installed. This substantial increase in solar thermal panel and storage size can be seen (Figure 8.10). The additional heat demand is covered by renewable energy, the amount of energy delivered by the boiler remains on average constant (Figure 8.11). The operating cost increase comes from the increase in gas price. Compared to the study of [*Moret et al.*, 2014a], the investment to operating cost ratio changes in the other direction: the increase in gas price is too strong and

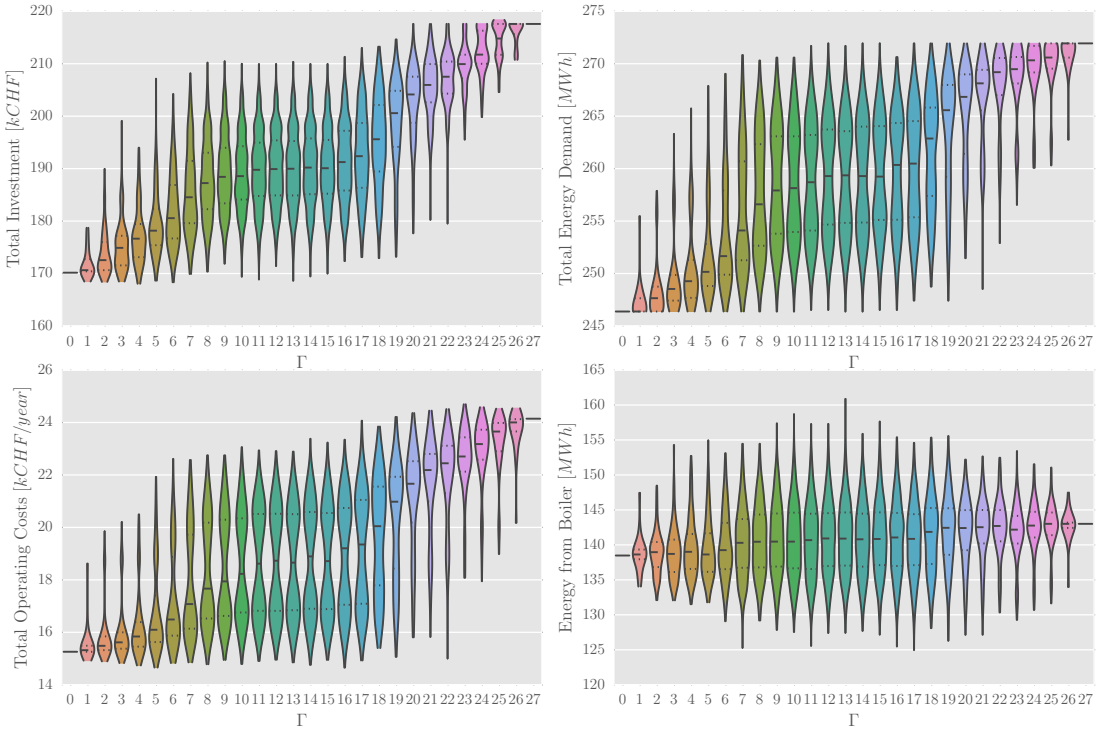


Figure 8.11 – 4 additional outputs of total investment and operating costs, the increasing energy demand and the energy delivered by the boiler with increasing Γ values after 25 000 runs

the trade off to solar long-term thermal energy storage is too expensive. Only renewable energy based solutions are proposed when the gas price goes over 0.20 CHF/kWh. Looking at long time horizon as 25 years for this case study, such a value might not even be that unrealistic. In [*Moret et al.*, 2014a], the boiler disappears, here no alternative peak power technology is available so that even with high gas prices, the boiler still delivers energy.

This approach can help to size the system while quantifying the additional costs of designing a robust system. With only few samples, the convex hull around all Γs can already be estimated around the violin plots, providing a first idea of how much variation to expect when increasing uncertainty for each output.

8.6 Conclusion

In a nutshell, the contributions of this work are:

1. real data for uncertainty classification,
2. uses a high time resolution with storage integration,
3. two methods for a global sensitivity analysis and
4. the simulation of robust optimization.

The here presented approach uses and classifies field data to specify the uncertainty of the inputs. With this approach, in a first step, the impact of each parameter on the system design can be quantified. Rarely, coherent data sets for all parameters exist. This allows to focus on the important ones for additional data collection or research, if possible. Simulating robust optimization as proposed in Chapter 8.5.1 is not found in literature: during the simulation of the robust optimization, the additional cost for a robust system can be determined. This allows to replace current sizing rules where simply an experienced based factor is added on the size of all equipment, especially the boilers. It can evolve into a methodology to design energy system by systematically taking uncertainty into account.

The robust optimization shows a shift towards a system with increasing renewable energy use and thermal energy storage. This simulation based approach to robust optimization can quantify envelops such as the overall costs to give an idea of the what it costs to build a robust energy system without having to reformulate the model.

8.7 Future Works

A more complex energy system should be analyzed with this approach. The sampling strategy can refined, for example based on sensitivity analysis results: using the insight of the proxy μ^* calculated with the EE or Saltelli's total sensitivity index S_{Ti} could significantly reduce the sampling, e.g. a parameter with a x times higher S_{Ti} value should be changed x times more. Further, a choice can be made to either put uncertain parameters in the objective function or in the constraints.

For a more enterprise oriented approach, the uncertain parameters could be separated into parameters that can be controlled by the enterprise such as efficiency of an installation compared to the heat demand or the gas prize which can not be controlled. Fixing the parameters that are out of the direct control at the worst case could provide insides on how much risk exposure the enterprise has: with the presented approach, the controllable but uncertain parameters can be varied to show from which Γ on losses are to be expected.

9 Conclusion

This thesis deals with the application of mixed integer linear programming techniques for the design urban energy systems with a special emphasis on the integration of heat storage in the system. Several aspects have been studied: from the problem formulation to the results analysis. Example from practical applications in Switzerland have been used to illustrate the developed methods.

Available data sources in Switzerland were compared and different modeling approaches were used to test the match between measurement and modeled estimate. In the current situation, the data quality is not high enough to take advantage of a detailed physical building model at urban scale. An energy signature model was as imprecise as the detailed physical building model. However the physical building model can be more easily updated to use better data. It is therefore worthwhile to use a building model that is at the same level of detail as the data found. This ensures quicker calculation and avoids to fix more default parameters for the model

The urban energy system design has been modeled as a mixed integer linear programming problem that aims at defining the equipment to be used in the system, i.e. choosing the technologies to be used, their sizes as well as their optimal operation. The definition of the optimal sizes requires to incorporate in the problem on the one hand a measure of the total costs over the life time of the equipment and on the other hand to define the size of the equipment such that the extreme conditions of the service will be satisfied. A generic framework using MILP formulation has been developed and several problem formulation have been investigated with the aim of limiting the computing time and the problem size in order to be able to solve problems compatible with the size of urban systems problems.

The problem has been formulated as a multi-period mixed integer linear programming problem. The multi-period problem represents the life time of the equipment by a set of successive typical periods, each being represented by a sequence of time slices that reproduces a certain number of times. In order to reduce the number of periods, data reduction techniques have been investigated.

Data reduction techniques were used to reduce an annual hourly data set significantly. Compared to the often used k-means approach, the *k*-medoids approach with PAM chooses from existing data points instead of looking for an averaging of them. Changing the distance measure between the points of the data set ensures that extreme points are well represented. We have demonstrated that the method allows one to represent both power and energy balance with an error of less then 5%. In general, between 7 to 12 days are enough to represent consistently a complete year including the seasonal and daily dynamics. These days, however do not cover the sequence of real alternations. This sequence still has a length of around 180 to 200 days when completely written. Using monthly representative days can bypass this problem. However, it is more difficult to represent a month with only one day, especially when a multi-dimensional input data set is used. For seasonal storage integration, representing each month with a single day is a compromise that has to be made in order to obtain a solution to the optimization problem.

A mixed integer linear programming model is proposed for the integration of solar thermal panels and the integration of 2 storage types: a daily and a seasonal storage. The model uses an optimal strategy to maximize the use of the solar energy in the system. This model is an alternative to the typical simulation approaches where the solar resources are investigated but not sized optimally. The proposed model allows to consider simultaneously the integration of solar heat with the use of advanced energy conversion technologies like heat pumps, revealing therefore the synergies between those technologies. It is worth mentioning as well that the integration of predictive control strategies is needed to make a better usage of the storage capabilities.

Both stochastic resource integration and seasonal storage can be treated at an adequate time resolution representing seasonal and daily variations. Instead of solar energy, other stochastic resources such as an industrial waste heat source can be considered. It has been shown that to be efficient, it is important to set up a storage system that has two dimensions: the daily variations being considered with a smaller storage tank while the seasonal being treated in a separate tanks. This allows to realize a better peak shaving during the daily variations while still storing heat across seasons in a larger tank. The model does not consider part load restrictions. In a first screening exercise in the pre-design phase, these solutions are justified under the hypothesis that equipment can also be used at full load due to a sufficiently big storage absorbing the excess heat.

A third storage type is proposed: when the building wall's composition is known, the building's heat capacity can be quickly and accurately estimated based on the method presented in Chapter 6. This capacity can be introduced as a heat storage into design of the energy system. It does not have an impact on the sizing of the peak load equipment, but it can save non-renewable energy by maximizing the use of renewable energy when available. When different feed-in-tariffs exist, the same optimization model can be used to decrease the operating costs. Because the model has almost no impact on the sizing of the peak load equipment, it can also be run after fixing the equipment.

In Chapter 5, different formulations have been investigated with the aim of reducing the size of the problem without compromising the precision of the results. In particular, the use of integer variables has been investigated and a linear programming that makes a limited use of integer variables has been tested. The LP formulation always overtakes the MILP when it comes to computing time which also indicates the possibility of formulating larger problems.

When the problem is well defined or reduced in (period-) size, the MILP formulation can be used to provide a solution that is more precise representing, the storage lid losses, ramp up and down times as well as part load efficiency by piecewise linearization. From a mechanical point of view, when equipment has ramp up and down times, a storage must at least be able to hold the ramp up and down times including the minimum run time for all equipment to allow for the hypothesis of ignoring them to be valid.

From the numerical perspective, significant computing time can be saved by increasing the MILP gap. However, the price of the reduced computing time is often high: solutions with almost similar objective value but very different values for the equipment size can be obtained. When the gap is lowered, they do not appear any more. It is therefore recommended to take attention to the problem definition, especially the bounds for the problem in order to define the list of competing solutions. With this respect to use of sensitivity analysis techniques as described in Chapter 8 of this thesis will become very important. In addition, depending on what input data is used, the MILP model shows very different resolution times. This is often linked to the integer variable being activated and whether it is easy or not to distinguish between their respective solutions. It is also be linked to the heuristics of the solver: when solving the same problem multiple times, the starting point (when not provided) can change because the solver heuristics chooses a different point leading to different resolution times.

The use of Cumulative Exergy Demand (CExD) as an objective function has been investigated with the goal of a more holistic view of the energy system. The aim was to represent the trade-off between the energy savings and the renewable energy integration and the grey energy needed for the production of the equipment. Changing the objective function to the CExD shows a higher penetration of renewables and bigger storage installations. Using CExD for the energy system introduces a holistic view on the energy system able to identify shifting exergy consumption from the consumption to construction or recycling. The case study of retrofitting shows that it can be uses to reduce the overall impact of the energy system considering the demand and the supply variable.

Solutions that perform poorly under the CExD criteria should be further studied to understand the source of an under-performance. The problem might be linked to the data source of the CExD values: Not many reference data sets exists up to now and on the few existing ones, often questions remain unanswered on the way the data set created, leading to a potential source of error when using it. When this source of error can be excluded, CExD solutions with poor performance should be discarded.

It is advantage to be able to re-run a model frequently, because uncertainty on parameter exists in addition to the fact that the build environment changes continuously. Fast and simple but not simplistic models ensure that more possibilities can be exploited.

Uncertainty on the economic and technical conditions has been investigated to show the importance of them on the decisions made for the design of the system. When introducing uncertainty into the case study, the use of more renewable energy combined with storage has been observed. The analysis of Morris was shown as an interesting proxy for a global sensitivity analysis with the advantage of a relative low computing time. The method identifies the key parameters of the solution. With the simulated robust optimization, the cost of a robust system can be quantified. Sizing rules, that are often applied based on experience, can be replaced by a systematic uncertainty classification approach. During the simulated robust optimization, the additional equipment needed for a robust system design can be quantified. Here it is a clear advantage (or even necessity) to have model with reduced integers or a LP model so that the number of runs necessary can be performed.

For the tool development, a simple and extendable framework has been created. The important technologies have been modeled including the impact of a thermal storage with a long sequence of time steps for optimization problems. When the data collection contains an uncertainty classification, all the steps shown here can be performed. The used simplified models are complementary to the existing simulation tools for heat planning. Besides the optimization, where all technology choices are activated, a user can also manually introduce them to show different scenarios. Both features are important to gain confidence in a tool. Of course, the usability and the user interface need work to be independent of the authors.

9.1 Perspectives

The results depend on the data quality. Even though, data is collected and studied carefully, it is difficult to ensure the quality. Compared to previous works, heat demand is based on dynamic building simulation, that on the input side needs more data but should on the output side provide a much higher accuracy than a linear signature. More detailed simulation requires more data but will become mandatory when one wants to represent and profit from the building's dynamics to maximize the use of renewable energy in the buildings. This will also be true when one wants to study the trade-off between building refurbishment and renewable energy sources integration. Research will therefore be needed to better represents the demand when buildings are becoming more efficient. However the question remains open how in this part the input data uncertainty changes the result.

For large scale MILP problems, it might be worth exploring the solvers tuning features. Different settings might increase the solving time. settings might increase the solving time while others will decrease the computing time. It is also important to study carefully the problem definition and the pertinence of the parameters used in the problem formulation in order to frame the solutions towards the solutions that have engineering significance. In particular, it

would be worth to study in more detail the definition of the objective function (cost) and the bounds as well as the pertinence of the linearization techniques to avoid generating solutions that at the end are not pertinent at the level of the decision variables. The risk, however, is that this procedure might have to be repeated with new input data every time.

After having integrated the heat perspective with different levels of detail, the electricity side would add the complementary part for a complete decision support on local level. With electricity, the integration of renewable energy sources using heat pumps, solar photovoltaic and co-generation technologies (including electrical storage) will take another dimension and create trade-off between thermal and electrical storage that will be worth investigating. Especially the PV model would allow to calculate the trade-off between solar thermal and PV.

For a case study of Verbier, it would also be useful to have the consumption of the ski-lifts. It then needs to be tested again to show whether the model is still manageable in terms of size and resolution time. When electricity is being included, transport might be next one to integrate because it might be linked to the electric (or gas) network deping on the type of engine used in the vehicle.

As mentioned the user-interface side should be developed ensuring intuitive usability and bringing the use of the models to the practical level.

Bibliography

A. D. Gordon. *Classification*, volume 82 Ed.2. Chapman and Hall, Boca Raton, 2nd ed. edition, 1999. ISBN 1-58488-013-9.

Hans-Jérg Althaus, Christian Bauer, Gabor Doka, Roberto Dones, Rolf Frischknecht, Stefanie Hellweg, Sébastian Humbert, Niels Jungbluth, Thomas Köllner, Yves Loerincik, Manuele Margni, and Thomas Nemecek. Implementation of Life Cycle Impact Assessment Methods. Technical Report 3, Swiss Centre for Life Cycle Inventories, St.Gallen, 2010.

Giovanni Angrisani, Michele Canelli, Carlo Roselli, and Maurizio Sasso. Calibration and validation of a thermal energy storage model: Influence on simulation results. *Applied Thermal Engineering*, 67(1–2):190–200, June 2014. ISSN 1359-4311. doi: 10.1016/j.applthermaleng.2014.03.012. URL http://www.sciencedirect.com/science/article/pii/S135943111400177X.

Roberto Aringhieri and Federico Malucelli. Optimal Operations Management and Network Planning of a District Heating System with a Combined Heat and Power Plant. *Annals of Operations Research*, 120(1-4):173–199, April 2003. ISSN 0254-5330, 1572-9338. doi: 10.1023/A:1023334615090. URL http://link.springer.com/article/10.1023/A%3A1023334615090.

A. Arteconi, N. J. Hewitt, and F. Polonara. State of the art of thermal storage for demand-side management. *Applied Energy*, 93:371–389, May 2012. ISSN 0306-2619. doi: 10.1016/j.apenergy.2011.12.045. URL http://www.sciencedirect.com/science/article/pii/S0306261911008415.

D. Bauer, R. Marx, J. Nußbicker-Lux, F. Ochs, W. Heidemann, and H. Müller-Steinhagen. German central solar heating plants with seasonal heat storage. *Solar Energy*, 84(4):612–623, April 2010. ISSN 0038-092X. doi: 10.1016/j.solener.2009.05.013. URL http://www.sciencedirect.com/science/article/pii/S0038092X09001224.

Helen Becker and François Maréchal. Energy integration of industrial sites with heat exchange restrictions. *Computers & Chemical Engineering*, 37(0):104–118, February 2012. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2011.09.014. URL http://www.sciencedirect.com/science/article/pii/S0098135411002857.

Helen Carla Becker. *Methodology and Thermo-Economic Optimization for Integration of Industrial Heat Pumps*. PhD thesis, STI, Lausanne, 2012. URL http://infoscience.epfl.ch/record/175150/files/EPFL_TH5341.pdf.

Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer, New York, 2nd ed. 2011 edition edition, June 2011. ISBN 978-1-4614-0236-7.

Grégoire Blanc, Gaëtan Cherix, and Loïc Darmayan. How to plan the desirable development of the energy use and supply of a local territory with the use of GIS tool. Dubrovnik, Croatia, September 2013.

Michel Bonvin. bsol. Technical report, 2004.

Lucien Borel and Daniel Favrat. *Thermodynamics and Energy Systems Analysis: From Energy to Exergy*. EPFL Press, 2010. ISBN 978-1-4398-3516-6.

S Brunold, R Frey, and U Frei. A comparison of three different collectors for process heat applications. Technical report, SPF-ITR, 1994.

Stephane L Bungener, François Maréchal, and Greet Van Eetvelde. A methodology for creating sequential multi-period base-case scenarios for large data sets. *Chemical Engineering Transaction*, 35, 2013. doi: http://dx.doi.org/10.3303/cet0918162.

Michael E. Bösch, Stefanie Hellweg, Mark A. J. Huijbregts, and Rolf Frischknecht. Applying cumulative exergy demand (CExD) indicators to the ecoinvent database. *The International Journal of Life Cycle Assessment*, 12(3):181–190, May 2007. ISSN 0948-3349, 1614-7502. doi: 10.1065/lca2006.11.282. URL http://link.springer.com/article/10.1065/lca2006.11.282.

Meinrad Bürer. *Multi-criteria optimization and project-based analysis of integrated energy systems for more sustainable urban areas*. PhD thesis, STI, Lausanne, 2003.

Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10):1509–1518, October 2007. ISSN 1364-8152. doi: 10.1016/j.envsoft.2006.10.004. URL http://www.sciencedirect.com/science/article/pii/S1364815206002805.

Massimiliano Capezzali, Gaëtan Cherix, Diane Perez, Jakob Rager, and Alain Duc. Swiss Federal Office of Energy 102775 FP WKK, March 2013. URL http://www.bfe.admin.ch/dossiers/03613/index.html?lang=de.

KW Childs, GE Courville, and EL Bales. Thermal mass assessment: an explanation of the mechanisms by which building mass influences heating and cooling energy requirements. Technical report, Oak Ridge National Lab., TN (USA), 1983.

Silvia Coccolo, Jérôme Kämpf, and Jean-Louis Scartezzini. The EPFL campus in Lausanne: new energy strategies for 2050. Turino, Italy, 2015.

Andrés Collazos, François Maréchal, and Conrad Gähler. Predictive optimal management method for the control of polygeneration systems. *Computers & Chemical Engineering*, 33 (10):1584–1592, October 2009. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2009.05.009. URL http://www.sciencedirect.com/science/article/pii/S0098135409001318.

Commune de Bagne. Evolution de la population, 2014. URL http://www.bagnes.ch/fr/Decouvrir/Statistiques-de-la-population/.

D. Connolly, H. Lund, B.V. Mathiesen, and M. Leahy. A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied Energy*, 87(4): 1059–1082, April 2010. ISSN 0306-2619. doi: 10.1016/j.apenergy.2009.09.026. URL http://www.sciencedirect.com/science/article/pii/S0306261909004188.

Reinerus Louwrentius Cornelissen. *Thermodynamics and sustainable development; the use of exergy analysis and the reduction of irreversibility*. info:eu-repo/semantics/doctoralThesis, Universiteit Twente, Enschede, November 1997. URL http://doc.utwente.nl/32030/. 00340.

Cynthia A. Cruickshank and Stephen J. Harrison. Heat loss characteristics for a typical solar domestic hot water storage. *Energy and Buildings*, 42(10):1703–1710, October 2010. ISSN 0378-7788. doi: 10.1016/j.enbuild.2010.04.013. URL http://www.sciencedirect.com/science/article/pii/S0378778810001544.

Alison C. Cullen and H. Christopher Frey. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Springer Science & Business Media, 1999. ISBN 978-0-306-45957-3. URL http://product-page-live-online.live.cf.i.springer.com/gp/book/9780306459566.

Vinicio Curti, Michael R. von Spakovsky, and Daniel Favrat. An environomic approach for the modeling and optimization of a district heating network based on centralized and decentralized heat pumps, cogeneration and/or gas furnace. Part I: Methodology. *International Journal of Thermal Sciences*, 39(7):721–730, July 2000. ISSN 1290-0729. doi: 10.1016/S1290-0729(00)00226-X. URL http://www.sciencedirect.com/science/article/pii/S129007290000226X.

A. Dalla Rosa, H. Li, and S. Svendsen. Method for optimal design of pipes for low-energy district heating, with focus on heat losses. *Energy*, 36(5):2407–2418, May 2011. ISSN 0360-5442. doi: 10.1016/j.energy.2011.01.024. URL http://www.sciencedirect.com/science/article/pii/S0360544211000259.

A. Dalla Rosa, R. Boulter, K. Church, and S. Svendsen. District heating (DH) network design and operation toward a system-wide methodology for optimizing renewable energy solutions (SMORES) in Canada: A case study. *Energy*, 45(1):960–974, September 2012. ISSN 0360-5442. doi: 10.1016/j.energy.2012.06.062. URL http://www.sciencedirect.com/science/article/pii/S0360544212005142.

Tony Day. *CIBSE TM41 Degree Days: Theory and Application*. CIBSE, London, 2006. ISBN 978-1-903287-76-7. URL http://www.cibse.org/knowledge/cibse-tm/tm41-degree-days-theory-application. 00000.

B. De Meester, J. Dewulf, S. Verbeke, A. Janssens, and H. Van Langenhove. Exergetic life-cycle assessment (ELCA) for resource consumption evaluation in the built environment. *Building and Environment*, 44(1):11–17, January 2009. ISSN 0360-1323. doi: 10.1016/j.buildenv.2008.01.004. URL http://www.sciencedirect.com/science/article/pii/S0360132308000085. 00057.

Richard de Neufville. Real Options: Dealing with Uncertainty in Systems Planning and Design. *Integrated Assessment*, 4(1):26–34, 2003.

Jo Dewulf, Herman Van Langenhove, Bart Muys, Stijn Bruers, Bhavik R. Bakshi, Geoffrey F. Grubb, D. M. Paulus, and Enrico Sciubba. Exergy: Its Potential and Limitations in Environmental Science and Technology. *Environmental Science & Technology*, 42(7):2221–2232, April 2008. ISSN 0013-936X. doi: 10.1021/es071719a. URL http://dx.doi.org/10.1021/es071719a.

Ibrahim Dincer and Marc A. Rosen. Exergy as a Driver for Achieving Sustainability. *International Journal of Green Energy*, 1(1):1–19, December 2004. ISSN 1543-5075. doi: 10.1081/GE-120027881. URL http://dx.doi.org/10.1081/GE-120027881. 00106.

Fernando Domínguez-Muñoz, José M. Cejudo-López, Antonio Carrillo-Andrés, and Manuel Gallardo-Salazar. Selection of typical demand days for CHP optimization. *Energy and Buildings*, 43(11):3036–3043, November 2011. ISSN 0378-7788. doi: 10.1016/j.enbuild.2011.07.024. URL http://www.sciencedirect.com/science/article/pii/S037877881100329X.

Roberto Dones, Christian Bauer, Rita Bolliger, Bastian Burger, Thomas Heck, Alexander Röder, Mireille Faist-Emmenegger, Rolf Frischknecht, Niels Jungbluth, and Matthias Tuchschmid. Life Cycle Inventories of Energy Systems: Results for Current Systems in Switzerland and other UCTE Countries. Technical report, Swiss Centre for Life Cycle Inventories, 2007.

V. Dorer, R. Weber, and A. Weber. Performance assessment of fuel cell micro-cogeneration systems for residential buildings. *Energy and Buildings*, 37(11):1132–1146, November 2005. ISSN 0378-7788. doi: 10.1016/j.enbuild.2005.06.016. URL http://www.sciencedirect.com/science/article/pii/S0378778805001076.

M.S. Drew and R.B.G. Selvage. Sizing procedure and economic optimization methodology for seasonal storage solar systems. *Solar Energy*, 25(1):79–83, 1980. ISSN 0038-092X. doi: 10.1016/0038-092X(80)90408-9. URL http://www.sciencedirect.com/science/article/pii/0038092X80904089.

Matthias Dubuis and François Maréchal. Proposition of methodology for optimization of energy system design under uncertainty. In Ian David Lockhart Bogle and Michael Fairweather, editor, *Computer Aided Chemical Engineering*, volume Volume 30, pages 302–306. Elsevier, 2012. ISBN 1570-7946. URL http://www.sciencedirect.com/science/article/pii/B9780444595195500617.

John A Duffie and William A Beckman. *Solar engineering of thermal processes*. Wiley, Hoboken, 2013. ISBN 978-0-470-87366-3.

EPIQR and ESTIA. Un Outil d'aide à la décision pour réhabilitation des bâtiments d'habitation. Technical report, EPIQR, Lausanne, 2004. URL http://www.epiqr.ch/pdf/EPIQR_methode_04.pdf.

Heike Erhorn-Kluttig, Reinhard Jank, Ludger Schrempf, Armand Dütz, Friedrun Rumpel, Johannes Schrade, Hans Erhorn, Carsten Beier, Christina Sager, and Dietrich Schmidt. *Energetische Quartiersplanung: Methoden - Technologien - Praxisbeispiele*. Fraunhofer-IRB-Verl., Stuttgart, 2011. ISBN 978-3-8167-8411-1 3-8167-8411-9.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. Detecting Clusters Graphically. In *Cluster Analysis*, pages 15–41. John Wiley & Sons, Ltd, 2011a. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.ch2.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. Some Final Comments and Guidelines. In *Cluster Analysis*, pages 257–287. John Wiley & Sons, Ltd, 2011b. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.ch9.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. Front Matter. In *Cluster Analysis*, pages i–xv. John Wiley & Sons, Ltd, 2011c. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.fmatter.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. An Introduction to Classification and Clustering. In *Cluster Analysis*, pages 1–13. John Wiley & Sons, Ltd, 2011d. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.ch1.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. Measurement of Proximity. In *Cluster Analysis*, pages 43–69. John Wiley & Sons, Ltd, 2011e. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.ch3.

Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. Miscellaneous Clustering Methods. In *Cluster Analysis*, pages 215–255. John Wiley & Sons, Ltd, 2011f. ISBN 978-0-470-97781-1. URL http://dx.doi.org/10.1002/9780470977811.ch8.

Ralph Evins. A review of computational optimisation methods applied to sustainable building design. *Renewable and Sustainable Energy Reviews*, 22:230–245, June 2013. ISSN 1364-0321. doi: 10.1016/j.rser.2013.02.004. URL http://www.sciencedirect.com/science/article/pii/S1364032113000920.

D. Favrat, F. Maréchal, and O. Epelly. The challenge of introducing an exergy indicator in a local law on energy. *Energy*, 33(2):130–136, February 2008. ISSN 0360-5442. doi: 10.1016/j.energy.2007.10.012. URL http://www.sciencedirect.com/science/article/pii/S0360544207001892.

Samira Fazlollahi. *Decomposition optimization strategy for the design and operation of district energy system*. PhD thesis, EPFL-STI, Lausanne, February 2014. URL http://infoscience.epfl.ch/record/197534/files/EPFL_TH6130.pdf.

Samira Fazlollahi, Gwenaelle Becker, and François Maréchal. Multi-Objectives, Multi-Period Optimization of district energy systems: II-Daily thermal storage. *Computers & Chemical Engineering*, 2012.

Samira Fazlollahi, Gwenaelle Becker, and François Maréchal. Multi-objectives, multi-period optimization of district energy systems: II—Daily thermal storage. *Computers & Chemical Engineering*, 71:648–662, 2014a.

Samira Fazlollahi, Stephane Laurent Bungener, Pierre Mandel, Gwenaelle Becker, and François Maréchal. Multi-objectives, multi-period optimization of district energy systems: I. Selection of typical operating periods. *Computers & Chemical Engineering*, 65: 54–66, June 2014b. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2014.03.005. URL http://www.sciencedirect.com/science/article/pii/S0098135414000751.

Matteo Fischetti and Michele Monaci. Exploiting erraticism in search. *Operations Research*, 62 (1):114–122, 2014.

Aurélie Foucquier, Sylvain Robert, Frédéric Suard, Louis Stéphan, and Arnaud Jay. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23:272–288, July 2013. ISSN 1364-0321. doi: 10.1016/j.rser.2013.03.004. URL http://www.sciencedirect.com/science/article/pii/S1364032113001536. 00047 bibtex: Foucquier2013.

C.A. Frangopoulos, M.R. Von Spakovsky, and E. Sciubba. A brief review of methods for the design and synthesis optimization of energy systems. *International Journal of Applied Thermodynamics*, 5(4):151–160, 2002. URL http://www.scopus.com/inward/record.url?eid=2-s2.0-0036963331&partnerID=40&md5=6ee02937f2226aba7b719f53f6496f2c. cited By (since 1996)24.

Alexandros Gasparatos, Mohamed El-Haram, and Malcolm Horner. A critical review of reductionist approaches for assessing the progress towards sustainability. *Environmental Impact Assessment Review*, 28(4–5):286–311, May 2008. ISSN 0195-9255. doi: 10.1016/j.eiar.2007.09.002. URL http://www.sciencedirect.com/science/article/pii/S019592550700131X. 00206.

Léda Gerber. *Integration of Life Cycle Assessment in the conceptual design of renewable energy conversion systems*. Phd Thesis, EPFL, L, December 2012. URL http://infoscience.epfl.ch/record/182673?ln=en.

Léda Gerber, Martin Gassner, and François Maréchal. Systematic integration of LCA in process systems design: Application to combined fuel and electricity production from lignocellulosic biomass. *Computers & Chemical Engineering*, 35(7):1265–1280, July 2011. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2010.11.012. URL http://www.sciencedirect.com/science/article/pii/S0098135410003595.

Luc Girardin. *A GIS-based Methodology for the Evaluation of Integrated Energy Systems in Urban Area*. PhD thesis, STI, Lausanne, 2012. URL http://library.epfl.ch/theses/?nr=5287.

Luc Girardin, François Maréchal, Matthias Dubuis, Nicole Calame-Darbellay, and Daniel Favrat. EnerGis A geographical information based system for the evaluation of integrated energy conversion systems in urban areas. *Energy*, 35(2):830–840, February 2010. ISSN 0360-5442. doi: 10.1016/j.energy.2009.08.018. URL http://www.sciencedirect.com/science/article/pii/S0360544209003582.

Mei Gong and Göran Wall. On exergy and sustainable development—Part 2: Indicators and methods. *Exergy, An International Journal*, 1(4):217–233, 2001. ISSN 1164-0235. doi: 10.1016/S1164-0235(01)00030-9. URL http://www.sciencedirect.com/science/article/pii/S1164023501000309. 00139.

Ignacio E. Grossmann. Advances in mathematical programming models for enterprise-wide optimization. *Computers & Chemical Engineering*, 47:2–18, December 2012. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2012.06.038. URL http://www.sciencedirect.com/science/article/pii/S0098135412002220.

Ignacio E. Grossmann and Jorge Santibanez. Applications of mixed-integer linear programming in process synthesis. *Computers & Chemical Engineering*, 4(4):205–214, 1980. ISSN 0098-1354. doi: 10.1016/0098-1354(80)85001-0. URL http://www.sciencedirect.com/science/article/pii/0098135480850010.

Jean-Christophe Hadorn. Guide to Seasonal Heat Storage. Technical report, SIA and Federal Energy Office, Zürich, 1990.

Jean-Christophe Hadorn. ADVANCED STORAGE CONCEPTS FOR ACTIVE SOLAR ENERGY. Technical Report IEA-SHC Task 32 2003-2007, IEA, Geneva, 2008.

Eric Hahne. Cumulative Energy and Emission Balance of Large Solar Heating Systems. *Int.J. Applied Thermodynamics*, 2(1):37–43, 1999. ISSN 1301-9724. URL http://www.ijoticat.com/index.php/IJoT/article/viewFile/12/10.

P.-A. Haldi and D. Favrat. Methodological aspects of the definition of a 2 kW society. *Energy*, 31(15):3159–3170, December 2006. ISSN 0360-5442. doi: 10.1016/j.energy.2006.02.011. URL http://www.sciencedirect.com/science/article/pii/S0360544206000557.

Stig Hammarsten. A critical appraisal of energy-signature models. *Applied Energy*, 26(2):97–110, 1987. ISSN 0306-2619. doi: 10.1016/0306-2619(87)90012-2. URL http://www.sciencedirect.com/science/article/pii/0306261987900122.

Mehrdad Heidari Tari and Mats Söderström. Modelling of thermal energy storage in industrial energy systems the method development of MIND. *Applied Thermal Engineering*, 22(11):1195–1205, August 2002. ISSN 1359-4311. doi: 10.1016/S1359-4311(02)00044-3. URL http://www.sciencedirect.com/science/article/pii/S1359431102000443.

Jon Herman. Sensitivity Analysis Library (SALib), March 2015. URL http://jdherman.github.io/SALib/.

James S Hodges, James A Dewar, and Arroyo Center. *Is it you or your model talking?: A framework for model validation*. Rand Santa Monica, CA, 1992.

Volker H. Hoffmann. *Multi-objective decision making under uncertainty in chemical process design*. PhD thesis, ETH Zürich, 2001. URL http://www.library.ethz.ch/search/action/display.do?fn=display&doc=ebi01_prod004141517.

Christina J. Hopfe and Jan L. M. Hensen. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings*, 43(10):2798–2805, October 2011. ISSN 0378-7788. doi: 10.1016/j.enbuild.2011.06.034. URL http://www.sciencedirect.com/science/article/pii/S0378778811002830.

A. Hugo and E. N. Pistikopoulos. Environmentally conscious long-range planning and design of supply chain networks. *Journal of Cleaner Production*, 13(15):1471–1491, December 2005. ISSN 0959-6526. doi: 10.1016/j.jclepro.2005.04.011. URL http://www.sciencedirect.com/science/article/pii/S0959652605001083.

International Atomic Energy Agency, United Nations Department of Economic and Social Affairs, International Energy Agency, Eurostat, and European Environment Agency. *Energy Indicators for Sustainable Development: Guidelines and Methodologies*. International Atomic Energy Agency, Vienna, 2005. ISBN 92-0-116204-9. URL http://www-pub.iaea.org/mtcd/publications/PubDetails.asp?pubId=7201.

R. R. Iyer and I. E. Grossmann. Synthesis and operational planning of utility systems for multiperiod operation. *Computers & Chemical Engineering*, 22(7–8):979–993, July 1998. ISSN 0098-1354. doi: 10.1016/S0098-1354(97)00270-6. URL http://www.sciencedirect.com/science/article/pii/S0098135497002706.

Egle Jaraminiene and Egidijus Juodis. Heat demand uncertainty evaluation of typical multi flat panel building. *Journal of Civil Engineering and Management*, 12(1):69–75, 2006. ISSN 1392-3730. doi: 10.1080/13923730.2006.9636375. URL http://www.tandfonline.com/doi/abs/10.1080/13923730.2006.9636375.

Mark Jennings, David Fisk, and Nilay Shah. Modelling and optimization of retrofitting residential energy systems at the urban scale. *Energy*, 64:220–233, January 2014. ISSN 0360-5442. doi: 10.1016/j.energy.2013.10.076. URL http://www.sciencedirect.com/science/article/pii/S0360544213009432.

Eberhard Jochem, Daniel Favrat, Göran Andersson, Heinz Gutscher, Konrad Hungerbühler, Phillip Rudolf von Rohr, Daniel Spreng, and Mark Zimmermann. *Steps towards a sustainable development. A White Book for R&D of energy-efficient technologies*. Novatlantis, Altstätten, March 2004. URL http://www.novatlantis.ch/en/2000-watt-society/white-book.html.

Heinrich Jäckli. Carte de la consommation calorifique de la Suisse en rapport avec la superficie du terrain bâti, Karte des flächenspezifischen Wärmebedarfs der Schweiz, 1970.

Jouri Kanters, Miljana Horvat, and Marie-Claude Dubois. Tools and methods used by architects for solar design. *Energy and Buildings*, June 2012. ISSN 0378-7788. doi: 10.1016/j.enbuild.2012.05.031. URL http://www.sciencedirect.com/science/article/pii/S0378778812003040.

Leonard Kaufman and Peter J Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. Wiley, Hoboken, N.J., 2005. ISBN 0-471-73578-7 978-0-471-73578-6.

M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic. A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, 45(7):1683–1697, July 2010. ISSN 0360-1323. doi: 10.1016/j.buildenv.2010.01.021. URL http://www.sciencedirect.com/science/article/pii/S0360132310000338.

M. Kavgic, D. Mumovic, A. Summerfield, Z. Stevanovic, and O. Ecim-Djuric. Uncertainty and modeling energy consumption: Sensitivity analysis for a city-scale domestic energy model. *Energy and Buildings*, 60:1–11, 2013. ISSN 0378-7788. doi: 10.1016/j.enbuild.2013.01.005. URL http://www.sciencedirect.com/science/article/pii/S0378778813000224.

KBOB. Données des écobilans dans la construction. Technical report, Conférence de coordination des services de la construction et des immeubles des maîtres d'ouvrage publics KBOB c/o Office fédéral des constructions et de la logistique OFCL, 2014. URL https://www.kbob.admin.ch/kbob/fr/home/publikationen/nachhaltiges-bauen.html.

James Keirstead and Carlos Calderon. Capturing spatial effects, technology interactions, and uncertainty in urban energy and carbon models: Retrofitting newcastle as a case-study. *Energy Policy*, 46:253–267, July 2012. ISSN 0301-4215. doi: 10.1016/j.enpol.2012.03.058. URL http://www.sciencedirect.com/science/article/pii/S0301421512002637.

James Keirstead and Nilay Shah. *Urban Energy Systems: An Integrated Approach*. Routledge, Urban Energy Systems, 2013. ISBN 978-0-415-52902-0 0-415-52902-6.

James Keirstead, Mark Jennings, and Aruna Sivakumar. A review of urban energy system models: Approaches, challenges and opportunities. *Renewable and Sustainable Energy Reviews*, 16(6):3847–3866, August 2012. ISSN 1364-0321. doi: 10.1016/j.rser.2012.02.047. URL http://www.sciencedirect.com/science/article/pii/S1364032112001414.

Claudia Kemfert. Applied Economic- environment- energy Modeling for quantitative impact assessment. In *Proceedings of the EU Advanced Summer Course in Integrated Assessment methodology*, page 149, Maastricht, 2003. International Centre for Integrated assessment and Sustainable development. URL https://lehre.wiwi.hu-berlin.de/Professuren/vwl/uoe/lehre/ws0809/KemfertModelling.pdf.

Andreas Kemmler, Ale Piégsa, Andrea Ley, Mario Keller, Jakob Martin, and Giacomo Catenazzi. Analyse des schweizerischen Energieverbrauchs 2000 - 2011 nach Verwendungszwecken. Technical report, Swiss Federal Office of Energy, Bern, August 2012. URL http://www.bfe.admin.ch/themen/00526/00541/00542/02167/index.html?lang=en&dossier_id=02169.

Ian C. Kemp. *Pinch analysis and process integration : a user guide on process integration for the efficient use of energy*. Elsevier Butterworth-Heinemann, Amsterdam, 2nd ed. / ian c. kemp edition, 2007.

IC Kemp and EK MacDonald. Energy and process integration in continuous and batch processes. In *IChemE Symposium Series*, volume 105, pages 185–200, 1987.

Tom Kerr. CHP/DHC Country Scorecard: Denmark, 2015. URL http://www.iea.org/media/files/chp/profiles/denmark.pdf.

Karsten-Ulrich Klatt and Wolfgang Marquardt. Perspectives for process systems engineering—Personal views from academia and industry. *Computers & Chemical Engineering*, 33 (3):536–550, March 2009. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2008.09.002. URL http://www.sciencedirect.com/science/article/pii/S0098135408001737.

Richard E. Klosterman. Commentary: simple and complex models. *Environment and Planning B: Planning and Design*, 39(1):1 – 6, 2012. doi: 10.1068/b38155. URL http://www.envplan.com/abstract.cgi?id=b38155.

Ed Klotz and Alexandra M Newman. Practical guidelines for solving difficult mixed integer linear programs. *Surveys in Operations Research and Management Science*, 18(1):18–32, 2013. URL http://inside.mines.edu/~anewman/MIP_practice120212.pdf.

M. Kovarik. Comments on "sizing procedure and economic optimisation methodology for seasonal storage solar systems" by M. S. Drew and R. B. G. Selvage. *Solar Energy*, 26(2):187, 1981. ISSN 0038-092X. doi: 10.1016/0038-092X(81)90084-0. URL http://www.sciencedirect.com/science/article/pii/0038092X81900840.

Pierre Krummenacher. *Contribution to the heat integration of batch processes (with or without heat storage)*. PhD thesis, STI, Lausanne, 2002.

Jérôme Henri Kämpf and Darren Robinson. Optimisation of Urban Energy Demand using an Evolutionary Algorithm. In *Proceedings of the Eleventh International IBPSA Conference*, pages 668–673, Glasgow, 2009.

Roland Köhler and Sigrid Hanke, editors. *Schweizer Energiefachbuch 2013*. Number 30 in Schweizer Energiefachbuch. Kömedia, St, 2012. ISBN 978-3-9523902-5-2. URL http://www.xn--kmedia-wxa.ch/presse/pressemitteilung-energiefachbuch/.

Lesosai. Lesosai, March 2013. URL http://www.lesosai.com/.

Jens Libbe, Hadia Köhler, and Klaus J. Beckmann. *Infrastruktur und Stadtentwicklung*. Deutsches Institut für Urbanistik (Berlin), 2010.

D Lindenberger, T Bruckner, H.-M Groscurth, and R Kümmel. Optimization of solar district heating systems: seasonal storage, heat pumps, and cogeneration. *Energy*, 25(7):591–608, July 2000. ISSN 0360-5442. doi: 10.1016/S0360-5442(99)00082-1. URL http://www.sciencedirect.com/science/article/pii/S0360544299000821.

Bodo Linnhoff. *A user guide on process integration for the efficient use of energy*. Institution of Chemical Engineers, 1997.

Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, Junjie Wu, and Sen Wu. Understanding and Enhancement of Internal Clustering Validation Measures. *IEEE Transactions on Cybernetics*, 43(3):982–994, June 2013. ISSN 2168-2267. doi: 10.1109/TSMCB.2012.2220543.

Stuart Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):128–137, March 1982. URL https://www.google.ch/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CDUQFjAB&url=http%3A%2F%2Fwww.nt.tuwien.ac.at%2Ffileadmin%2Fcourses%2F389075%2FLeast_Squares_Quantization_in_PCM.pdf&ei=DR6RUomJBovo7Abak4HACQ&usg=AFQjCNH0o9GeLuW1FvP0Mr8qRJHBvyNBsw&sig2=khnRX7o_5a-nexVaOXzasg&bvm=bv.56988011,d.ZGU.

Miguel A. Lozano, Jose C. Ramos, and Luis M. Serra. Cost optimization of the design of CHCP (combined heat, cooling and power) systems under legal constraints. *Energy*, 35 (2):794–805, February 2010. ISSN 0360-5442. doi: 10.1016/j.energy.2009.08.022. URL http://www.sciencedirect.com/science/article/pii/S0360544209003624.

P.D. Lund. A general design methodology for seasonal storage solar systems. *Solar Energy*, 42(3):235–251, 1989. ISSN 0038-092X. doi: 10.1016/0038-092X(89)90015-7. URL http://www.sciencedirect.com/science/article/pii/0038092X89900157.

Peizheng Ma and Lin-Shu Wang. Effective heat capacity of interior planar thermal mass (iPTM) subject to periodic heating and cooling. *Energy and Buildings*, 47:44–52, April 2012. ISSN 0378-7788. doi: 10.1016/j.enbuild.2011.11.020. URL http://www.sciencedirect.com/science/article/pii/S0378778811005664.

Massimiliano Manfren, Paola Caputo, and Gaia Costa. Paradigm shift in urban energy systems through distributed generation: Methods and models. *Applied Energy*, 88(4): 1032–1048, April 2011. ISSN 0306-2619. doi: 10.1016/j.apenergy.2010.10.018. URL http://www.sciencedirect.com/science/article/pii/S0306261910004204.

Marc A. Rosen. *Energy storage*. Nova Science Publ, New York, 2012. ISBN 978-1-61324-708-2.

François Marechal and Boris Kalitventzeff. Targeting the integration of multi-period utility systems for site scale process integration. *Applied Thermal Engineering*, 23(14):1763–1784, October 2003. ISSN 1359-4311. doi: 10.1016/S1359-4311(03)00142-X. URL http://www.sciencedirect.com/science/article/pii/S135943110300142X.

Kurt Marti. U-Wert-Berechnung und Bauteilekatalog. Technical report, Bundesamt für Energie BFE, Bern Switzerland, 2002.

François Maréchal and Boris Kalitventzeff. Process integration: Selection of the optimal utility system. *Computers & Chemical Engineering*, 22, Supplement 1:S149–S156, March 1998. ISSN 0098-1354. doi: 10.1016/S0098-1354(98)00049-0. URL http://www.sciencedirect.com/science/article/pii/S0098135498000490.

Ramanunni P Menon, Mario Paolone, and François Maréchal. Study of optimal design of polygeneration systems in optimal control strategies. *Energy*, 55:134–141, 2013.

Glenn W Milligan and Martha C Cooper. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2):159–179, 1985.

GlennW. Milligan and MarthaC. Cooper. A study of standardization of variables in cluster analysis. *Journal of Classification*, 5(2):181–204, 1988. ISSN 0176-4268. doi: 10.1007/BF01897163. URL http://dx.doi.org/10.1007/BF01897163.

Stefano Moret, Michel Bierlaire, and François Maréchal. Robust Optimization for Strategic Energy Planning. Technical report, 2014a. URL http://transp-or.epfl.ch/documents/technicalReports/MoreBierMare14.pdf.

Stefano Moret, Victor Codina Gironès, François Maréchal, and Daniel Favrat. SwissEnergyScope.ch: a Platform to Widely Spread Energy Literacy and Aid Decision-Making. In Ps Varbanov, Jj Klemes, Py Liew, Jy Yong, and P Stehlik, editors, *Pres 2014, 17th Conference On Process Integration, Modelling And Optimisation For Energy Saving And Pollution Reduction, Pts 1-3*, volume 39 of *Chemical Engineering Transactions*, pages 877–882, Milano, 2014b. Aidic Servizi Srl. ISBN 978-88-95608-30-3. doi: 10.3303/Cet1439147.

Max D. Morris. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2):161–174, May 1991. URL http://www.abe.ufl.edu/Faculty/jjones/ABE_5646/2010/Morris.1991%20SA%20paper.pdf.

Bernd Möller. A Danish Heat Atlas for Supply Strategies and Demand Side Management. In *DHC13*, volume 1, page 296, Copenhagen, Denmark, April 2012. District Energy Development Center. URL http://vbn.aau.dk/files/72417224/A_Danish_Heat_Atlas_for_Supply_Strategies_and_Demand_Side_Management.pdf.

A. Müller, L. Kranzl, P. Tuominen, E. Boelman, M. Molinari, and A. G. Entrop. Estimating exergy prices for energy carriers in heating systems: Country analyses of exergy substitution with capital expenditures. *Energy and Buildings*, 43(12):3609–3617, December 2011. ISSN 0378-7788. doi: 10.1016/j.enbuild.2011.09.034. URL http://www.sciencedirect.com/science/article/pii/S0378778811004361.

Andreja Nemet, Jiří Jaromír Klemeš, Petar Sabev Varbanov, and Zdravko Kravanja. Methodology for maximising the use of renewables with variable availability. *Energy*, 44(1):29–37, August 2012. ISSN 0360-5442. doi: 10.1016/j.energy.2011.12.036. URL http://www.sciencedirect.com/science/article/pii/S0360544211008620.

Klaus Neumann and Martin Morlock. *Operations Research*. Hanser, 2002. ISBN 978-3-446-22140-6.

F. Ochs, W. Heidemann, and H. Müller-Steinhagen. Effective thermal conductivity of moistened insulation materials as a function of temperature. *International Journal of Heat and Mass*

Transfer, 51(3–4):539–552, February 2008. ISSN 0017-9310. doi: 10.1016/j.ijheatmasstransfer.2007.05.005. URL http://www.sciencedirect.com/science/article/pii/S0017931007003614.

Fabian Ochs. *Modelling Large-scale Thermal Energy Stores*. Shaker Verlag GmbH, Germany, Aachen, February 2010. ISBN 978-3-8322-8834-1.

Fabian Ochs and Hans Müller-Steinhagen. Weiterentwicklung der Erdbecken-Wärmespeichertechnologie. Technical report, INSTITUT FÜR THERMODYNAMIK UND WÄRMETECHNIK, UNIVERSITÄT STUTTGART, Stuttgart, Germany, April 2008. URL http://www.itw.uni-stuttgart.de/dokumente/Publikationen/publikationen_08-05.pdf.

Torben Ommen, Wiebke Brix Markussen, and Brian Elmegaard. Comparison of linear, mixed integer and non-linear programming methods in energy system dispatch modelling. *Energy*, 74:109–118, September 2014. ISSN 0360-5442. doi: 10.1016/j.energy.2014.04.023. URL http://www.sciencedirect.com/science/article/pii/S0360544214004368.

J. Ortiga, J.C. Bruno, and a. Coronas. Selection of typical days for the characterisation of energy demand in cogeneration and trigeneration optimisation models for buildings. *Energy Conversion and Management*, 52(4):1934–1942, April 2011. ISSN 01968904. doi: 10.1016/j.enconman.2010.11.022. URL http://linkinghub.elsevier.com/retrieve/pii/S0196890410005315.

Jessen Page, Sébastien Dervey, and Gilbert Morand. Aggregating building energy demand simulation to support urban energy design. Ahmedabad, India, December 2014.

Soterios A. Papoulias and Ignacio E. Grossmann. A structural optimization approach in process synthesis—I: Utility systems. *Computers & Chemical Engineering*, 7(6):695–706, 1983. ISSN 0098-1354. doi: 10.1016/0098-1354(83)85022-4. URL http://www.sciencedirect.com/science/article/pii/0098135483850224.

Diane Perez. *A framework to model and simulate the disaggregated energy flows supplying buildings in urban areas*. PhD thesis, ENAC, Lausanne, 2014. URL http://infoscience.epfl.ch/record/197073/files/EPFL_TH6102.pdf.

Diane Perez, Jérôme Henri Kämpf, and Jean-Louis Scartezzini. Urban Area Energy Flow Microsimulation for Planning Support: a Calibration and Verification Study. *International Journal On Advances in Systems and Measurements*, 6(3 and 4):260–271, 2013.

Stefan Pfenninger, Adam Hawkes, and James Keirstead. Energy systems modeling for twenty-first century energy challenges. *Renewable and Sustainable Energy Reviews*, 33:74–86, May 2014. ISSN 1364-0321. doi: 10.1016/j.rser.2014.02.003. URL http://www.sciencedirect.com/science/article/pii/S1364032114000872.

Patrice Pinel, Cynthia A. Cruickshank, Ian Beausoleil-Morrison, and Adam Wills. A review of available methods for seasonal storage of solar thermal energy in residential applications.

Renewable and Sustainable Energy Reviews, 15(7):3341–3359, September 2011. ISSN 1364-0321. doi: 10.1016/j.rser.2011.04.013. URL http://www.sciencedirect.com/science/article/pii/S136403211100150X.

S.D. Pohekar and M. Ramachandran. Application of multi-criteria decision making to sustainable energy planning—A review. *Renewable and Sustainable Energy Reviews*, 8 (4):365–381, August 2004. ISSN 1364-0321. doi: 10.1016/j.rser.2003.12.007. URL http://www.sciencedirect.com/science/article/pii/S1364032104000073.

F. Porges. 8 - Heating systems. In F. Porges, editor, *{HVAC} Engineer's Handbook (Eleventh Edition)*, pages 119 – 134. Butterworth-Heinemann, Oxford, eleventh edition edition, 2001. ISBN 978-0-7506-4606-2. URL http://www.sciencedirect.com/science/article/pii/B9780750646062500081.

Fabien Poumadère, Gabriel Ruiz, and Gaëtan Cherix. CostDBCREM - A costing database for realized Swiss Projects with a detailed cost breakdown, 2015. URL www.crem.ch.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL http://www.R-project.org.

Jakob Rager, O. Epelly, and Francois Maréchal. From science to application in Geneva: best practice example from theoretical concept on exergy over a law into everyday life of a citizen. Geneva, 2011.

Jakob Rager, Cédric Dorsaz, and François Maréchal. ECOS: Integrated and practical energy planning for new districts. 2013a. URL http://infoscience.epfl.ch/record/188137?ln=en.

Jakob Rager, Luc Girardin, and François Maréchal. ECOS: Solar Potential Assessment on Roofs of Urban Areas and its Integration. 2013b. URL http://infoscience.epfl.ch/record/188136?ln=en.

Jakob Rager, Damien Rebeix, Gaëtan Cherix, François Maréchal, and Massimiliano Capezzali. MEU: An urban energy management tool for communities and multi-energy utilities. 2013c. URL http://infoscience.epfl.ch/record/188138?ln=en.

Urs Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1(3):244–256, November 1972. ISSN 0146-664X. doi: 10.1016/S0146-664X(72)80017-0. URL http://www.sciencedirect.com/science/article/pii/S0146664X72800170.

Hongbo Ren and Weijun Gao. A MILP model for integrated plan and evaluation of distributed energy systems. *Applied Energy*, 87(3):1001–1014, March 2010. ISSN 0306-2619. doi: 10.1016/j.apenergy.2009.09.023. URL http://www.sciencedirect.com/science/article/pii/S0306261909004152.

G. Reynders, T. Nuytten, and D. Saelens. Potential of structural thermal mass for demand-side management in dwellings. *Building and Environment*, 64:187–199, June 2013. ISSN 0360-1323. doi: 10.1016/j.buildenv.2013.03.010. URL http://www.sciencedirect.com/science/article/pii/S0360132313000905.

A. P. Reynolds, G. Richards, B. de la Iglesia, and V. J. Rayward-Smith. Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. *Journal of Mathematical Modelling and Algorithms*, 5(4):475–504, December 2006. ISSN 1570-1166, 1572-9214. doi: 10.1007/s10852-005-9022-1. URL http://link.springer.com/article/10.1007/s10852-005-9022-1.

LuisMiguel Rios and NikolaosV. Sahinidis. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293, 2013. ISSN 0925-5001. doi: 10.1007/s10898-012-9951-y. URL http://dx.doi.org/10.1007/s10898-012-9951-y.

Stephen A. Roosa. *Sustainable development handbook*. Fairmont Press, 2010.

Marc A Rosen and Ibrahim Dincer. Exergy as the confluence of energy, environment and sustainable development. *Exergy, An International Journal*, 1(1):3–13, 2001. ISSN 1164-0235. doi: 10.1016/S1164-0235(01)00004-8. URL http://www.sciencedirect.com/science/article/pii/S1164023501000048.

Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, November 1987. ISSN 0377-0427. doi: 10.1016/0377-0427(87)90125-7. URL http://www.sciencedirect.com/science/article/pii/0377042787901257.

A. Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis: The Primer*. Wiley-Interscience, Chichester, England ; Hoboken, NJ, 1 edition edition, February 2008. ISBN 978-0-470-05997-5.

Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. The Screening Exercise. In *Sensitivity Analysis in Practice*, pages 91–108. John Wiley & Sons, Ltd, 2002. ISBN 978-0-470-87095-2. URL http://onlinelibrary.wiley.com/doi/10.1002/0470870958.ch4/summary.

Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, February 2010. ISSN 0010-4655. doi: 10.1016/j.cpc.2009.09.018. URL http://www.sciencedirect.com/science/article/pii/S0010465509003087.

Arno Schlueter and Frank Thesseling. Building information model based energy/exergy performance assessment in early design stages. *Automation in Construction*, 18(2):153–

163, March 2009. ISSN 0926-5805. doi: 10.1016/j.autcon.2008.07.003. URL http://www.sciencedirect.com/science/article/pii/S0926580508001064.

T. Schmidt, D. Mangold, and H. Müller-Steinhagen. Central solar heating plants with seasonal storage in Germany. *Solar Energy*, 76(1–3):165–174, January 2004. ISSN 0038-092X. doi: 10.1016/j.solener.2003.07.025. URL http://www.sciencedirect.com/science/article/pii/S0038092X03002937.

Thomas Schütz, Rita Streblow, and Dirk Müller. A comparison of thermal energy storage models for building energy system optimization. *Energy and Buildings*, 93:23–31, April 2015. ISSN 0378-7788. doi: 10.1016/j.enbuild.2015.02.031. URL http://www.sciencedirect.com/science/article/pii/S037877881500136X.

Enrico Sciubba. Beyond thermoeconomics? The concept of extended exergy accounting and its application to the analysis and design of thermal systems. *Exergy, an international journal*, 1(2):68–84, 2001.

Suresh Sethi and Gerhard Sorger. A theory of rolling horizon decision making. *Annals of Operations Research*, 29(1):387–415, December 1991. ISSN 0254-5330, 1572-9338. doi: 10.1007/BF02283607. URL http://link.springer.com/article/10.1007/BF02283607.

Atul Sharma, V. V. Tyagi, C. R. Chen, and D. Buddhi. Review on thermal energy storage with phase change materials and applications. *Renewable and Sustainable Energy Reviews*, 13(2):318–345, February 2009. ISSN 1364-0321. doi: 10.1016/j.rser.2007.10.005. URL http://www.sciencedirect.com/science/article/pii/S1364032107001402.

Bruce Sibbitt, Doug McClenahan, Reda Djebbar, Jeff Thornton, Bill Wong, Jarrett Carriere, and John Kokko. The Performance of a High Solar Fraction Seasonal Storage District Heating System – Five Years of Operation. *Energy Procedia*, 30(0):856–865, 2012. ISSN 1876-6102. doi: 10.1016/j.egypro.2012.11.097. URL http://www.sciencedirect.com/science/article/pii/S187661021201613X.

Gürkan Sin, Krist V. Gernaey, and Anna Eliasson Lantz. Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnology Progress*, 25(4):1043–1053, July 2009. ISSN 1520-6033. doi: 10.1002/btpr.166. URL http://onlinelibrary.wiley.com/doi/10.1002/btpr.166/abstract.

Allen L Soyster. Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.

Peter Steiger, editor. *PLENAR - Wärmeverbund Olten: (Vorprojekt im Auftrag des Nationalen Energieforschungsfonds NEFF)*. PLENAR-Vereinigung, 1979.

Douglas Steinley. Standardizing Variables in K-means Clustering. In David Banks, FrederickR. McMorris, Phipps Arabie, and Wolfgang Gaul, editors, *Classification, Clustering, and Data Mining Applications*, Studies in Classification, Data Analysis, and Knowledge Organisation,

pages 53–60. Springer Berlin Heidelberg, 2004. ISBN 978-3-540-22014-5. URL http://dx.doi.org/10.1007/978-3-642-17103-1_6.

Lukas G. Swan and V. Ismet Ugursal. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8):1819–1835, October 2009. ISSN 1364-0321. doi: 10.1016/j.rser.2008.09.033. URL http://www.sciencedirect.com/science/article/pii/S1364032108001949. 00453 bibtex: Swan2009.

Swiss Society of Engineers and Architects. Standard-Nutzungsbedingungen für die Energie- und Gebäudetechnik 2024, 2006. URL http://www.webnorm.ch/null/null/sia%202024/d/F/Product/.

Swiss Society of Engineers and Architects. *Certificat énergétique des bâtiments selon SN EN 15217 et SN EN 15603 Cahier technique 2031*. SIA 2031. SIA, Zürich, January 2009a.

Swiss Society of Engineers and Architects. *L'énergie thermique dans le bâtiment*. SIA 380/1. SIA, Zürich, January 2009b.

Jan Szargut. Analysis of cumulative exergy consumption. *International Journal of Energy Research*, 11(4):541–547, October 1987. ISSN 1099-114X. doi: 10.1002/er.4440110410. URL http://onlinelibrary.wiley.com/doi/10.1002/er.4440110410/abstract.

Jan Szargut. *Exergy Method: Technical and Ecological Applications*. WIT Press, 2005. ISBN 978-1-85312-753-3.

Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to data mining*. Pearson Addison Wesley, Boston, 2005. ISBN 0-321-32136-7 978-0-321-32136-7. URL http://www-users.cs.umn.edu/~kumar/dmbook/index.php.

Hideki Tanaka, Takashi Tomita, and Masaya Okumiya. Feasibility study of a district energy system with seasonal water thermal storage. *Solar Energy*, 69(6):535–547, 2000. ISSN 0038-092X. doi: 10.1016/S0038-092X(00)00122-5. URL http://www.sciencedirect.com/science/article/pii/S0038092X00001225.

Parfait Tatsidjodoung, Nolwenn Le Pierrès, and Lingai Luo. A review of potential materials for thermal energy storage in building applications. *Renewable and Sustainable Energy Reviews*, 18:327–349, February 2013. ISSN 1364-0321. doi: 10.1016/j.rser.2012.10.025. URL http://www.sciencedirect.com/science/article/pii/S1364032112005679.

Robert Tibshirani, Guenther Walther, and Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001. ISSN 1467-9868. doi: 10.1111/1467-9868.00293. URL http://dx.doi.org/10.1111/1467-9868.00293.

Laurence Tock. *Thermo-environomic optimisation of fuel decarbonisation alternative processes for hydrogen and power production*. PhD thesis, STI, Lausanne, 2013.

Herena Torío, Adriana Angelotti, and Dietrich Schmidt. Exergy analysis of renewable energy-based climatisation systems for buildings: A critical view. *Energy and Buildings*, 41(3): 248–271, March 2009. ISSN 0378-7788. doi: 10.1016/j.enbuild.2008.10.006. URL http://www.sciencedirect.com/science/article/pii/S0378778808002211. 00000.

Petar Sabev Varbanov and Jiří Jaromír Klemeš. Integration and management of renewables into Total Sites with variable supply and demand. *Computers & Chemical Engineering*, 35 (9):1815–1826, September 2011. ISSN 0098-1354. doi: 10.1016/j.compchemeng.2011.02.009. URL http://www.sciencedirect.com/science/article/pii/S0098135411000718.

Hrishikesh D. Vinod. Integer Programming and the Theory of Grouping. *Journal of the American Statistical Association*, 64(326):506–519, 1969. ISSN 0162-1459. doi: 10.1080/01621459.1969.10500990. URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1969.10500990.

Philip Voll. *Automated Optimization-Based Synthesis of Distributed Energy Supply Systems*. PhD thesis, RWTH Aachen, Aachen, 2013.

Göran Wall and Mei Gong. On exergy and sustainable development—Part 1: Conditions and concepts. *Exergy, An International Journal*, 1(3):128–145, 2001. ISSN 1164-0235. doi: 10.1016/S1164-0235(01)00020-6. URL http://www.sciencedirect.com/science/article/pii/S1164023501000206. 00216.

Peter Warren. A review of demand-side management policy in the UK. *Renewable and Sustainable Energy Reviews*, 29:941–951, January 2014. ISSN 1364-0321. doi: 10.1016/j.rser.2013.09.009. URL http://www.sciencedirect.com/science/article/pii/S1364032113006680.

Céline Weber. *Multi-objective design and optimization of district energy systems including polygeneration energy conversion technologies*. PhD thesis, EPFL, Lausanne, January 2008. URL http://infoscience.epfl.ch/record/114786/files/EPFL_TH4018.pdf.

H. Paul Williams. *Model building in mathematical programming*. Wiley, Chichester, 2013. ISBN 978-1-118-44333-0.

World Bank. Urban population, 2015. URL http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS.

World Commission On Environment and Development. *Our Common Future*. Oxford University Press, USA, May 1987. ISBN 0-19-282080-X.

Rui Xu and D.C. Wunsch. Clustering Algorithms in Biomedical Research: A Review. *Biomedical Engineering, IEEE Reviews in*, 3:120–154, 2010. ISSN 1937-3333. doi: 10.1109/RBME.2010.2083647.

Xue Xue, Shengwei Wang, Yongjun Sun, and Fu Xiao. An interactive building power demand management strategy for facilitating smart grid optimization. *Applied Energy*, 116:297–310, 2014. ISSN 0306-2619. doi: 10.1016/j.apenergy.2013.11.064. URL http://www.sciencedirect.com/science/article/pii/S0306261913009719.

Ryohei Yokoyama, Yasushi Hasegawa, and Koichi Ito. A MILP decomposition approach to large scale optimization in structural design of energy supply systems. *Energy Conversion and Management*, 43(6):771–790, April 2002. ISSN 0196-8904. doi: 10.1016/S0196-8904(01)00075-9. URL http://www.sciencedirect.com/science/article/pii/S0196890401000759.

Fengqi You, Jose M. Pinto, Elisabet Capón, Ignacio E. Grossmann, Nikhil Arora, and Larry Megan. Optimal Distribution-Inventory Planning of Industrial Gases. I. Fast Computational Strategies for Large-Scale Problems. *Industrial & Engineering Chemistry Research*, 50(5):2910–2927, March 2011. ISSN 0888-5885. doi: 10.1021/ie1017578. URL http://dx.doi.org/10.1021/ie1017578.

Jakob Rager

Curriculum Vitae

Education

2010–2015 **PhD**, *Ecole Polytechnique Fédérale de Lausanne*, Lausanne, Switzerland, *PhD in Energy*.
Urban Energy System Design

2003–2009 **Diplom - Wirtschaftsingenieur**, *Universität Karlsruhe (TH)*, Karlsruhe, Germany, *Diplom*.
Master of Science Equivalent

2007–2008 **Exchange**, *Ecole Polytechnique Fédérale de Lausanne and Université de Lausanne*, Lausanne, Switzerland.
Exchange in the Framework of the European Erasmus Programme

Vocational Experience

2009–2010 **Research and Teaching Assistant**, *Ecole Polytechnique Fédérale de Lausanne*, Lausanne, Switzerland.
Working in research projects and teaching small groups of students

2008 **Intern**, *Office cantonal de l'énergie Genève*, Geneva, Switzerland.
Working on a data verification system for urban energy studies

2005–2006 **Intern**, *United States Senate Environment and Public Works Committee*, Washington, D.C., USA.
Preparing hearings on environmental impacts of hurricane Katrina, future pipeline projects and nuclear power and working as a research assistant

Languages

German	C2	*native language*
English	C2	*high level in writing and speaking through study and work experience abroad*
French	C1	*high level in speaking, high technical writing skills*

Computer skills

Languages	Python, Matlab, R, C#, SQL
OS	Linux, Windows, Mac OS
Office	Latex, LibreOffice, Microsoft Office, Adobe InDesign

Interests

Sports	Rowing, Ski touring: Having a great experience in a team.
Traveling	Discovering new places and new cultures

Chemin des Clochetons 6 – 1004 Lausanne – Switzerland
jakob.rager@gmail.com