Understanding and accelerating the model-based design of multi-source energy systems

by

François Lédée
M.Eng., IMT Mines Albi-Carmaux, 2019

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Civil Engineering

We acknowledge and respect the Lək̓ʷəŋən (Songhees and Esquimalt) Peoples on whose territory the university stands, and the Lək̓ʷəŋən and W̱SÁNEĆ Peoples whose historical relationships with the land continue to this day.

Understanding and accelerating the model-based design of multi-source energy systems

by

François Lédée
M.Eng., IMT Mines Albi-Carmaux, 2019

Supervisory Committee

Dr. Ralph Evins, Supervisor
(Department of Civil Engineering)

Dr. Curran Crawford, Co-Supervisor
(Department of Mechanical Engineering)

Dr. Madeleine McPherson, Departmental Member
(Department of Civil Engineering)

Abstract

The transition towards sustainable energy system requires innovative tools and methodologies to efficiently address the design of resilient and cost-effective multi-energy systems (MES). The energy hub concept emerges as a powerful approach, capable to address optimal design and operation of energy systems while best exploiting potential synergies between technologies. Its scalability, limited by significant computational demands, is necessary to help capture the multi-scale nature of energy systems and support a harmonized transition across all levels. This thesis investigates two key strategies to leverage this limitation. The first is the use of representative days to reduce the temporal complexity of the core application. The second is the development of surrogate models to rapidly estimate optimal designs. These strategies are examined across diverse contexts and system design problems through four independent but complementary studies. Each study led to a publication in international scientific journal or conference proceedings.

The first study evaluates the robustness of selection methods for the use of representative days. The study highlights their sensitivity to context and the need for validation across multiple case studies. The second study examines the impact of shorter time horizons on the MES design. It emphasizes the strong influence on the system design, particularly on the design of storage. It also offers new insights into the design decision mechanisms. The third study compares district heating and cooling technologies. It demonstrates the superiority of the 5th generation network through diverse scenarios, emphasizing sensitivity to the topology and spatial demand distribution. The fourth study develops a novel surrogate modeling framework leveraging the use of machine-learning to address system design problems. The robustness of the framework is validated across diverse scenarios.

Overall, this work emphasizes the importance of context-specific validations and sensitivity analyses. These aim to improve the understanding of design decision processes of energy hub applications, necessary to develop robust and effective complexity reduction methodologies. These contributions provide a foundation to improve the scalability and accessibility of MES design methods and, eventually, support the energy transition.

List of publications

The research conducted during the development of this thesis leads to publications in international scientific journals and conference proceedings. This thesis manuscript is articulated around the four most impactful articles. Two are published, one in a scientific journal, the other in the proceedings of the eSim conferences. The two remaining articles are currently under review. Two additional articles published during the development of the thesis are not covered in the present manuscript.

The following list presents the articles in order of appearance in the thesis.

- **[P1]** Lédée.F., Crawford,C. Evins,R., Error incurred in multi-energy system design through selection of representative days. [Submitted to *Energy*]
- **[P2]** Lédée.F., Crawford,C. Evins,R., The influence of time horizon on energy hub sizing. *Proceedings of eSim22* (Vol.12). IBPSA.
- **[P3]** Lédée,F., Evins,R., 2024, A comparison of 4th and 5th generation thermal networks with energy hub. *Energy 311* (2024), doi: 10.1016/j.energy.2024.133336
- **[P4]** Lédée.F., Crawford,C. Evins,R., Improved surrogate modeling for multi-energy system design: model architecture, sampling and scaling choices [Submitted to *Applied Energy*]

For publications [P1] to [P4], the contributions are as follows: Lédée F.: conceptualization, methodology, data collection, experimentation, analysis, writing, revision; Crawford C.: writing, revision, supervision; Evins R.: writing, revision, supervision, funding.

Secondary publications

- Lédée.F., Faure,G., Crawford,C. Evins,R., Optimal energy system sizing with independent load and weather time series. *Proceedings of Building Simulation 2021* (Vol.17). IBPSA.
- Lédée,F., Padey.P, Goulouti,K, Lasvaux,S., Beloin-Saint-Pierre,D., 2023, EcoDynElec: Open Python package to create historical profiles of environmental impacts from regional electricity mixes. *SoftwareX* (Vol.23), doi: 10.1016/j.softx.2023.101485.

Nomenclature

Acronyms

4G/5G $4^{th}/5^{th}$ generation

AChill Absorption Chiller

ASHP Air Source Heat Pump

CHP Combined Heat and Power

COP Coefficient of Performance

DTN District thermal network

EB Electric Boiler

EChill Electric Chiller

EHub Energy Hub

GB Gas Boiler

HP Heat Pump

HWT Hot Water Tank

HX Heat Exchanger

MES Multi-energy system

MILP Mixed Integer Linear Programming

PSO Particle Swarm Optimization

PTES Pit Thermal Energy Storage

PV Photovoltaic

RnE Renewable Energy

ST Solar Thermal

WH Waste Heat

WiP Water in Pipes

Constants

α^R Surface per capacity unit

β Storage limit factor

Δp_x Max pressure gradient

ΔT Temperature difference

ΔT^{flex} Max. T variation of network water

ϵ width of pipe's house

η efficiency

η_{lin} Linearized pumping efficiency

κ Static storing efficiency

λ thermal conductivity

ρ Water density

A_{max} Max Roof area

C Price

c_p Water thermal capacity

D Diameter of pipes

f_D Darcy-Weisbach friction factor

G Carbon intensity

H conversion efficiency matrix

I^{sun} Sun irradiation

L Length of link

M Big M parameter

NPV Net Present Value

T_s Temperature of soil

T_w Temperature of water

X_{CO_2} Carbon emission limit

Ensembles

$\sigma \in \mathcal{S}$ Set of streams

$\tau \in \mathcal{T}$ Set of technologies or links

Ξ^{Net} Set of techs if connected

Ξ^{Node} Set of techs if standalone

$f \in \mathcal{F}$ Set of fuels

$l \in \mathcal{L}$ Set of links

$n \in \mathcal{N}$ Set of nodes

$t \in \mathbb{T}$ Set of time steps

Subscripts/Superscripts

ch charging

cv a converter

dch discharging

fix Relative to fixed part

H / C hot / cold

in entering

lin Relative to linear part

$load$ power demand / load

out exiting

ref main / of reference

s a storage

Variables

δp Pseudo-pressure level

δ binary for installation status

$\dot{Q}$ Power

$\dot{Q}^{\mathcal{L}}$ Power between node and network

$\dot{Q}^{\mathcal{N},in/out}$ Imported/Exported power

$\dot{q}^{el}$ circulation losses

$\dot{q}^{th}$ thermal losses

γ State on/off

J^{a}_{CAPEX} Annualized investment cost

J^{a}_{tot} Total annualized cost

J_{CO_2} Cost of carbon taxes

J_{fuel} Operating cost

$J_{O\&M}$ Cost for maintenance

P_{nom} Capacity of device

SOC State of Charge

SOC^{d} Daily resolution SOC (intra-day)

SOC^{h} Variation of SOC during a typical day (intra-day)

Contents

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Ralph Evins, for his supervision over my years of study at the University of Victoria. His guidance and advices have been of the most precious in the realization of this work.

My gratitude also goes to my co-supervisor, Dr Curran Crawford, for his constant availability and complete devotion to provide the best of help and most effective advises.

This whole work would not have been possible without the infinite support of my parents. To both of you, thank you for helping me go through the moments of deepest doubts, and for sharing your research experience.

And of course, thank you to all those with whom I had the honor to interact during this adventure: my dear lab mates.

François Lédée

Victoria BC, November 2024

Chapter 1

Introduction

1.1 Context

A strong correlation exists between economic growth, financial development, energy consumption and CO2 emissions [1]. The Covid19 pandemic led to a global reduction in carbon emissions of 8% in 2020 [2]. Since, the level of global emissions ramped back to its pre-covid trend within a single year [3]. Keeping the global temperature increase below two degrees would require an additional fall of 6% per year to 2050, while projections rather predict an average annual rate of 0.7% [2]. This reduction could be achieved by massive electrification along with a strong decline of fossil fuel use in electric generation and an important penetration of renewable energy and development of energy storage systems [2]. The energy production market is driven by financial considerations, and prices related to renewable energy generation systems [3, 4] and storage [5] are decreasing, rending their use more affordable at all geographical scales. The intermittent nature of wind turbines (WT) and photovoltaic (PV) panel production increases the risk of system failures and this production is not necessarily correlated to our needs. To facilitate the penetration of renewable electricity use, the development of storage technologies [2], the careful design of integrated energy systems [6, 7], as well as smart planning approaches incorporating the challenges relative to the management of energy networks [8–11] and of the energy demand [12, 13] at multiple scales [7] is of essence.

Energy systems are complex [14]. From the energy network at national level [15] to the local management and generation at community or building level [16], systems need to be seen as a whole. To phase out of the reliance on fossil fuels, massively integrate renewable energy sources of various nature, and decarbonize

the energy sector [17], adequately advanced modeling is required. On diverse energy systems, each actor and technology represents specific challenges and opportunities. Their interaction expands the fields of possibilities via new potential synergies [16,18,19]. It also increases the difficulties to carefully plan their smooth interaction [9]. This intensifies the challenges relative to the elaboration of modeling capable to accurately represent these complex interactions [16].

One computational framework representing this complexity and the interaction between all energy needs and renewable energy systems is called the Energy Hub (EH) [20]. It consists of an optimization framework acting as an interface between various energy infrastructures and loads, in which energy carriers can be converted, conditioned and stored [20]. Related concepts can be found under various names, such as "Multi-Energy Systems" (MES) [16, 21], "Distributed Energy System" (DES) [22,23], "Hybrid Renewable Energy Systems" [24], "Multi-Carrier Energy System" [25], "Integrated Energy Systems" [26] among many others [27]. From the modeling point of view, it has the advantage of being very flexible, and therefore applicable at multiple scales (from a building, neighborhood, city to a province or even a country) [27] and across time horizons [28,29]. This optimization framework couples both the ability to realize optimal scheduling [30] and sizing [31] of energy infrastructures. Optimization goes along with an objective function, this being the cost [19,32], the environmental impacts [33] including carbon equivalent emissions, the penetration of renewable energy in the designed system [27, 34] or level of autonomy [35,36] for the most common. But finding the extrema of a cost function is not the main purpose, and 3 uses of such a framework are predominant: (1) Energy System Design (buying or not buying), (2) Energy System Sizing (what capacity) and (3) Operation Planning (what is used when and for what amount) [16].

1.2 Scope, Motivation and Objectives

The electric network can be seen as a stratified structure, with the different levels corresponding to voltage levels. In a centralized system, the energy is fed in at the transmission level (high voltage), and travels through the network to reach the customer at the distribution level (low voltage). From a modeler's point of view, common practice consists in representing the network infrastructures at one voltage level (e.g. sub-transmission at a regional scale), aggregating the demand from lowest levels (e.g. distribution level at the city scale) [15, 37–39] and considering higher voltage levels (e.g. transmission level at the national scale) as an infinite energy source [16, 27, 37].

The penetration of renewable electricity has two consequences on the electrical network. First, it modifies the demand; a building equipped with solar panels will require less electricity from the grid during the day and this modification will be irregular. Second, an excess of renewable energy production is directly fed to the distribution network, and in an irregular way; this newly available energy must be considered to ensure the network stability and the balance between generation and demand. Expansion models at the transmission level consider both the generation (renewable or not) and the demand at the transmission level, potentially creating a bottleneck effect [15]. A recent report [15] demonstrated the importance of considering details and joining the highest voltage level (Transmission) to the lowest voltage level (Distribution) in the modeling. The key result of this study is the cost of transmission expansion that can be avoided if decentralized generation and balancing is considered in the expansion models (between $88 billion and $473 billion by 2050 at the USA scale).

The intermittent nature of renewable energy generation and variations in the demand both mainly occur at distribution level. As described in [15], the consideration of storage and networking in addition to renewable generation at the distribution level helps mitigate the need for transmission expansion. The optimal solution of expansion and dispatch planning problems is likely to be a hybrid between measures at the transmission and distribution levels. Therefore, to suggest better models representing larger scopes, it seems necessary to start emphasizing on model where the demand occurs and is shifted, i.e. at distribution, community or even building level [40,41].

Moreover, the nature of the energy demand shows a larger diversity in nature at lower levels, adding heating and cooling at different temperature levels [42,43] to the electricity demand. Recent efforts in building retrofitting [40,44] and electrification of the thermal demand [5,9] adding complexity, diversity and additional opportunities for synergies. However, their integration into larger models necessarily increases modeling difficulty, and consequently an increase in computation burden [45]. For this integration, a deeper understanding of decision mechanisms from energy system design models at local levels are needed to enhance their computational efficiency while retaining their accuracy.

Recent developments in sustainable energy systems emphasize the need for efficient optimization frameworks, usable throughout a multitude of contexts. Current literature reveal gaps in the characterization of robustness of suggested approaches, and in the search for explanations of design decisions within suggested frameworks. This thesis contributes to the global modeling efforts by suggesting new analyzes through various energy contexts and by exploring new strategies to

reduce the computational burden of energy system design via optimization.

The overall goal of this thesis is to evaluate and develop methodologies for the optimal design and operation of energy systems. The main aspects of interest are the trade-off between computational efficiency and design accuracy across system complexity. Through the exploration of use of problem reduction approaches and machine learning techniques, this work aims eventually to enhance the scalability of energy system design optimization. Particularly, the integration of diverse energy sources and the exploitation of their synergies within multi-energy systems (MES) to serve the decarbonization of the building energy sector at local scale is of prime interest.

1.3 Key contributions

The main research question driving this dissertation is: *How can methods for time-series aggregation and surrogate modeling be developed and evaluated to improve the accuracy, efficiency, and reliability of multi-energy system design optimization for buildings across diverse contexts?*

The thesis is divided into four studies, each addressing different aspects of the main research question. The state-of-the-art is presented in each chapter individually. Sections 1.3.1 to 1.3.4 summarize the most central literature and introduce the objectives and contributions for each chapter. Together, the four chapters propose the evaluations of reliability of model results, the understanding of decision processes and the development of time reduction approaches: via the usage of representative days as a complexity reduction approach, and via the leverage of machine-learning algorithms for the development of surrogate models.

1.3.1 Objective 1: Understanding the impact of context for the evaluation of accuracy and efficiency

*This objective is mainly addressed in Chapter 2, reflecting the publication **[P1]**.*

The use of representative design days is among the most popular time series aggregation approaches to reduce the computation burden inherent to the design of MES via optimization [46]. It reduces the size of the design problem while, in theory, exploiting a reduced amount of information that still retains key elements for design decisions. Methods to best address the crucial selection of representative

days are abundant in recent literature [46]. These include clustering-based approaches using various ways to assess similarity between days [47,48], approaches trying to reproduce the original temporal data with a reduced set of days [49,50], and the specific selection of days exhibiting extreme events [47,51]. Most suggested methods are compared with some others on specific application cases. However, there is a significant gap in understanding the reliability of these methods across varying contexts and throughout multiple cases. While prior research has evaluated and compared methods on isolated case studies, an assessment of the method capabilities for multiple systems, load types and techno-economic conditions is missing. The second chapter of this thesis aims at evaluating the ability to extrapolate observations on method reliability across contexts.

1.3.2 Objective 2: Providing a deeper understanding of decision processes

*This objective is mainly addressed in Chapter 3, reflecting the publication **[P2]**.*

The use of representative days for the design of MES induces errors. Even state-of-the-art methods to reformulate the design problem and extract the most relevant information, a priori, from original data fail at systematically replicate the expected behavior of the original and burdensome design problem without error [46]. A major difficulty arises from the selection of representatives, performed without considering the characteristics of the design problem itself. Alternatives are suggested to address this shortcoming and reduce error, by considering iterative frameworks refining the selection [52,53], or by incorporating characteristics of the system to design [51], e.g. storage. Despite these advances, the decision mechanisms within the MES design problem remains underexplored, and no attempt was made to identify key periods, time steps or elements driving design decisions. The third chapter of this thesis addresses this gap by evaluating how varying the time series horizon affects both the system sizing and the operational control strategies.

1.3.3 Objective 3: Assessing the importance of modeling details

*This objective is mainly addressed in Chapter 4, reflecting the publication **[P3]**.*

With the need to decarbonize heating and cooling demands in buildings, state-

of-the-art district thermal networks were identified as key technologies for future sustainable energy strategies [16]. The 4th generation network, centralized system using low temperature distribution, enables the valorization of waste heat and integration of renewable sources [43, 54, 55]. The 5th generation network, decentralized system operating with ultra-low temperatures (close to ambient), enhances the system flexibility by simultaneously addressing heating and cooling, and further support electrification and the integration of renewable sources [42, 43, 56]. The few studies addressing the comparison between the two technologies, exhibit shortcomings. Most focus on the network and neglect potential synergies among other technologies and sources part of the surrounding energy system [9]. Others fail to fully capture the decentralized features and thermal coupling unique to 5th generation networks [56]. The fourth chapter of the thesis addresses these gaps by developing a model to systematically compare the economic and environmental performances of the two network. Applying principles described in previous chapters, the comparison is repeated across diverse climatic and techno-economic contexts. Also, an investigation of the influence of key modeling parameters on the comparison is conducted.

1.3.4 Objective 4: Leveraging the use of machine-learning

*This objective is mainly addressed in Chapter 5, reflecting the publication **[P4]**.*

As the scale and detail level of MES design problems grow, the computational demand increases exponentially [45]. While leveraging the computational power of machine-learning (ML) algorithms directly for the design tasks is frequently mention as future paths to explore [45, 57], related state-of-the-art usage of ML focuses on surrogate modeling assisted optimization [58]. This method only uses ML to predict an objective function, to support the original design problem in making design decisions. A very restricted set of works attempted to study surrogates directly for the design of MES [44, 59, 60], however large errors were obtained. Based on observations on the decision mechanism of MES design problems, the fifth chapter of this thesis establishes a methodology to develop accurate surrogate models. Applying principles described in previous chapters, the methodology is tested across a large variety of cases.

1.4 Outline

The thesis is structured in six chapters. Due to the independent nature of each, their respective literature review and description of methodology are left within each chapter, creating occasional redundancies. This introduction constitutes the Chapter 1.

Chapter 2 evaluates the ability to extrapolate observations on the reliability of aggregation method to select representative days for energy system design applications. A systematic comparison of selection method on the induced quality of design is conducted exhaustively for the design of three systems. The experiment is repeated for different load types, and across different geographical and techno-economic scopes.

Chapter 3 evaluates how varying the time series horizon affects both the system sizing and operational control strategy of MES design problems. Particularly, it investigates whether a reduction of the time series horizon can still capture the essential dynamics of system performance, in order to connect exhibited properties of the time series with specific design elements.

Chapter 4 compares the economic and environmental performances of energy systems articulated around a 4th and a 5th generation thermal network. The model suggested for the comparison includes decentralized features and the thermal coupling unique to 5th generation networks, while relying on the Energy Hub concept to include synergies between the network, other technologies and energy sources in the system.

Chapter 5 establishes a new methodology to development of accurate surrogate models for the rapid and precise estimate of energy system design. The suggested method relies on diverse machine-learning and data preprocessing steps, justified by the joint understanding of key properties of the MES design problem and of the machine-learning model capabilities.

Finally, Chapter 6 summarizes the main findings of the thesis, and discusses them in perspectives with the objectives of the work. Possible extensions and future research opportunities are discussed to close the thesis.

Chapter 2

Error incurred in multi-energy system design through selection of representative days

Abstract

The design of multi-source energy systems is key to addressing the decarbonization of the energy sector. The use of representative days to shrink the size of design problems is common to reduce the computation burden. While numerous methods exist to select these representatives days, they all show limitations. First, by focusing mainly on the cost, these methods tend to overlook the sizing of assets, inducing large errors. Second, their validity when reused for different contexts is unknown. This study evaluates methods to select representative days across different system complexities, load types, geographical locations and techno-economic contexts.

The findings reveal a clear dependence between system performance and the context in which performance is measured. For an intended usage of the model within a specific scope (e.g. geographical or techno-economic), an extrapolation to broader or narrower scopes is likely acceptable. However, the extrapolation of conclusions to a different system types or load type is not recommended. Moreover, the study reveals that the quality of prediction of the objective value (e.g. cost) does not correlate with the quality of prediction of the overall system design. Finally, while no method consistently outperforms all others across all contexts, a few selection methods show more stable performance overall.

Ultimately, this research highlights the context-dependent nature of aggregation methods. It shows the need for a systematic framework to characterize induced errors across the full range of expected scenarios. Such a framework would help identify the key factors influencing MES optimization and guide the development of more reliable methods for representative day selection in future work.

*This chapter is related to the publication **[P1]***

2.1 Introduction

The optimization of complex Multi-energy system (MES) models for the sizing of infrastructures is a computationally demanding task, usually performed using yearly time series with hourly time steps. Time series aggregation (TSA) methods can be used to reduce model complexity and decrease computation time [46]. The selection of TSA approaches to retain the key properties of the time series that are important to the MES design problem is a major challenge. Storage, and particularly seasonal storage, is essential to model systems with high penetrations of fluctuating renewable energy, but significantly increases the difficulty to correctly select representative days and use TSA [61, 62].

The whole time horizon is required to simulate storage, therefore most existing studies only consider storage within a representative period [63]. This may result in an underestimation of the storage capacity for both short-term and long-term storage [64]. Renaldi and Friedrich [65] solve this limitation by proposing an optimization formulation based on multiple time grids. Further improvements were made by coupling with the typical periods method [61, 62]. Gabrielli et al. [62] propose two computational formulations, modeling storage and converters on distinct time grids. Kotzur et al. [61] suggest an adaptation to link representative days, and use two time grids to model the storage level.

Clustering techniques are used to group days and select representative periods, the k-medoids clustering [10, 46, 53, 56, 61, 63, 66, 67] being the most popular technique. The selection of representative periods within clusters has more influence on the optimization result than the clustering algorithm [28, 47, 48] or the distance metric used for the clustering [28, 47, 68]. Kotzur et al. [47] recommend the use of cluster medoids as representative days, as these are real data points.

Different approaches exist to select days to represent the full year of hourly data. A common approach uses clustering to group days by similarity, then select a representative in each cluster. The weights associated to each representative

days correspond to the size of the clusters. The selection of representative periods within clusters has more influence on the optimization result than the clustering algorithm [28,47,48] or the distance metric used for the clustering [28,47,68]. A few days exhibiting yearly extremes may be manually selected as cluster representatives [47]. For the rest, as the selection of a real day over an averaged day is recommended [47], the k-medoids clustering is the most popular technique to group days and select representatives simultaneously [10,46,53,56,61,63,66,67]. Evidence of some instability [47,68,69] argues in favor of alternative algorithms, for example the deterministic algorithm of agglomerative clustering [47]. However, no exploration was considered for cases including seasonal storage.

Another alternative approach numerically selects representative days and associated weights specifically to approximate the load duration curve of demands and renewable sources availability [50,69,70]. This kind of approach was proven superior to clustering-based approaches on numerous individual cases, including seasonal storage [36,69,71,72] or not [50,70,73]. Opinion based on single-cases analyses is divided, some claiming evidence of correlation between the a-priori error of the TSA on the a-posteriori error on the results [74], others claiming a lack of evidence [46,68,71]. An emerging practice suggests to further aggregate adjacent hours within representative days [36,53,69] to further reduce the problem size.

Communication about values and uncertainties embedded in the modeling [75] are of prime importance. However, the error induced by the TSA on the design results is lacking characterization in current literature. Kotzur et al. [47] investigates the selection of representative days for three systems excluding seasonal storage. Teichgräber et al. [68] compares selection methods on the objective function for two systems without seasonal storage, over two climate zones. Fleschutz et al. [71] characterizes the quality of aggregation for two systems with seasonal storage, for multiple decarbonization objectives. Hoffmann et al. [69] evaluates the selection of representatives on the objective function for two systems with seasonal storage. Risch et al. [36] studies possible energy autonomy of 11,000 German municipalities, with a global validation of the selection of representatives on the objective for the unique suggested system layout and selection method. To allow reproducibility and extrapolation to other case studies, a broader investigation of the influence of the selection of representative days across multiple modeling interests is required.

This study explores the consistency of methods selecting representative days for early stage MES design. Particularly, the four questions below are addressed:

How reliable are the design results with a selection method carefully chosen,

...

- ... if the potential MES design is slightly modified?
- ... if the nature of the load is changed significantly?
- ... if the load characteristics are slightly modified?
- ... if the techno-economic parameters of the problem fluctuate?

Section 2.2 details the implementations of the MES design problem, of the selections of representative days, and introduces the procedure to evaluate the reliability of design results. Section 2.3 addresses each question separately, which are further discussed in section 2.4.

Table 2.1: Contribution and comparison with studies including error characterization

	Kotzur et al. [47]	*Teichgraeber et al. [68]*	*Hoffmann et al. [69]*	*Fleschutz et al. [71]*	*Risch et al. [36]*	*This paper*
number of Σ	3	2	2	2	1	3
Compare Σ		✓				✓
Selection	✓	✓	✓			✓
Tech-eco context				✓		✓
Geogr. context	$1/\Sigma$	2	$1/\Sigma$		11,000	20
Load type	$1/\Sigma$		$1/\Sigma$		1/Geog.	3
Seasonal storage			✓	✓	✓	✓
Design parameter prediction	✓			✓		✓

2.2 Methodology

Figure 2.1 shows the flowchart of the methodology for this study. The different elements are detailed below. The building data used for all locations and load types are from the same dataset of archetypes [76].

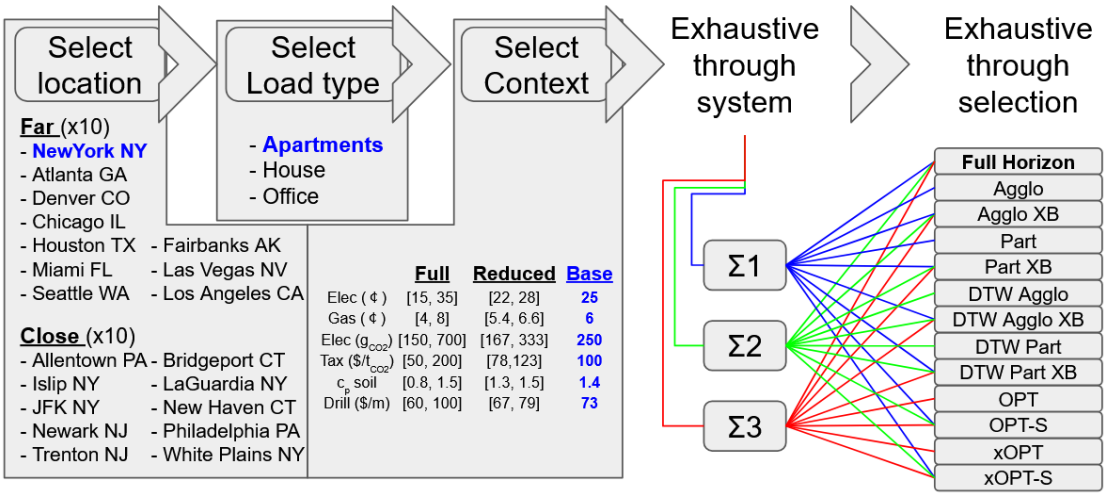


Figure 2.1: Methodology of the investigation

2.2.1 Comparison methodology

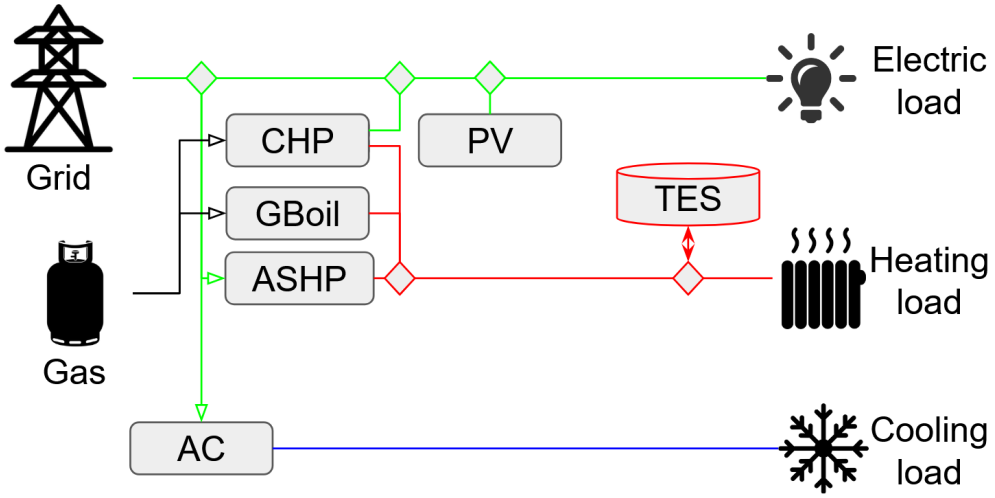
2.2.1.1 Base case

First, the three systems presented in Figure 2.2 are optimized for a base case representing a mid-rise apartments building in NewYork NY, with one specific set of techno-economic parameters.

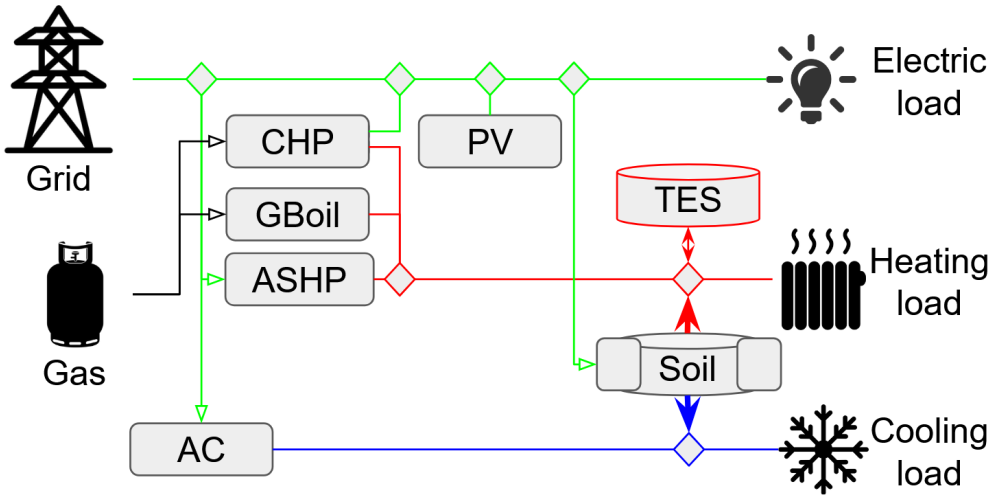
These systems cover a range of difficulty and different usages. The simplest system (Σ1) addresses the trade-off between different heating options, including a possible short-term load offset (via the TES) and the synergy with PV (Figure 2.2a). For the sake of comparison with the other systems, the cooling load was included, with only an AC to cover it.

The two other systems interact with the ground via a reversible ground-sourced heat pump (GSHP), capable of extracting heat and cold. The unit is constrained to yearly cyclic usage via modeling of a ground seasonal storage, detailed in Appendix A. The extraction of heat will cool down the ground, by transferring energy from a "warm" soil compartment to a "cold" compartment. The inverse principle occurs for system cooling.

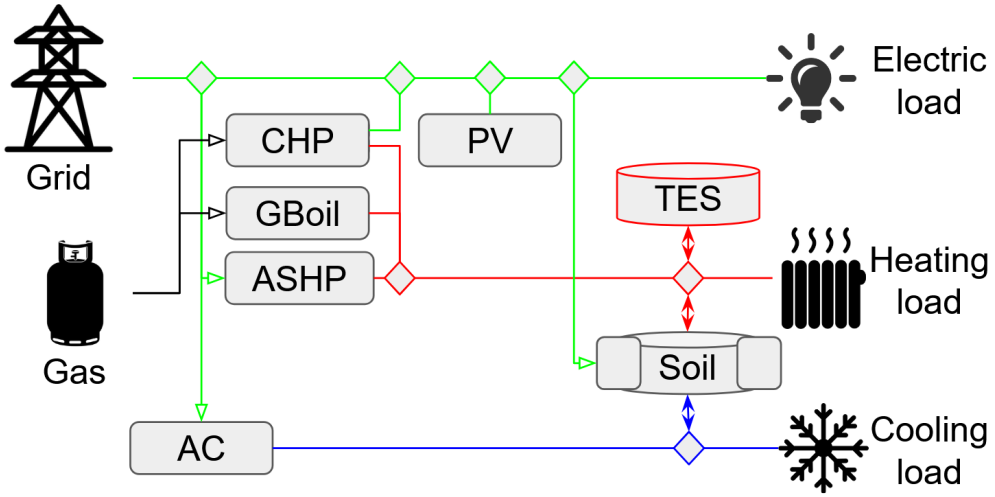
The second system (Σ) can extract heat or cold, but cannot actively store excess heat or excess cold (Figure 2.2b). The third system (Σ3) is almost identical (Figure 2.2c), but can actively store excess heat and cooling via the GSHP, using an additional component as described in Appendix A.



(a) Simplest system (Σ1)



(b) System with ground-sourced heat pump (Σ2)



(c) System with long term storage. (Σ3)

Figure 2.2: The three system configurations considered.

2.2.1.2 Building and load type

The consistency of selection approaches is investigated through two other building types (Table 2.2), whose load data is published in [77]. The objective is to show the impact when load changes. A single family house is considered as an example of a smaller residential load, i.e. for which the scheduling is comparable to the base mid-rise apartments building. An office building is also considered, to investigate an example with different scheduling. This investigation is only conducted with the base techno-economic context in the base location. It relies on data from

Table 2.2: Annual energy demand per building type in New York NY [77]

(MWh/y)	Heating	Cooling	Electricity
Mid-rise apartments	166.6	96.9	206.8
Offices	141	292.2	548.3
House (SFH)	21.7	2.7	8.7

2.2.1.3 Sensitivity to location

The study explores multiple geographical locations to investigate the sensitivity of geographical contexts. The base case considered one unique location (New York NY). This unique location is also considered through all cases investigating the sensitivity to load type and context. The sensitivity to the location considers two sets of locations: a distant set including 10 cities covering the US and its climates as widely as possible, and a local set including 10 locations neighbor to the base case.

Table 2.3: Locations considered for the sensitivity (base 18.3^{o}C)

Far			Close		
City	HDD	CDD	City	HDD	CDD
New York NY	2753	617	Allentown PA	3121	430
Atlanta GA	1570	956	Bridgeport CT	3049	411
Denver CO	3282	349	Islip NY	2759	529
Chicago IL	3188	652	JFK NY	2707	598
Houston TX	817	1556	La Guardia NY	2444	715
Miami FL	77	2330	Newark NJ	2794	667
Seattle WA	2451	84	New Haven CT	3024	365
Fairbanks AK	7236	21	Philadelphia PA	2677	638
Las Vegas NV	1294	1842	Trenton NJ	2750	648
Los Angeles CA	724	178	White Plains NY	3195	324

2.2.1.4 Sensitivity to contextual parameters

The study covers multiple contexts to investigate the sensitivity to impactful parameters that may vary, be uncertain, or depend on specific situations. The base case considered one unique context. This unique context is also considered through all cases investigating the sensitivity to load type and to location. The sensitivity considers two ranges of variations of the context variables: a large range covering most US contexts, and a reduced range of ±15% of the full range centered around the base context.

A total of 8 large and 8 reduced contexts, in addition to the base context, are selected using Sobol sequences. These contexts are considered through all three systems, the unused context parameters (e.g. electricity related for the off-grid system) are ignored by the system.

Table 2.4: Context parameters and considered ranges

Name	Sign	Range/Value			Unit
		Large	Reduced	Base	
Price Elec	C^{fuel}_{Elec}	[15, 35]	[22, 28]	25	¢/kWh
Price Gas	C^{fuel}_{Gas}	[4 , 8]	[5.4, 6.6]	6	¢/kWh
CO_2 intensity Elec	G_{Gas}	[150, 700]	[167, 333]	250	gCO_2eq/kWh
Carbon Tax	C^{CO_2}	[50, 200]	[78, 123]	100	\$/tCO_2eq
Soil Thermal Capacity	c_p	[0.8, 1.5]	[1.3, 1.5]	1.4	kJ/kg.K
Drilling cost	C^{drill}	[60, 100]	[67, 79]	73	\$/m

2.2.2 Error metrics and variables of interest

The design variables of interest are grouped by categories (Table 2.5). The following error metrics are considered:

- Root Mean Squared Error (RMSE) on the normed duration curve (DC) of each series. This error is summed per category $\mathscr{C}$. Only for *TsSource* and *TsStore*.
- Absolute Error. It is considered between the true values y_{nu} and predictions $\tilde{y}_{nu}$ for each design variable ν, then summed per category. Considered for all but *TsSource* and *TsStore*.
- Percentage Error (eq.(2.1)). The error is relative to the sum of true values y_ν of design variables ν within their category $\mathscr{C}$. This formulation avoids divisions by zero when the true value of one design variable is null (e.g.

capacity not installed). It is only used to highlight contributions of design variables to the error in a category.

- Absolute Percentage Error. Same as the percentage error, but using the absolute error. The error is summed through all design variables for error measurement per category.

$$\epsilon_{\%,\nu} = \frac{(y_\nu - \tilde{y}_\nu)}{\sum_{\upsilon \in \mathscr{C}} y_\upsilon} \tag{2.1}$$

Table 2.5: Design variables of interest

Category	**Unit**	**Design variables**
Cost	k$/y	total annualized cost
Carbon	tCO_2eq/y	yearly carbon emissions
CapCold	kW_{th}	- AC capacity - GSCP capacity
CapHeat	kW_{th}	- ASHP capacity - GSHP capacity - GB capacity - CHP heating power* - TES heating power**
CapElec	kW_{el}	- Grid *** - PV capacity - CHP capacity
EnSource	MWh/y	- yearly Gas purchase - yearly Electricity purchase - yearly PV production
EnStore	MWh/y	- yearly usage of the TES - yearly heat intake from ground - yearly cold intake from ground
TsSource	-	- Error on normed DC of Gas purchase - Error on normed DC of Elec purchase - Error on normed DC of PV production
TsStore	-	- Error on normed DC of TES usage - Error on normed DC of ground heat intake - Error on normed DC of ground cold intake

(*) Thermal capacity deduced from Elec capacity and efficiency
(**) Thermal power capacity is 30% of capacity (based on SOC)
(***) Yearly peak demand of grid electricity

Name	Grouping Algorithm	Selection of Representatives	Extremes	Balanced	Distance Metric	Data Format	
Part	Partitioning	Medoids			L2	original	[47, 68]
PartXB	Partitioning	Medoids	✓	✓	L2	original	[47]
Agglo	Hierarchical	Medoids			L2	original	[47]
AggloXB	Hierarchical	Medoids	✓	✓	L2	original	[47]
DtwPart	Partitioning	Medoids			DTW	original	[68]
DtwPartXB	Partitioning	Medoids	✓	✓	DWT	original	
DtwAgglo	Hierarchical	Medoids			DTW	original	
DtwAggloXB	Hierarchical	Medoids	✓	✓	DTW	original	
OptV	OPT	OPT			L1	bins(t)	[50, 70]
OptVS	OPT	OPT			L1	bins(t,dt)	[50, 70]
XOptVOr	Hierarchical	OPT			Mixed	bins(t)	
XOptVSOr	Hierarchical	OPT			Mixed	bins(t,dt)	

L1 refers to the L1 norm, i.e. the Manhattan distance
L2 refers to the L2 norm, i.e. the Euclidean distance

Table 2.6: Methodologies for days grouping and representative days selection

2.2.3 Clustering and selection of representative days

Twelve methods for the selection of representative days are compared in this study (Table 2.6). The distinction between these methods is presented via six key characteristics.

The grouping of days and selection of representatives may be intertwined. The OPT algorithm [70] directly selects representatives and associated weights to minimize the error on approximating the duration curve (DC). Partitioning (k-Medoids) and Hierarchical (agglomerative) clustering group days by similarity, then allow a custom selection of representatives within clusters: cluster medoids and the selection using the OPT algorithm to select one representative and weight per cluster are considered in this study. Cluster-based approaches also allow selecting days of occurrence of yearly extremes (highest hourly heat, cold, electric demand) to represent their cluster [47]. Cluster-based approaches may over- or underestimate the yearly representation of loads and renewable resources, thus a balancing mechanism [63] may modify weights of representatives to equal the yearly total of clustered series. All selection procedures rely on distance metrics, between days for clustering based methods, and to the original DC. For the latter, the Manhattan distance is used to retain linearity in the underlying optimization algorithm. Clustering based methods may rely on the Euclidean distance [47] or

the Dynamic Time Warping (DTW) distance [68]. Finally, the OPT algorithm sorts hourly time series values of each day into bins. The range of values covered by each bin covers 1/24 of the range of values over the whole year. The variations (derivative) of each series during each day may also be considered (bins(t+dt)). For clustering algorithms, the choice was made to use original daily values, each hour being a dimension.

Part and *Agglo* use k-Medoids and agglomerative clustering respectively with Euclidean distance. Their counterparts *PartXB* and AggloXB include the distinction of extreme days and balance the yearly values. All four have a DTW counterpart, using the DTW distance as this measure is common in time-series clustering [78] and suggested to use for MES design problems [68] suggests using it for MES design problems. *Opt* and *OptS* use the OPT algorithm [49, 50] to select representatives, the latter considering bins of the original series and their derivatives. The same distinction is made for *xOpt* and *xOptS*, in which clusters are formed with agglomerative clustering to restrict the selection via the OPT algorithm to one representative per cluster, as per recommendation from [47] to put more emphasis on the selection of representatives.

2.2.4 MES problem formulations

The quality of a selection method is measured against the results of the optimization using the entirety of the time series (called "reference"). This subsection introduces the design problem formulation using the entire time series (section 2.2.4.1), then the changes induced to use uncoupled design days (section 2.2.4.4) and design days linked for seasonal storage modeling (section 2.2.4.5). The MES design problem follows the energy hub concept [20, 29], i.e. intertwine the energy purchase, consumption, production, transformation and storage of energy.

2.2.4.1 Reference formulation using the full time series

The cost J^a_{tot} in eq.(2.2) is the variable to be minimized during the MES design problem. It combines OPEX (operational expenditure) and annualized CAPEX (capital expenditure). The CAPEX considers the purchase and installation of the whole system considering prices C and installed capacities P^{nom}, and is annualized (eq.(2.3)). The OPEX includes energy bills J_{fuel}, carbon taxes J_{CO_2} and maintenance costs $J_{O\&M}$ and is not annualized. Terms in eq.(2.4) and emissions in eq.(2.5) consider net purchases of fuel and power $\dot{Q}$, prices, and carbon intensities G for the latter.

$$J^a_{tot} = J^a_{CAPEX} + J_{fuel} + J_{CO_2} + J_{O\&M} \tag{2.2}$$

$$J^a_{CAPEX} = \sum_{\tau \in \mathcal{T}} NPV_\tau (C^{fix}_\tau + P^{nom}_\tau C^{lin}_\tau) \tag{2.3}$$

$$J_{fuel} = \sum_{f \in \mathcal{F}} \sum_{t \in \mathcal{T}} C^{fuel}_f \ (\dot{Q}^{\mathcal{N},in}_{\sigma_f,t} - \dot{Q}^{\mathcal{N},out}_{\sigma_f,t}) \tag{2.4}$$

$$J_{CO_2} = C^{CO_2} \sum_{f \in \mathcal{F}} G_f \sum_{n \in \mathcal{N}} (\dot{Q}^{\mathcal{N},in}_{\sigma_f} - \dot{Q}^{\mathcal{N},out}_{\sigma_f}) \tag{2.5}$$

2.2.4.2 Stream balance

The system considers different energy streams. Each stream σ must be balanced between generation ($H_{\tau_{cv},\sigma} \times \dot{Q}^{\mathcal{T},in}_{\tau_{cv},t}$), consumption ($\dot{Q}^{load}_{\sigma,t}$), storage interactions ($\dot{Q}^{ch/dch}_{\sigma,\tau_s,t}$), exports ($\dot{Q}^{\mathcal{N},out}$) to an external grid and imports ($\dot{Q}^{\mathcal{N},in}$) from an external source or grid at any time t.

$$\dot{Q}^{load}_{\sigma,t} = \sum_{\tau_{cv} \in \mathcal{T}_{cv}} (H_{\tau_{cv},\sigma} \ \dot{Q}^{\mathcal{T},in}_{\tau_{cv},t}) + \dot{Q}^{\mathcal{N},in}_{\sigma,t} - \dot{Q}^{\mathcal{N},out}_{\sigma,t} + \sum_{\tau_s \ on \ \sigma} (\eta^{dch}_{\tau_s} \dot{Q}^{dch}_{\sigma,\tau_s,t} - \eta^{ch}_{\tau_s} \dot{Q}^{ch}_{\sigma,\tau_s,t}) \tag{2.6}$$

2.2.4.3 Converters

A converter τ_{cv} transforms power from one or multiple streams into one or multiple others. All converters are considered linear. This is common for most technologies (e.g. Boilers, HP, PV, ST) [56, 61, 62, 67, 79, 80]. GSHP and CHP units [61, 62] are also assumed linear for simplicity [80], the latter including a minimum partial load of 40%.

A matrix-based formulation describes the operation (eq.(2.7)). The unitless coefficient $H_{(\tau_{cv},\sigma)}$ binds each converter τ_{cv} to each stream σ. For HP, AC, GSHP

and photovoltaic (PV) panels, this coefficient is precomputed for every time step t, it is constant for other converters (Appendix B). It is negative for inputs and positive for outputs. A reference input stream σ_{ref} is used to determine other input and output powers of each converter τ_{cv} (eq.(2.6)). PV cannot be curtailed and must convert all received irradiation; a transformed factor $\xi^{sun}_{PV,t}$ (in kW/kWp) is precomputed to account for direct and diffuse irradiation, air temperature, tilt and orientation [81].

$$\begin{cases} 0 \leq H_{(\tau_{cv},\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{\tau_{cv},t} \leq P^{nom}_{\tau_{cv}}, \\ \qquad\qquad\qquad\qquad \text{if } \tau_{cv} \text{ is dispatchable} \\ H_{PV,\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{PV,t} = \xi^{sun}_{PV,t}\ P^{nom}_{PV} \end{cases} \tag{2.7}$$

The Big$M$ method[1] is used to impose a minimum partial load on the operation of the CHP unit. The modeling uses a binary γ and an arbitrarily large number M to constrain the CHP unit to only turn on to generate power above $z = 40\%$ of its nominal capacity (eq.(2.8)).

$$\begin{cases} z\ P^{nom}_{CHP} \leq H_{(CHP,Elec)}\ \dot{Q}^{\mathcal{T},in}_{CHP,t} + M\ (1 - \gamma_{CHP,t}) \\ \dot{Q}^{\mathcal{T},in}_{CHP,t} \leq \gamma_{CHP,t}\ M \\ \gamma_{CHP,t} \leq \dot{Q}^{\mathcal{T},in}_{CHP,t}\ M \end{cases} \tag{2.8}$$

A pair of GSHP-GSCP is used to model a reversible Ground-Sourced Heat Pump. We call GSCP a GSHP designed for cooling purposes. These units cannot operate simultaneously (eq.(2.9)), one is connected to the warm stream, the other to the cold stream. Both interact with the ground, modeled as a couple of storage units (c.f. Appendix A). The reversible GSHP unit is not included in the layout of $\Sigma 1$ and included for $\Sigma 2$. for $\Sigma 3$, it is included with an additional component enabling active storage of excess heat and cold, as detailed in appendix.

$$\gamma_{GSHP,t} + \gamma_{GSCP,t} \leq 1 \tag{2.9}$$

The storage units link time steps and displace productions to better match the demand. The thermal energy storage (TES) and components of the ground model comply to the following sets of rules. The model describes a yearly cycle,

[1] https://en.wikipedia.org/wiki/Big_M_method

thus the State Of Charge (SOC) at the start and end of the time horizon $\mathbb{T}$ must be identical for every storage (eq.(2.10)). The maximum charging and discharging power of every storage is set proportional to its capacity (eq.(2.11-2.12)). The capacity of every storage binds its state of charge, as depicted by eq.(2.13) The time coupling equations are expressed via eq.(2.14), in which the decay κ, charging η^{ch} and discharging η^{dch} efficiencies are considered.

$$SOC_{\sigma,\tau_s,\ t^{max}+1} = SOC_{\sigma,\tau_s,\ t=0} \tag{2.10}$$

$$0 \leq \dot{Q}^{ch}_{\sigma,\tau_s,t} \leq \beta^{ch}_{\tau_s} \times P^{nom}_{\tau_s} \tag{2.11}$$

$$0 \leq \dot{Q}^{dch}_{\sigma,\tau_s,t} \leq \beta^{dch}_{\tau_s}\ P^{nom}_{\tau_s} \tag{2.12}$$

$$0 \leq SOC_{\sigma,\tau_s,t} \leq P^{nom}_{\tau_s} \tag{2.13}$$

$$SOC_{\sigma,\tau_s,t+1} = (1-\kappa_{\tau_s})SOC_{\sigma,\tau_s,t} + \eta^{ch}_{\tau_s}\dot{Q}^{ch}_{\sigma,\tau_s,t} - \frac{1}{\eta^{dch}_{\tau_s}}\dot{Q}^{dch}_{\sigma,\tau_s,t} \tag{2.14}$$

Finally, all converters and storage are subject to a maximum capacity limitation. A rooftop area A^{max} limits the installation of PV, assuming a factor of $\alpha^R_{PV} = 8m^2/kWp_{el}$ (eq.(2.15)). For the ground model, limits are set to reflect acceptable ground temperature variations (Appendix A). The limit for other technologies are based on the peak demand of each building (Appendix B).

$$\alpha^R_{PV}\ P^{nom}_{PV} \leq A^{max} \tag{2.15}$$

2.2.4.4 Formulation with uncoupled days

The consideration of representative days induces necessarily changes to the problem formulation. First, the original time frame $t \in \mathbb{T}$ is subdivided into days

$\delta \in \{0, ..., \Delta\}$, each with 24 time steps $h \in \{0, ..., H\}$. The set of days is reduced to only include representative days $d \in \{0, ..., D\}$. All time-related constraints are transformed accordingly, trading the variable t for the couple (d,h).

Each representative day is associated with a weight w_d. The energy bill (eq.(2.4)) and carbon emissions (eq.(2.5)) are reformulated to use a weighted sum of each representative day instead of a sum over all time steps.

In the consideration of uncoupled representative days [47,63], the storage looping from eq.(2.10) is replaced by one equation per representative day eq.(2.16), forcing a daily cycle.

$$SOC_{\sigma,\tau_s,d,t_i^{max}+1} = SOC_{\sigma,\tau_s,\ d,t=0} \tag{2.16}$$

2.2.4.5 Formulation with coupled days

In the formulation for coupling representative days proposed by Kotzur et al. [61], additional modeling efforts are put on the storage modeling. All other operation and totals are similar to the formulation with uncoupled days.

For the storage units, the SOC is divided into two variables operating on two distinct time grids: the intra-day (hourly resolution) and the inter-day (daily resolution) time grids (Figure 2.3). The original sequence of days is retained from the selection of representative days to ensure a meaningful representation on the inter-day grid.

On the *intra-day* time grid (within a day), all converters and storage (charge, discharge) are operated similarly to the uncoupled days formulation. For each storage, the variations of SOC within a day ($SOC^h_{\tau_s}$) is initialized at the start of each day (eq.(2.17)) and evolves with the daily operations (eq.(2.18)) without looping constraint on the day. The variation of SOC SOC^h is considered within representative days.

$$SOC^h_{\sigma,\tau_s,d,h=0} = 0 \tag{2.17}$$

$$SOC^h_{\sigma,\tau_s,d,h+1} = (1-\kappa_{\tau_s})SOC^h_{\sigma,\tau_s,d,h} + \eta^{ch}_{\tau_s}\dot{Q}^{ch}_{\sigma,\tau_s,d,h} - \frac{1}{\eta^{dch}_{\tau_s}}\dot{Q}^{dch}_{\sigma,\tau_s,d,h} \tag{2.18}$$

The *inter-day* time grid has a daily resolution. On this time grid, the low-resolution part of the SOC (SOC^d) is considered for each day $\delta \in \Delta$ in the original

time line. This introduces 365 variables per storage, regardless the number of representative days. The variable represents the state of charge at the start of each day. This variable binds representative days together via equation eq.(2.19): its value for the next day ($SOC^d_{\delta+1}$) considers 24 hours ($H+1$) of decay over its value on the current day (SOC^d_{δ}), and the variations at the end of the current day ($SOC^h_{h=H+1}$). The term $\phi : \delta \mapsto d$ is used to denote the function associating each original day to its corresponding representative day. At any moment, the true SOC can be obtained through the sum of the variations of SOC (SOC^h) on the *intra-day* time grid (eq.(2.18)) and the low-resolution SOC (SOC^d) on the *inter-day* time grid (eq.(2.19)), adjusted with the adequate decay.

$$SOC^d_{\sigma,\tau_s,\delta+1} = (1-\kappa_{\tau_s})^{H+1}\ SOC^d_{\sigma,\tau_s,\delta} + SOC^h_{\sigma,\tau_s,\phi(\delta),h=H+1} \tag{2.19}$$

A yearly looping is enforced (eq.(2.20)).

$$SOC^d_{\sigma,\tau_s,\Delta+1} = SOC^d_{\sigma,\tau_s,\delta=0} \tag{2.20}$$

Eq.(2.21-2.22) express the relation between storage capacity and SOC.

$$0 \leq (1-\kappa_{\tau_s})^{h}\ SOC^d_{\sigma,\tau_s,\delta} + SOC^h_{\sigma,\tau_s,\phi(\delta),h} \leq P^{nom}_{\tau_s} \tag{2.21}$$

$$0 \leq \dot{Q}^{ch/dch}_{\sigma,\tau_s,\phi(\delta),h} \leq \beta^{ch/dch}_{\tau_s}\ P^{nom}_{\tau_s} \tag{2.22}$$

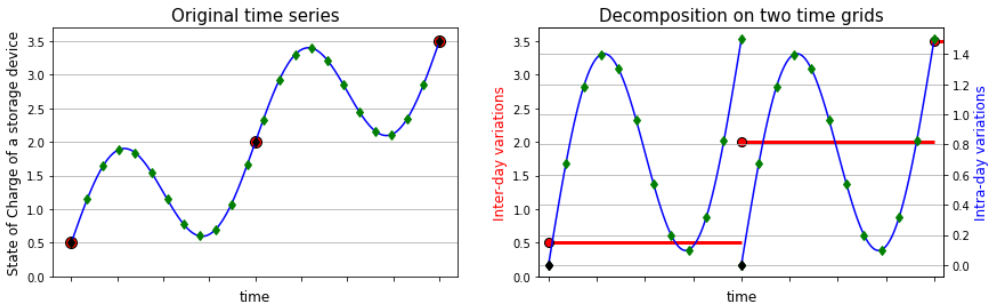
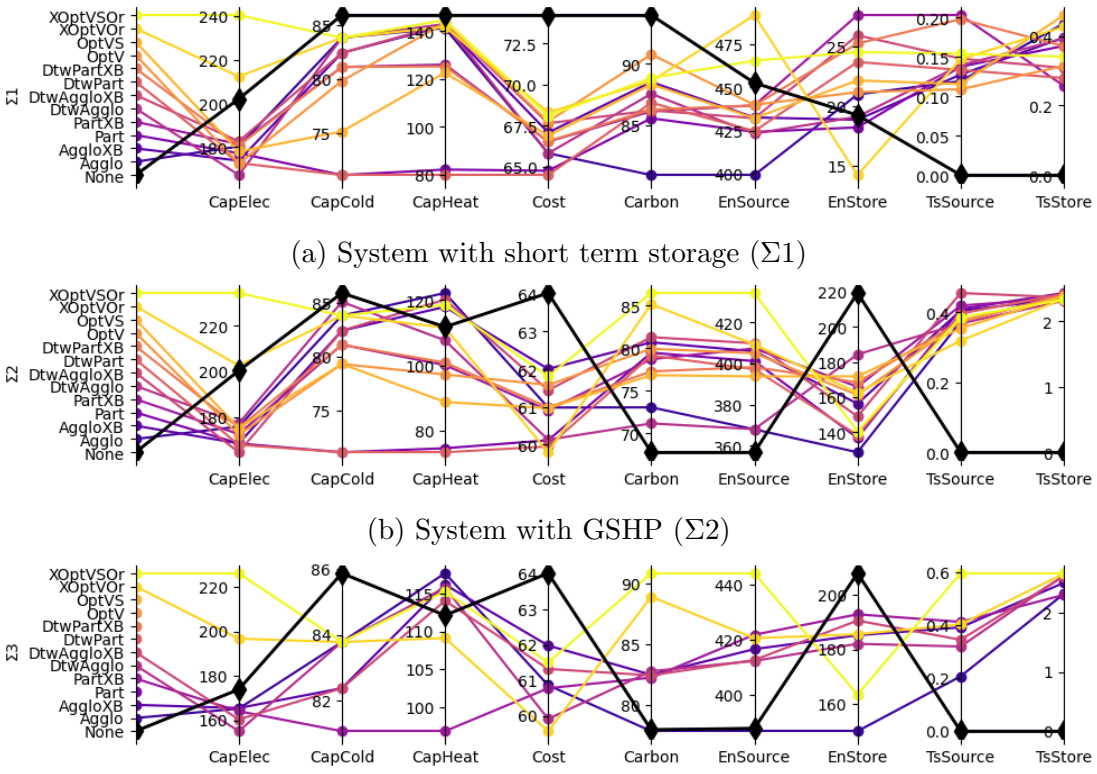


Figure 2.3: Principle of separation of the original state of charge into two different variables on different time grids: intra-period and inter-period [61]. The original hourly SOC (left) is split into days. One variable with hourly resolution represents the variations of SOC during the day, starting at zero (blue on the right). One variable with daily resolution represents the SOC at the start of each day (red).

2.3 Results

Section 2.3.1 analyzes the results of the base case only. The methodology is demonstrated to draw first conclusions with this one unique case. Later sections 2.3.2, 2.3.3 and 2.3.4 show how the conclusions from section 2.3.1 are altered when considering larger scopes.

2.3.1 Base case: validation across systems



(a) System with short term storage (Σ1)

(b) System with GSHP (Σ2)

(c) System with active long-term ground thermal storage (Σ3)

Figure 2.4: Results of the design optimization for the base case. The colors allow the methodology to be backed across the variables; black is the full-resolution case.

Figure 2.4 presents parallel-coordinates plots showing the values of design variables obtained by the reference using the year of data and the values obtained after use of representative days. Results for all variables follow similar generic trends, as the orders of magnitudes of values correspond. Some biases emerge, consistent through all systems in different magnitude: the cost, total capacity of coolers and electric generators is systematically underestimated. Most approaches using representative days estimate the total heating capacity adequately. Differences also arise between systems: the use of representative days underestimates the carbon

emissions for Σ1, while it overestimates the carbon emissions, energy purchases, and underestimates the use of storage for Σ2 and Σ3.

Observations confirm the expected similarities between Σ2 and Σ3, in reference values and trends when using representative days. Despite similarities, several selection methods result in infeasible solutions in the slightly more complex Σ3. This indicates a potential reduction of robustness in using representative days when increasing the model complexity.

The TsSource and TsStore variables already represent differences with the reference. The *OPT* and *xOPT* selection methods specifically aim to minimize the a-priori error on the duration curve (DC). However, their resulting error level on the energy imports and on the storage operation a-posteriori is not better than these of other selection methods. This highlights that minimizing the error on input series does not always correlate with improvements in output accuracy.

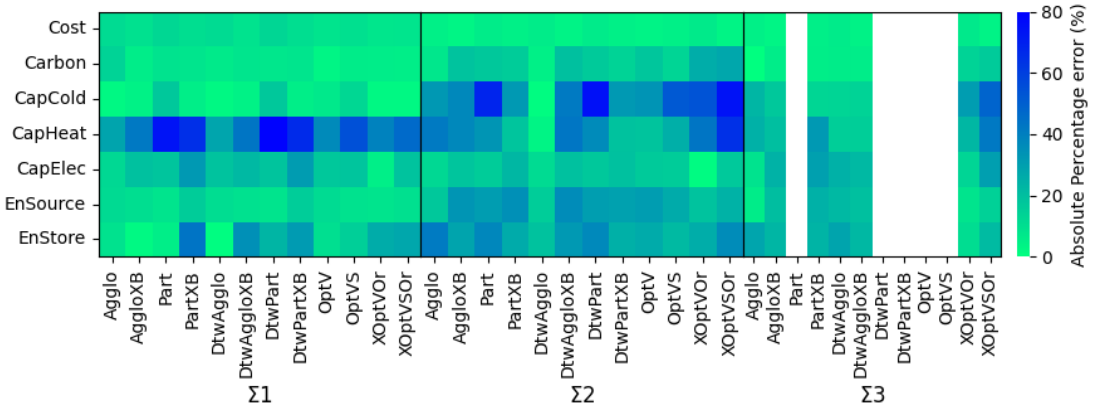


Figure 2.5: Absolute Percentage Error to the reference for all results on the base case

Figure 2.5 shows the absolute percent error between references and agglomerated optimizations, averaged across the systems. The figure shows that the cost is always the most faithfully approximated with representative days, ranging mainly between 0% and 10% error with all aggregation methods. Capacities for heating overall and for cooling in Σ2 show larger errors for almost all methods. Overall, error levels are more consistent through aggregation methods than across variables or systems.

Figure 2.6 presents heatmaps of the correlations between absolute percent errors for different design variables across all representative day selection methods. It complements the findings from Figure 2.5. A strong correlations between two variables suggests that, when a selection method reduces the error for one variable, it often reduces the error for the other variable as well. Across all systems, two

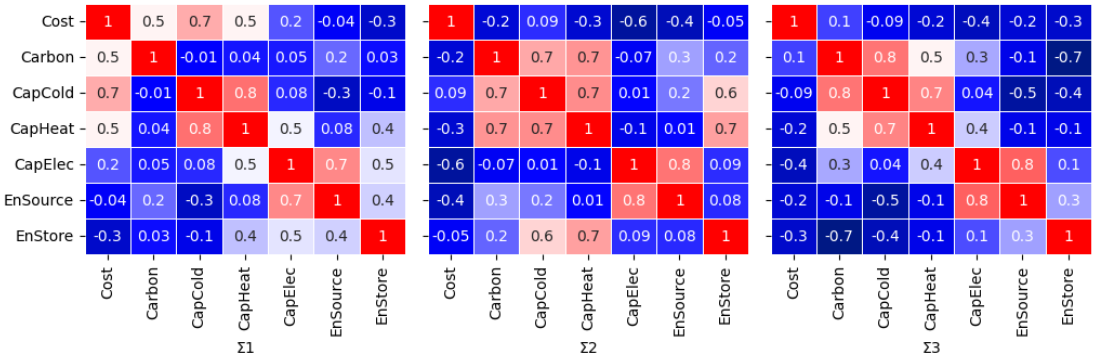


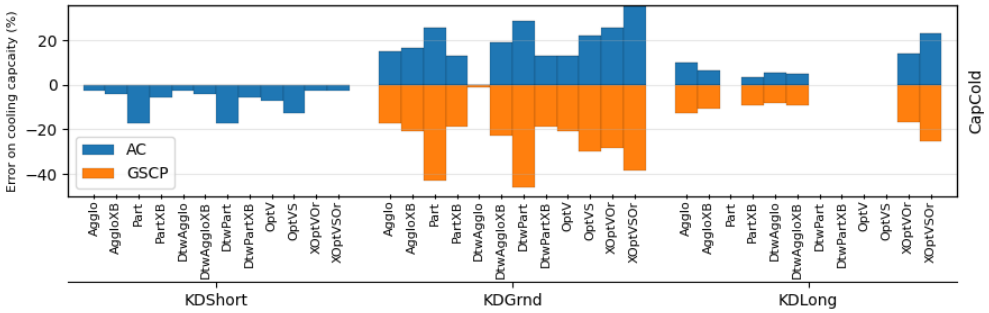
Figure 2.6: Correlation coefficients between the scores of different variables

consistent patterns emerge: a correlation between errors in heating and cooling capacities (CapHeat and CapCold) and another between errors in electrical capacity (CapElec) and energy sourcing (EnSource). For Σ1, the correlation between errors in cost and cooling capacity is mild. It is non-existent for other systems. This indicates that lower cost errors do not necessarily translate to better performance in other design aspects. In contrast, for Σ2 and Σ3, the error in carbon emissions shows a slight correlation with errors in thermal capacities. Additionally for Σ2 alone, a weak correlation between errors in thermal capacities and storage usage is found.

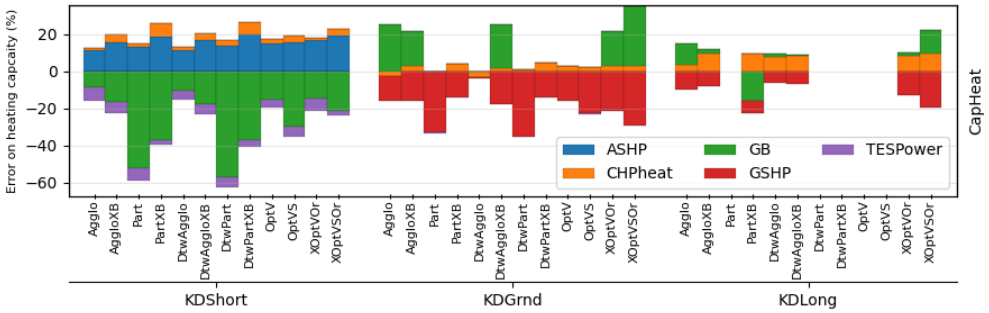
Figure 2.7 shows the percentage errors for the detailed design variables within each system. Notably, error levels are more consistent across different selection methods for a given system than they are for the same method applied across different systems. A surprising observation is the substantial variation in design errors between Σ2 and Σ3, despite their structural similarities.

For Σ2 and Σ3, the GSCP (Figure 2.7a) and GSHP (Figure 2.7b) capacities are consistently underestimated. This indicates a misrepresentation of ground heat exchanges when using representative days. These underestimations are often compensated by an overestimate of the use of AC for cooling, and either by gas boiler or CHP unit for the heating. The nature of error on the thermal aspect of Σ1 differs significantly. Gas boiler capacities are underestimated, with large variations (10% to 50%) across selection methods. ASHP capacities are overestimated of 10% to 20% as compensation.

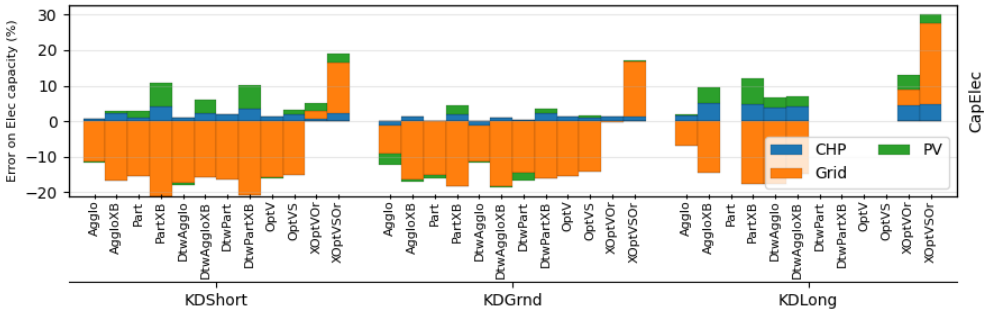
Regarding electrical capacities, most selection methods display similar errors, with grid import capacities being underestimated by 10%-20% across most methods. *xOPT* methods are the exceptions. The use of representative days appears to reduce the peak demand, with few methods compensating this underestimation. This induces a minor underestimate of the yearly electricity purchases. For Σ2



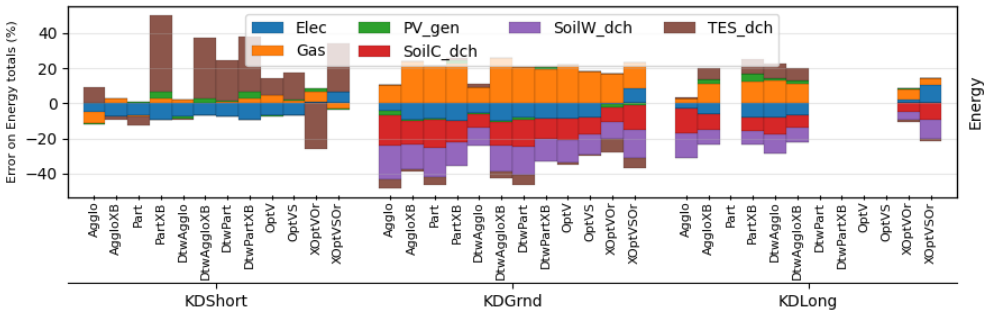
(a) Percentage error on Cooling capacities



(b) Percentage error on Heating capacities



(c) Percentage error on Electric capacities



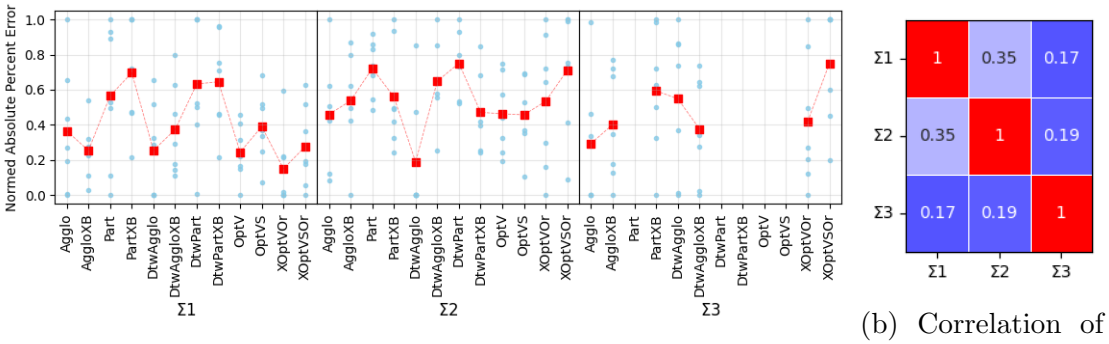
(d) Percentage error on yearly energy (purchase and battery usage)

Figure 2.7: Detail of the error across each variable.

and $\Sigma 3$, this is also explained by the compensatory overestimate in gas usage due to underestimates on the GSHP.

Overall, the correlation patterns from Figure 2.6 are confirmed: methods un-

derestimating AC capacity on Σ1 also underestimate gas boiler usage, while overestimating ASHP capacity. In Σ2 and Σ3, the consistent underestimation of GSHP and GSCP capacities is compensated by overestimates of AC and gas boiler capacities, thus methods better capturing heating capacities also tend to perform better for cooling capacities.



(a) Average error through (normed) variables

(b) Correlation of averages

Figure 2.8: Consistency of selection methods through systems

Figure 2.8 presents the average performance score of each selection method across the three systems. This score is calculated using the absolute percent error, normed within the pool of error results for each variable, then averaged through variables. The metric only serves the comparison between systems and selection methods. As shown in Figure 2.8a, there is no clearly superior aggregation method for all systems. The performance of methods varies significantly from one system to another: for instance, the DtwAgglo method excels in Σ2 but does not stand out for systems 1 and 3. Similarly, the *OPT* and *xOPT* methods show strong performance in Σ1 but are generally mediocre across the other systems.

Figure 2.8b further illustrates this lack of consistency by displaying the correlation coefficients of performance across systems. The absence of strong correlations indicates that a method performing well on one system does not necessarily excel on others. The results highlight a lack of robustness: no single method is universally effective across systems. This suggests that increasing system complexity does not guarantee the continued reliability of a given selection method. It is important to note that these findings are based on a single case study, indicating the need for further investigation to generalize these observations.

All aggregation methods underestimate the costs with levels comparable to these found in literature [61]. The simplest system also underestimates emissions and capacities, the two other systems underestimate emissions. Selection methods better at estimating heating capacities also perform good for cooling capacities and carbon emissions. However, the error of aggregation methods is more consistent within systems than between systems. We find no evidence of consistency of

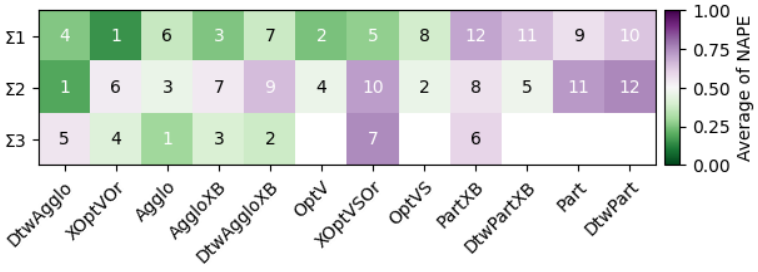


Figure 2.9: Ranking of selection methods for each system, based on the average of standardized error through design variables.

error between the cost and other variables. We find no evidence of consistency of error between systems, i.e. throughout evolution of the system complexity. Based on observations on the base case, Figure 2.9 proposes a ranking of aggregation methods using the average of the absolute percent error through variables. Strong performances of the DtwAgglo method on Σ2, with average results on other systems, place the method as first choice. These preliminary conclusions derived from the base case will be further validated in the subsequent sections.

2.3.2 Sensitivity to building type

In this section, we extend the analysis to two additional building types: an office building with approximately twice the energy load of the base mid-rise apartment building and a single-family house.

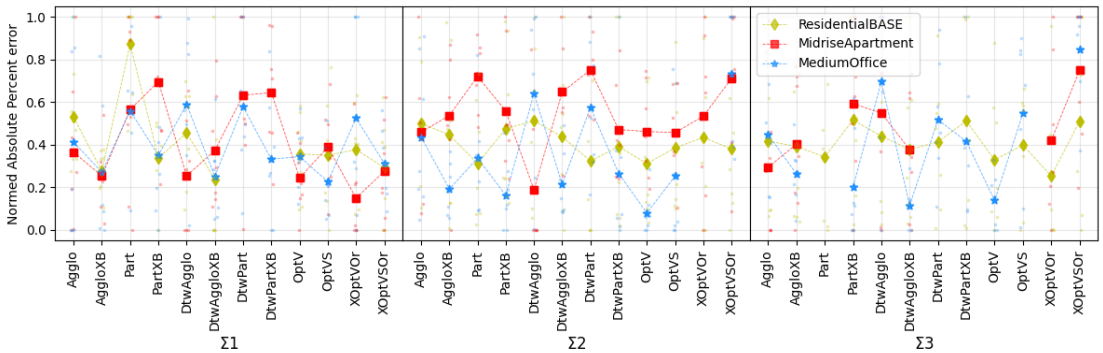
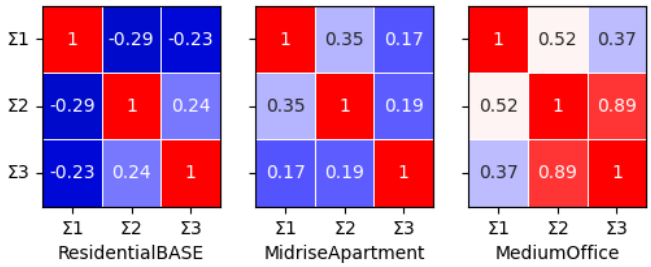


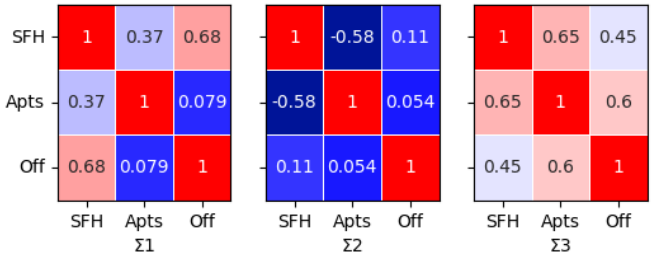
Figure 2.10: Average error through (normed) variables

Figure 2.10 shows the average performance score of each selection method across the three systems for these new building types. It provides a broader perspective beyond the base case results shown in Figure 2.8a. For the single-family house, the differences are small between selection methods for Σ2 and Σ3, with scores between 0.35 and 0.45. The variability increases for Σ1, allowing

more differentiation between methods. For the office building, strong variability in performance across all systems is found, with scores ranging from 0.1 to 0.65. While the DtwAgglo method performed well for the mid-rise apartment case, it struggles with the new building types. For example, the method ranks the worst for the two other building types on the Σ2. While the performance of the AggloXB method is similar for all buildings on Σ1, no other selection method shows any consensus among building types on any other system.



(a) Correlation of error between systems



(b) Correlation of error between buildings

Figure 2.11: Pearson correlation coefficients of the absolute error

Figures 2.10 and 2.11 confirm the lack of correlation between errors across different systems (Figure 2.11a) and across different building types (Fig 2.11b). For the office building, a moderate correlation is observed between the Σ2 and Σ3. This observation is expected due to the structural similarity of these systems. Surprisingly however, no correlation was found between the sesystems for other building types. Additionally, errors across the two residential building types show no apparent correlation, further underscoring the variability in method performance across different contexts.

Figure 2.12 presents the performance ranking of each selection method across all building types and systems. The methods are ordered based on the cumulative sum of their ranks. Notably, this updated ranking differs from the results observed in Figure 2.9. While the DtwAgglo method was the top performer for the mid-rise apartment case, it demonstrates regular but poor performance across systems for the two other buildings.

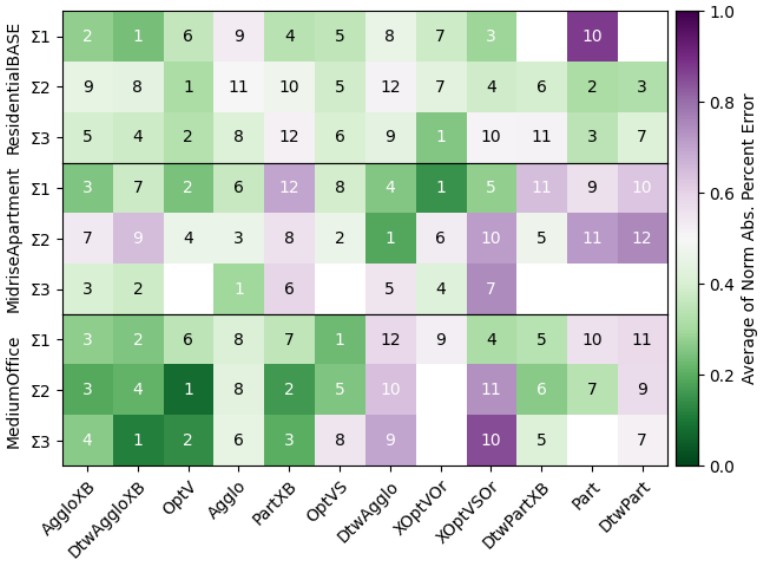


Figure 2.12: Ranking of selection methods for each system and building type

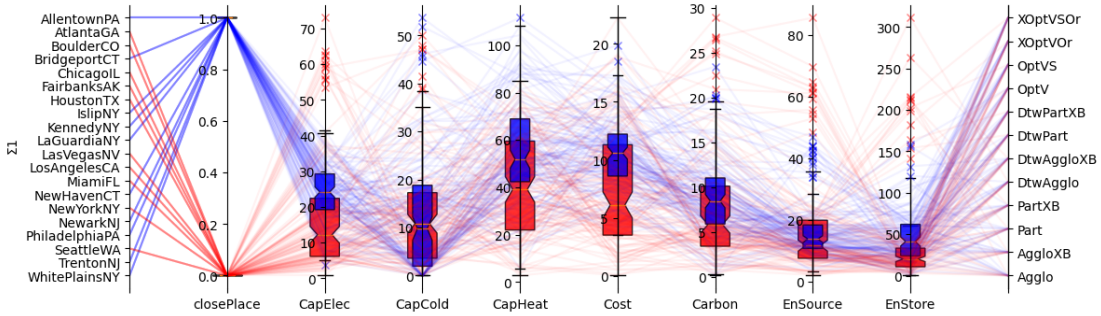
No aggregation method consistently performs better or worse than others across all three systems and building types. If a ranking were to be made, three aggregation methods perform slightly more consistently overall. However, the results highlight that performance of a method for one load type does not necessarily translate to another.

2.3.3 Sensitivity to location

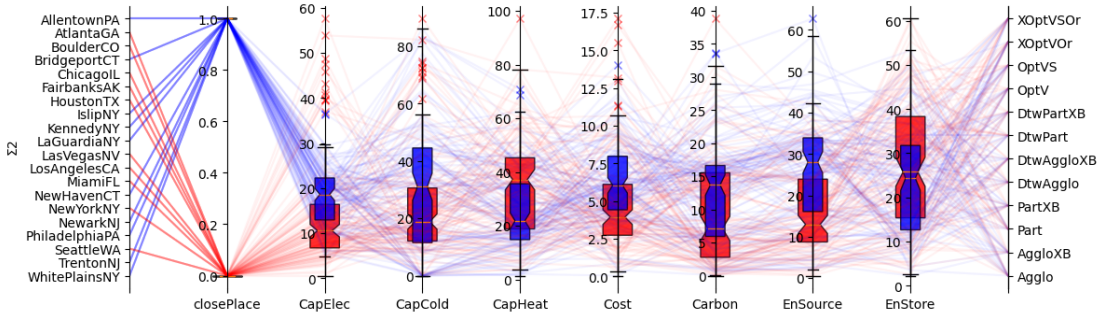
The sensitivity of design choices is tested across multiple locations for each aggregation method. The locations were divided into two categories. “Close” locations cover 10 sides within 100 miles of the base case. “Far” locations span diverse climates and regions across the whole US.

Figure 2.13 presents a combination of parallel-coordinates and box plots, capturing the diversity of errors on design variables across all aggregation methods and systems. This figure gives an ensemble view, but does not detail the performance of individual aggregation methods.

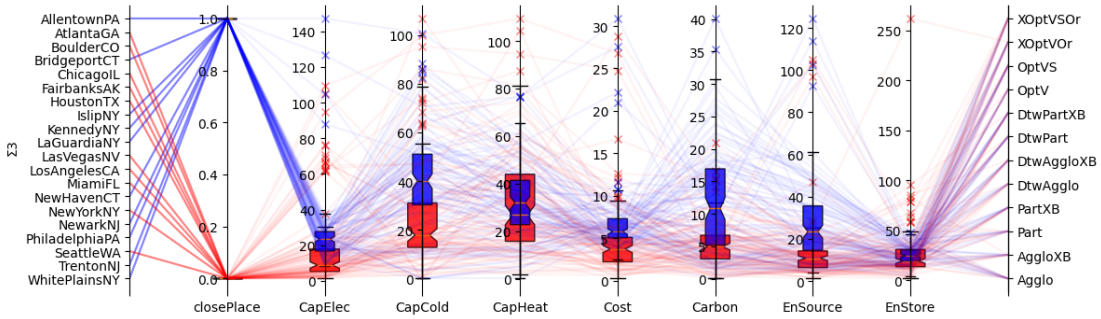
Across all three systems, error ranges remain relatively consistent for each variable. For example, thermal capacities tend to exhibit higher error levels, while cost and carbon emissions are predicted with greater accuracy. Depending on the system analyzed, differences in values and spread are observed between the “Close” and “Far” cases. Observations of individual lines reveal that isolated cases achieve perfect predictions for specific variables, though no single case excels



(a) Simple system (Σ1)



(b) System with ground interaction (Σ2)



(c) System with active ground storage (Σ3)

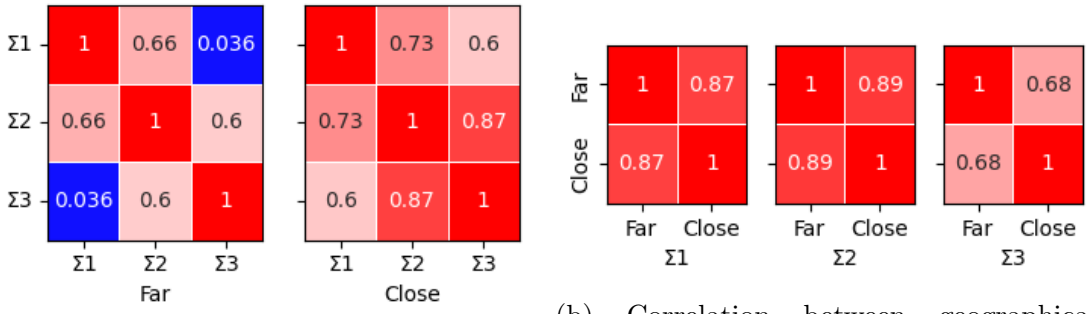
Figure 2.13: Absolute percentage errors through locations and systems

across all variables.

For Σ1 (Figure 2.13a), the "Close" cases exhibit less variability but higher overall errors. This suggests that the geographical region around the base case poses greater challenges in modeling of Σ1. The cooling capacity is an exception, though the median error remains high for a case characterized by the absence of design choice on the cooling system.

For Σ2 (Figure 2.13b), the error spread is comparable between "Close" and "Far" cases. However, the median error tends to be higher for the "Close" locations, except for storage usage (similar performance) and heating capacity (lower error). The levels of error for cooling capacities, heating capacites, and energy management strategies remain within similar ranges.

For Σ3 (Figure 2.13c), the variables related to cost, carbon emissions, and energy sourcing/storage show lower error variability compared to other systems. The error spread for carbon emissions aligns closely with that of the Σ2, while the errors for cooling capacity are higher but with reduced variability. Additionally, isolated cases with exceptionally high errors are more frequent and tend to be more extreme in Σ3 than in the other systems.



(a) Correlation between systems

(b) Correlation between geographical scopes

Figure 2.14: Correlations of errors across different cases

Figure 2.14 presents correlation matrices that evaluate the consistency of errors across systems and proximity groups. Following the methodology from Sections 2.3.1 and 2.3.2, higher correlations indicate better alignment of error levels (averaged across all variables) for each aggregation method, either between systems (Figure 2.14a) or between proximity groups (Figure 2.14b). Given the larger set of cases analyzed here, the scores were averaged by proximity to focus on the overall performance within each group.

Overall, correlations between systems are stronger than those previously observed when comparing different building types. The averaging across multiple locations smooths out the variations in performance variations previously observed. In contrast, it helps highlighting methods with more stable behavior across diverse geographical scopes. Notably, Figure 2.14a reveals a strong correlation between Σ2 and Σ3 for "Close" cases, with weaker correlations involving Σ1. For "Far" cases, the correlation between Σ1 and Σ3 decreases. This indicates that greater geographical diversity may reduce the ability of aggregation methods to effectively capture system-specific information.

Figure 2.14b shows that errors align more consistently between "Close" and "Far" cases for systems 1 and 2. The correlation between proximity groups is weaker for Σ3, which can be attributed to the absence of correlation noted between systems 1 and 3 in the "Far" cases in Figure 2.14a. While some aggregation methods show resilience to the change of geographical scope, this observation is not systematic. Particularly, the results suggest that an increase in model complexity

strengthen the bounds to the geographical scope.

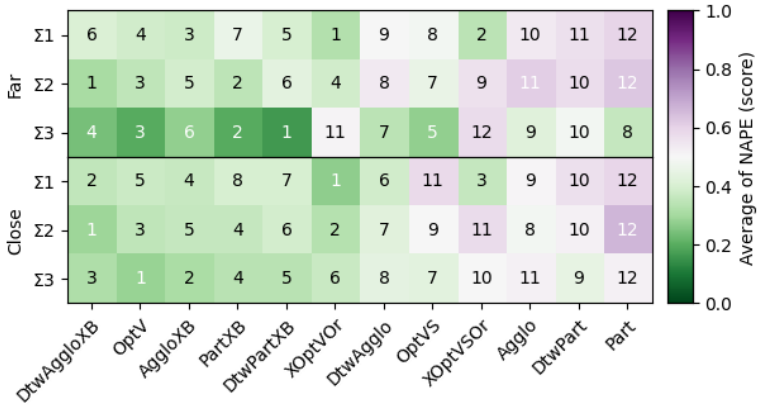


Figure 2.15: Ranking of methods based on distance to the reference

Figure 2.15 presents a ranking of the aggregation methods based on their average performance scores across systems and proximity levels. The horizontal ordering reflects the overall average scores. Displayed ranks and performance levels are specific to each system and proximity context.

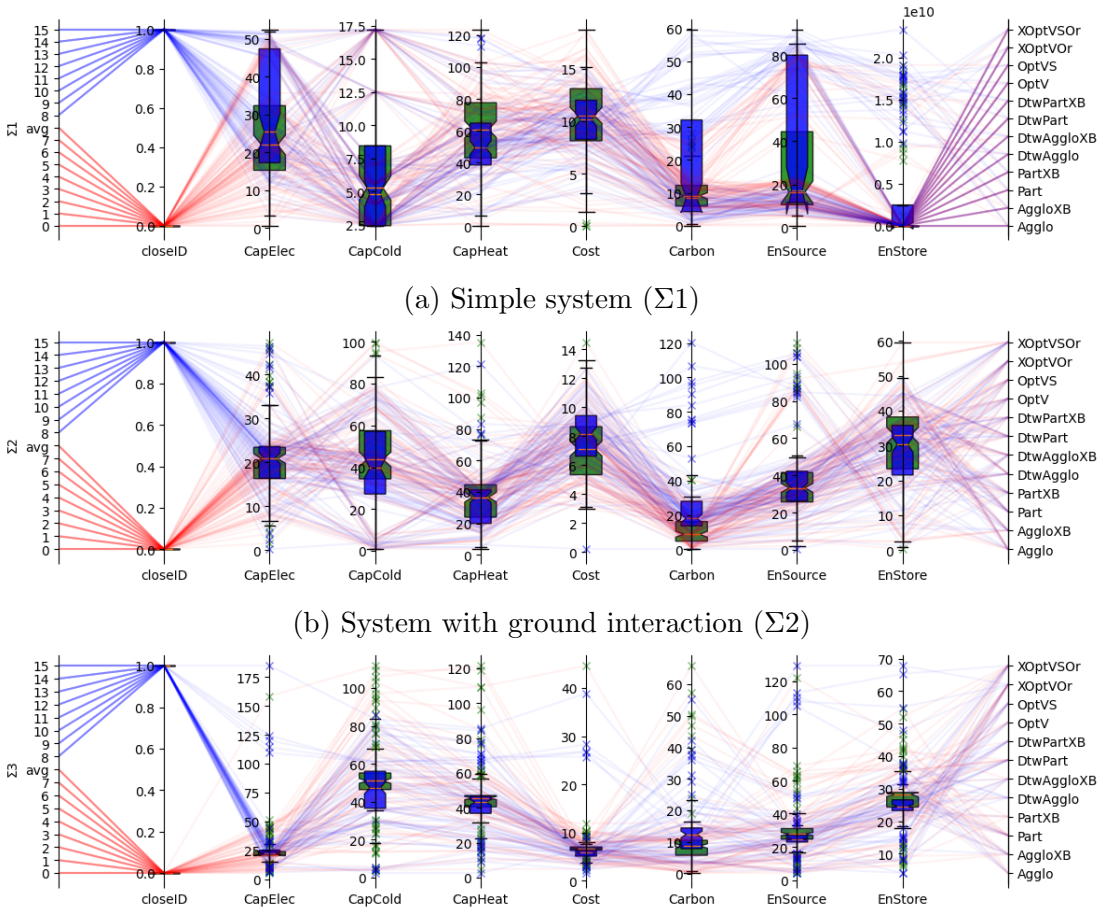
The averaging of multiple cases within each proximity category smooths out some of the differences between methods. As a result, the color distinctions in Figure 2.15 are fainter. However, a few top and bottom performers still stand out. Notably, methods that excel (or perform poorly) tend to maintain their relative ranking across systems and proximity groups. This indicates some level of consistency when a geographical scope is considered.

In particular, three methods (AggloXB, DtwAgglo, and OptV) regularly outperform the others. Their overall ranking aligning with earlier observations across different building types. In contrast, Part and DtwPart methods are among the worst performers, confirming observations from both the base case and the building type analysis. While the DtwAgglo method showed strong performance in the base case, it now emerges as an average method for all systems and geographical scopes in this broader analysis.

Interestingly, spreads of errors for geographically “Close” cases are occasionally larger than for “Far” cases. Three aggregation methods appear to perform slightly better overall. The results suggest that, in a context when geographical diversity is considered, extrapolating performance conclusions to different systems and geographic scopes and systems is generally acceptable, but offers not guarantees.

2.3.4 Sensitivity to context

The sensitivity of the design choices to varying techno-economic contexts (detailed in Table 2.4) was analyzed for each aggregation method. The contexts are categorized into two groups: those covering a broad range of values ("Far") and those restricted to a narrower range around the base case values ("Close").



(a) Simple system (Σ1)

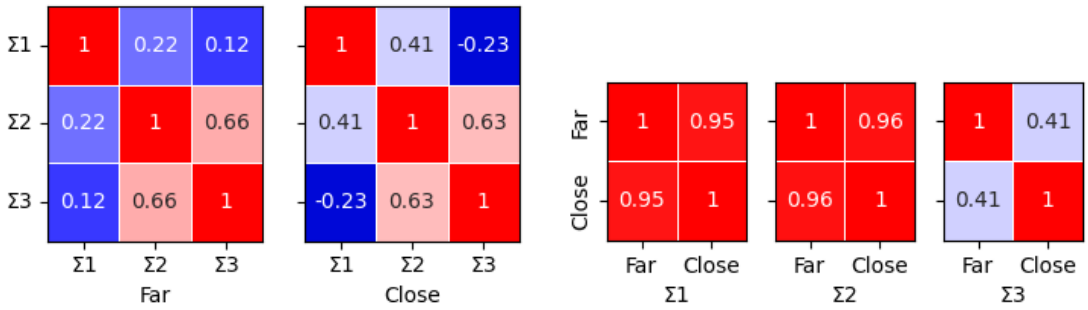
(b) System with ground interaction (Σ2)

(c) System with active ground storage (Σ3)

Figure 2.16: Abs pct errors through sampilngs and systesms

Figure 2.16 shows the diversity of error on design outcomes across all selection methods, separated by system. The Figure provides a broad overview of performance without detailing the specifics of each method. Across all three systems, error ranges remain relatively consistent for each variable. Thermal capacities exhibit higher levels of error, while Cost and Carbon are more accurately predicted overall. This aligns with previous observations. Clear differences in the spread of errors between "Close" and "Far" cases are observed for each system. Surprisingly, "Close" cases often show a greater range of errors than "Far" cases. Examining the plotted lines reveals that rare cases may perfectly predict certain variables, but that none are consistently accurate across all variables. Particularly for Σ1

(Figure 2.16a), the variability in error is notably higher for the "Close" cases for the carbon emissions, yearly imports of energy (EnSource) and electric generation capacityes (CapElec).



(a) Correlation between systems (b) Correlation between context scopes

Figure 2.17: Correlations of errors across different cases

Figure 2.17 shows correlation matrices analyzing the consistency of errors across contexts. Similar to Sections 2.3.1- 2.3.3, higher correlations indicate better alignment of error levels (averaged across all variables) between aggregation methods. The analysis includes correlations between systems (Figure 2.17a) and across different ranges of techno-economic contexts (Figure 2.17b). Due to the larger number of cases analyzed, results were averaged by context resemblance to focus on overall performance within each group.

The correlation matrix in Figure 2.17a reveals generally weak correlations between errors across different systems. A moderate correlation is only observed between Σ2 and Σ3, likely due to their structural similarities. This pattern is consistent for both "Close" and "Far" contexts. This indicates that a change of techno-economic scope does not significantly alter the lack of correlation between systems.

Therefore, Figure 2.17b shows a strong correlation between "Close" and "Far" contexts for Systems 1 and 2. This suggests that performance of selection methods can be somewhat be extrapolated across contexts for these systems. However for Σ3, the absence of correlation indicates that the effectiveness of individual selection methods does not translate well between broader and narrower techno-economic scopes. This divergence of observations between Σ2 and Σ3 is surprising given their structural similarity. Overall, this highlights the difficulty in drawing clear conclusions on the applicability of findings to different techno-economic contexts.

Figure 2.18 presents a ranking of aggregation methods based on their performance across different techno-economic scopes. The horizontal order of methods reflects the average performance score across systems and context resemblance.

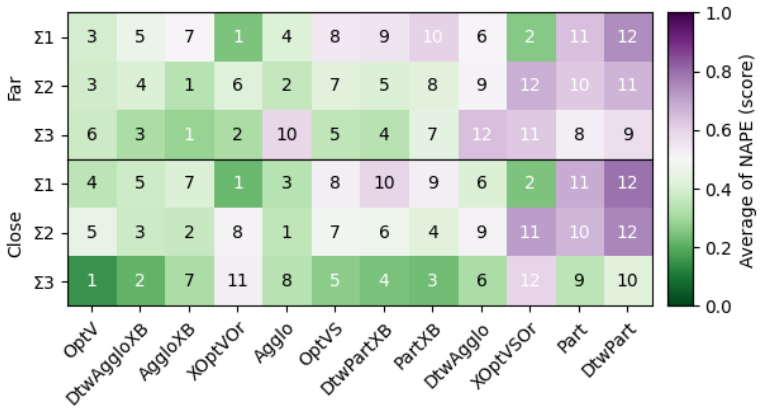


Figure 2.18: Ranking of methods based on difference in context to the reference

Specific rankings and performance levels are specific to each system and context.

The averaging of multiple cases within each proximity category smooths out some of the differences between methods. As a result, the color distinctions in Figure 2.18 are fainter. While a few aggregation methods demonstrate consistent rankings across systems and contexts, trends are less pronounced than with variations of geographical scopes. The three same methods (AggloXB, DtwAgglo, and OptV) are found more consistently reliable. This aligning with earlier findings from building types and geographical sensitivity analyses. Likewise, Part and DtwPart methods consistently rank lower.

Narrower scopes of techno-economic contexts may induce larger variations of performances of selection method than wider scopes. While a few methods perform slightly better overall, their robustness is not guaranteed for specific isolated cases. Additionally, not all systems allow to extrapolate conclusions on the performance of aggregation method across different contexts. This emphasizes the challenge of generalizing method effectiveness when faced with varying techno-economic scopes.

2.4 Discussion

This study explores the consistency of aggregation selection to perform early-stage design of MES relying on representative days. Many different contexts were tested. However the exploration is limited due to the large numbers of degrees of freedom of such application, we only consider some degrees of freedom. Moreover, usages and interests may differ. In this study, the choice is to include all design modeling aspects (sizing of converters, storage, total cost and carbon emissions, total usage

of sources and storage), however most frequent usages of such application usually focus on one variable (e.g. incorporation of one specific technology).

On these limited sets of degree of freedom, the level of reliability was explored. Trends tend to indicate a mild superiority of a few sets of selection methods. Results also show the difficulty to extrapolate conclusion on selection methods form a case to another, whether this case is a slightly different system layout, energy load or range of variation in techno-economic context. This indicates the need to choose the selection method in a manner as close as possible from the actual intended usage of the reduced MES design problem, while acknowledging the level of reliability within that range of usage.

A central interest is the understanding of the link between the selection method and the MES design results. Each selection method is based on the believe, a priori, that the essential information is retained despite the time series aggregation. As understanding this link must be addressed in future research, a validation across diverse contexts is of prime importance. This study reveals that no tested method reaches a systematic level of performance, neither overall nor in comparison to other methods. The three methods found slightly more consistent in their performances are of different nature: the first selects days to reduce the error on the load duration curve of the energy demand and solar availability, the second use euclidean distance and the third a DTW distance to characterize difference between days and group them hierarchically.

The sensitivity exploration across geographical and techno-economic contexts shows that the range of error committed by aggregation methods may be larger for smaller variations of contexts. This counter-intuitive finding is partially explained by the nature of the application. An early stage MES design addresses choices between technologies and energy sources. The non-robust approach may lead to hard bounds in the design spaces, on either sides of which different design choices are made. The problem simplification, necessary to use representative days, result in approximations of the location of such bound. This is dependent on the problem formulation (i.e. sets of constraints), but not on the context (building loads, techno-economic parameters). A narrower scope thus has more chances to be significantly affected by a variation of location of such bound.

2.5 Conclusions

This study assessed the performance of methods to select representative days for the early-stage design of multi-energy system design. Specifically, the study

focuses on their robustness across varying system complexities, building types, geographical locations, and techno-economic contexts.

This exploration finds that the accuracy of selection methods at reproducing the objective function (cost) does not correlate with the accuracy on other design variables. The identification of a best suited selection method is possible for a specific problem in a specific context. The extrapolation of conclusions to a more difficult design problem is not recommended. Mild variations in scheduling, intensity of the demand (e.g. via change of geographical scope) or techno-economic context may retain the validity of the identification, but conducting a validation over the scope of interest is recommended.

This study also concludes that none of the considered selection method perform consistently on all explored contexts, nor does it consistently outperforms all other methods. Three selection methods show slightly more regular performances across all problems, however their difference in formulation limits the understanding of the reasons for their superiority.

In conclusion, this study demonstrates that the effectiveness of selection methods is highly context-dependent. All existing methods to select days are based on a belief, a priori, that these retain the most important information. Isolated validations on single cases may be misleading, and a systematic characterization of the error over the range of expected usages of the reduced problem is found necessary.

Chapter 3

The influence of time horizon on the design of multi-energy system

Abstract

Designing multi-energy systems with interdependent loads, energy sources and devices is a computationally intensive task requiring at least a year of hourly data per load and source. Methods (e.g. using representative days) address this shortage, however none of these attempt to identify the important statistical features in the time series influencing the system design.

This study investigates the influence of using shorter time horizons on the design of multi-energy systems for single buildings, and tries to identify critical periods or time steps.

We find the operational control trajectory of a building energy system is not influenced by the horizon, while the design is. Storage technologies play a dominant role in the results, and are extensively used, while the storage sizing results from a complex interplay between charging/discharging speed, storage capacity and use of other technologies in the system.

*This chapter is related to the publication **[P2]***

3.1 Introduction

The integration of renewable energy to reduce the carbon footprint of building energy loads requires a careful sizing from the earliest design stages. To this end, the modeling of multi-energy systems (MES) using linear programming is gaining interests among practitioners [46, 82]. However these methods are slow as they require a solution over at least a full year of time series data.

To reduce the computation burden, exploiting the quasi-redundancy of building energy demands and solar irradiation [46, 53] allows the usage of a reduced number of pieces of information, from months to days. Using representative days selected via clustering algorithms has now become common practice [47].

The selection process for these representative periods is usually based on time series information, i.e. independently from the system itself. Some iterative frameworks were suggested to reduce this shortcoming [52, 53], e.g. to include extreme periods [51].

Although the results of all suggested methods are considered as is and the importance of each piece of information remains unknown [46].

This study performs and analyzes system design and operation results for different time series horizons (day to 3 months). Questions driving this study are:

- What influence does the time series horizon have on the sizing and operation of an energy system?
- Does the sizing with a reduced horizon allow identification of decisive periods or time steps?
- What are the consequences on computation time?

A methodology section introduces the framework, hypotheses and study cases. Then results are exposed and discussed. The conclusion summarizes our findings.

3.2 Methodology

3.2.1 Energy Hub modeling

In this work, the multi-energy systems are modeled as energy hubs. An energy hub (EH) is an interface between various facilities, sources and loads. It represents a framework to model the management (conversion, conditioning, storage) of multiple energy carriers [20]. This approach considers the interaction between elements of the energy system during the sizing procedure.

Mixed-integer linear programming (MILP) is used to represent and size the energy system. The objective function to minimize and governing set of equations are described by Evins et al. [29]. A key equation when considering different time horizons is the storage looping constraint (eq.(3.1)), forcing the storage state of charge (SOC) to be identical at the start ($t = 0$) and end ($t = T + 1$) of a period.

$$SOC(T + 1) = SOC(0) \tag{3.1}$$

We also highlight the capacity-related constraints, further analyzed in this study. The energy output ($p_{out,c}$) of a converter c cannot exceed the capacity of the device P_c (eq.(3.2)). Similarly, the SOC of a storage device s does not exceed its capacity P_s (eq.(3.3)). A storage ramping constraint limits the energy in/outflow $Q_{s,+/-}$ to a fraction λ of the capacity (eq.(3.4)).

$$0 \leq p_{out,c}(t) \leq P_c \ \forall c, t \tag{3.2}$$

$$0 \leq SOC_s(t) \leq P_s \ \forall s, t \tag{3.3}$$

$$0 \leq Q_{s,+/-}(t) \leq \lambda \times P_s \ \forall s, t \tag{3.4}$$

The total carbon emissions are also computed using carbon emission factors of grid electricity and gas. The ϵ-constraint [67] is used on the carbon emissions during the sizing step to explore optimal solutions at minimum cost, -50% CO_2 emissions and net-zero (NZ). Note that -50% CO_2 emission scenarios force CO_2 emissions to be *at least* reduced by 50% relatively to the minimum cost solution, though optimal solutions may result in even lower CO_2 emission levels.

3.2.2 The experiment

The experiment compares the sizing and operation of an EH for different time horizons: day, week, month, quarter. All are viewed as *estimates* and compared to the optimization using a year of data, considered as ground truth.

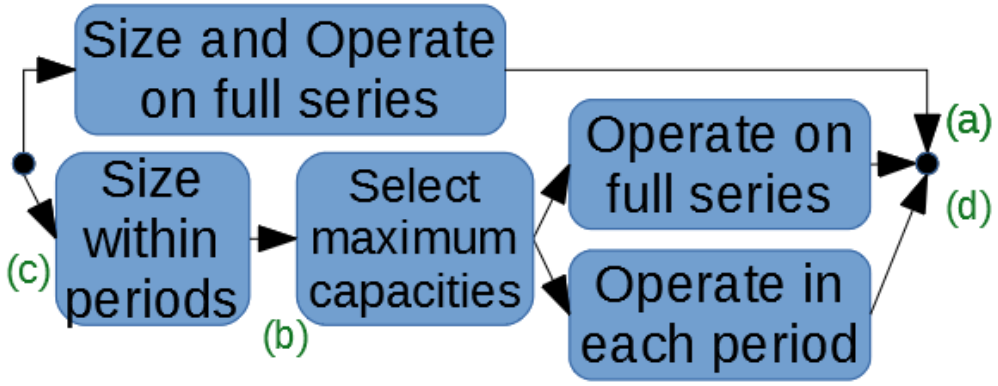


Figure 3.1: Methodology

The experiment consists of 3 steps: a sizing step in each period, the selection of capacities from these periods, and an operation step (Figure 3.1). For each horizon length, the timeline is subdivided into periods of one horizon. During the sizing step, optimal device capacities and operations are computed independently over each period. The resulting CO_2 emissions at minimum cost in each period are used to set the emission threshold of the -50% CO_2 scenarios. During the operation step, the highest capacities obtained over all periods (of equal length horizon) are then used and the resulting energy system is operated, both for one year of hourly data and independently in each shorter period length, estimating the cost and carbon emissions again. The operation in each period uses the same CO_2 emission threshold as during the sizing, while operations on the full time series use the sum of CO_2 emission obtained after sizing of all periods. In every case, the total cost is minimized.

The study first compares overall results (a). The sizing results from each periods is further analyzed to capture the decisive periods (b). To identify critical elements for the sizing, we localize the time steps at which equations (3.2-3.4) are binding (c). Finally, computation times are also compared (d).

3.2.3 Case Study

The study is conducted for 5 locations in different climate zones (Duluth MN, Seattle WA, Baltimore MD, Los Angeles CA, Houston TX), using data from Hourly

Load Profiles of Commercial and Residential Buildings dataset[1] [77]. In each location, 2 building types are considered (residential and small-office) across 4 scenarios (Table 3.1). Other modeling parameters were based on energy hub studies and manufacturer data. All codes and data are made available on a git repository[2].

The energy system to be sized is shown in Figure 3.2, and includes photovoltaic (PV) panels, a heat-pump (HP), Boiler, a hot water tank (HWT) and a battery. The PV efficiency is precomputed using the Python package Gsee[3] [81]. The EH framework uses the Besos[4] [83] package and the Gurobi solver [84].

Table 3.1: Price scenarios used.

#	Price ($\frac{\$}{kWh}$)			Carbon ($\frac{gCO_{2,eq}}{kWh}$)	
	Electricity	Gas	Feed-in	Electricity	Gas
1	0.1	0.02	0.05	200	400
2	0.33	0.02	0.05	600	400
3	0.1	0.1	0.05	200	400
4	0.33	0.1	0.05	600	400

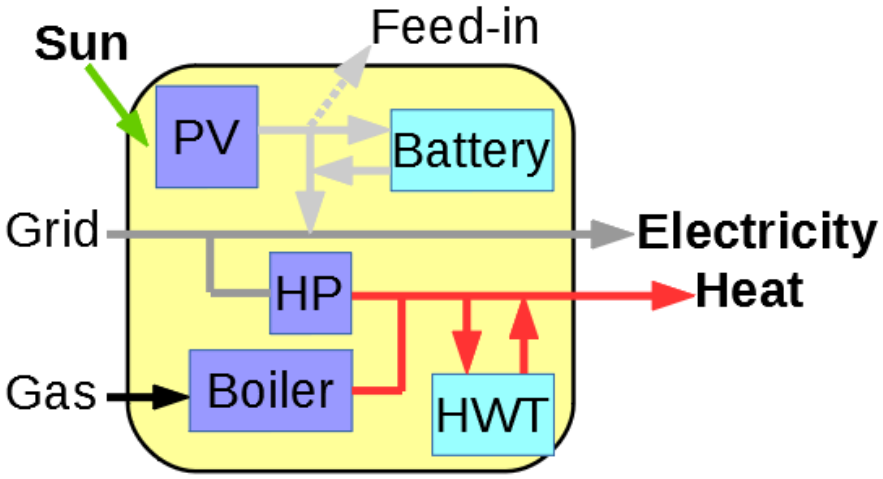
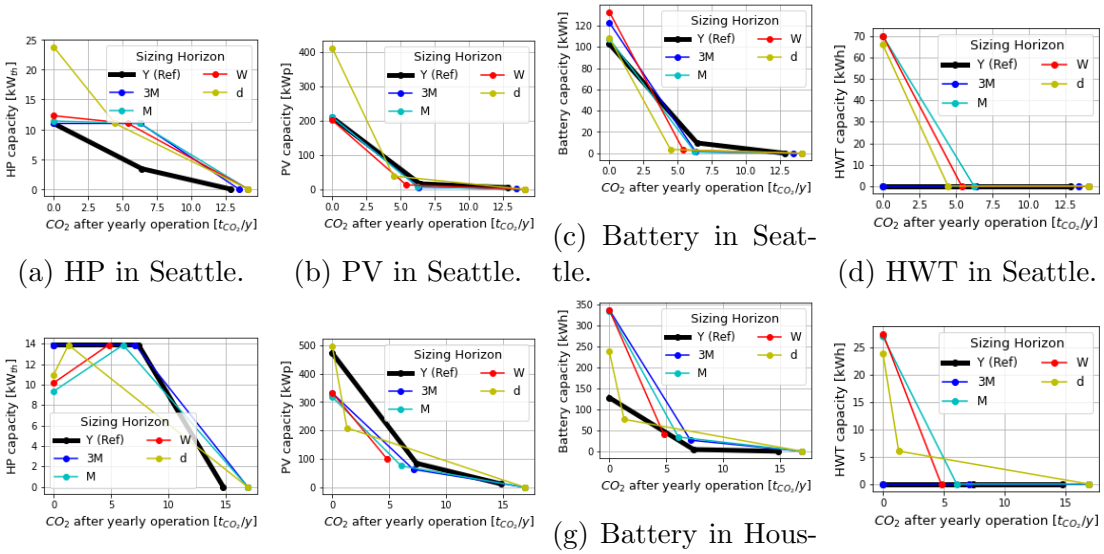


Figure 3.2: System design

[1] https://data.openei.org/submissions/153
[2] https://gitlab.com/fledee/horizon_ehub
[3] https://pypi.org/project/gsee/
[4] https://pypi.org/project/besos/



(a) HP in Seattle. (b) PV in Seattle. (c) Battery in Seattle. (d) HWT in Seattle.

(e) HP in Houston. (f) PV in Houston. (g) Battery in Houston. (h) HWT in Houston.

Figure 3.3: Renewable capacities optimal front for all carbon emission scenarios for Seattle (3.3a-3.3d) and Houston (3.3e-3.3h)

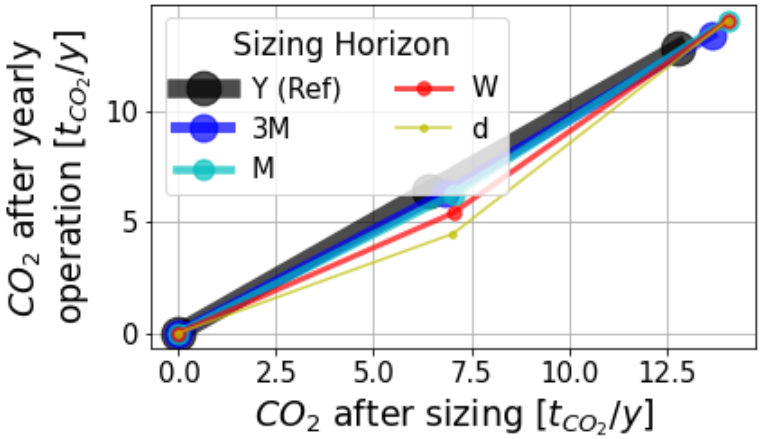
3.3 Results

3.3.1 Influence of horizons on the estimations

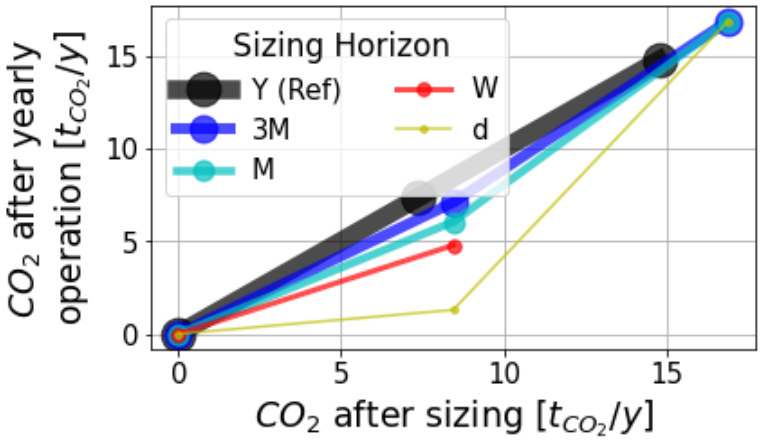
Figure 3.3 shows the optimal capacity-carbon fronts obtained after the sizing and operation steps in Seattle and Houston. Capacities are determined after the sizing step, CO_2 emission levels are obtained after operating the so sized system on one entire year.

The obtained fronts are similar in shape for every time horizon, except for the HWT (Figures 3.3d & 3.3h) and HP in Seattle (Figure 3.3a). PV and Battery show an exponential capacity increase with decreasing CO_2 limits for all horizons. This reveals that the importance of both technologies and their interplay can be correctly captured in specific shorter time spans. The HWT is usually only installed at net-zero and for horizons of a month or less. HP capacity profiles show a binary decision to either fully electrify the heating demand (-50% CO_2 and NZ) or discard the technology (minimum cost). At -50% CO_2 in Seattle (Figure 3.3a) the reference shows an intermediate HP capacity, this reveals that considering an entire year of data incentivizes a joint use of HP and Boiler. The presence of HWT at NZ also matches with a reduced need of HP capacity (Figure 3.3e for days to months) or an increased need (Figure 3.3a), as developed in the 3^{rd} results subsection.

Profiles may also differ in CO_2 emission levels. For the -50% CO_2 emissions scenarios, the shorter the horizon for the sizing, the lower the emissions. At minimum cost, shorter horizons increase the CO_2 emissions. Sizing in Houston (Figures 3.3e-3.3h) illustrates that shorter time horizons result in energy systems with higher PV and Battery capacities than the yearly reference. Optimal operation of these systems result in significantly lower CO_2 emissions than the reference case.



(a) CO_2 emissions in Seattle



(b) CO_2 emissions in Houston

Figure 3.4: Carbon emission right after sizing vs. after operation of the so sized system with a yearly horizon.

Figure 3.4 explains this phenomenon: the expected CO_2 emissions right after the sizing step are systematically higher with sub-yearly horizons than with the yearly reference. Conversely after operating so sized systems, the emission level become smaller than the reference. The shorter the horizon, the greater the reduction. Using shorter time horizons results either in a larger electrification of the heating demand (Figure 3.3a) or a higher use of PV and battery (Figures 3.3f, 3.3g & 3.5) to meet the carbon reduction constrain in periods with less resources.

These greater capacities become available for all periods during the operation step, and their usage results in CO_2 emissions lower than the reference case. This is highlighted in Figure 3.5, where a storage is only installed for a few independent days but used extensively during the operating phase.

The horizon of the operation step does not influence our results, regardless of the location, scenario or building type. Figure 3.5 shows a battery system operated with a year and a day of horizon, where the same daily cycles are observed. The strong daily cycles and low variability of profiles induced by simulated single building archetypes can explain this observation, as well as the use of constant energy prices.

3.3.2 Informative value of sizing results

Figure 3.6 illustrates the diversity of sizing results obtained throughout all study cases, with a ±5% margin regarding the results from the full optimization. A margin of 5% is chosen, similarly to the optimality gap in the optimization problem. Some cases allow a clear identification of decisive periods, narrowing down the scope as shorter horizons are considered (e.g. NZ in Figure 3.6). Other show systematic divergences with the full optimization (e.g. min. cost and -50% CO_2 in Figure 3.6c) or a mix of periods leading to accurate sizing, oversizing and undersizing (e.g. Figures 3.6a-3.6b at -50% CO_2).

Regardless of the horizon and carbon scenario, the optimization in most periods lead to a capacity underestimation (Figure 3.6). Notice that some periods of shorter horizon do yield accurate results. There generally exists periods for which the Heat-Pump is accurately sized, regardless of the horizon, location or constraint on carbon reduction (Figure 3.6a). We also find cases where using a year horizon systematically leads to different sizings (Figure 3.6c at -50% CO_2), or where shorter horizons lead to using a technology that otherwise wouldn't have been recommended (Figure 3.6c at min. cost).

The existence of at least one of these periods is not guaranteed, while being more likely with shorter horizons (Table 3.2). Highest probabilities in all locations are found with the heat-pump sizing. The existence of such periods is also likely for PV and Battery at NZ, when the system mostly rely on them. With a carbon reduction constraint, the likelihood that periods yield a sizing similar to the yearly horizon one increases with shorter periods. At min. cost, the horizon has little effect on the sizing, as the system rely on grid and gas imports in most cases.

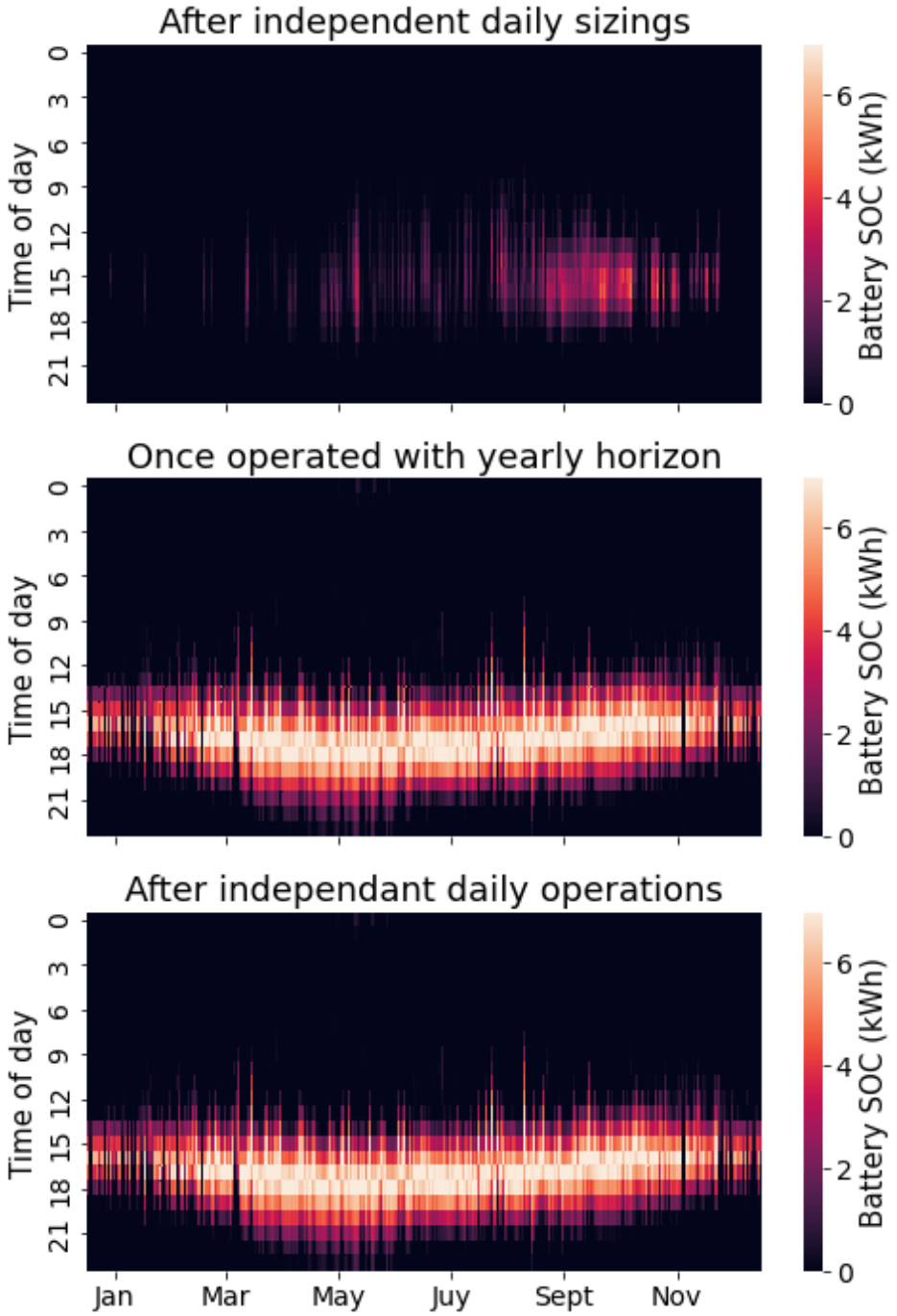
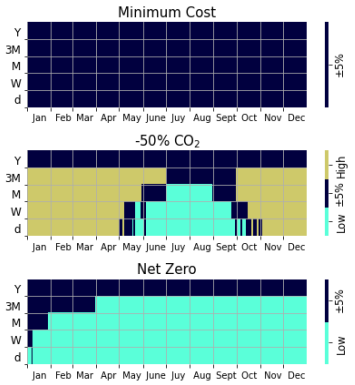


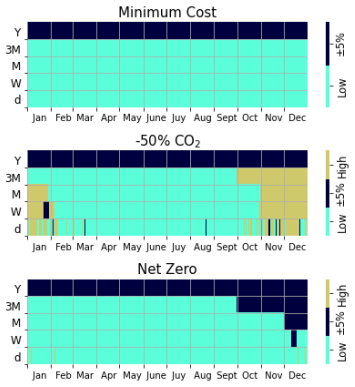
Figure 3.5: Example of the expected usage of a battery right after the sizing step (top), then after operating the so sized system with horizons of a year (middle) and a day (bottom) in Baltimore (MD).

3.3.3 Decisive elements

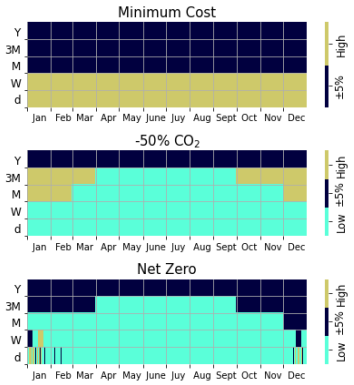
Figure 3.7 shows an example of a heat-pump usage *during the sizing step* and the moments when the devices operate at their maximum, i.e when the capacity constraint (3.2) is binding. In almost all cases, the maximum capacity constraint



(a) An HP sizing in Los Angeles



(b) A PV sizing in Seattle



(c) A battery sizing in Baltimore

Figure 3.6: Independent periods in which the sizing issues undersized capacities (Low), oversized (High) or similar to the full sizing (± 5%).

Table 3.2: Ratio of computed cases for which at least one subperiod of the year issued an optimal systems sizing within 5% of reference.

		Duluth			Seattle			Baltimore			LosAngeles			Houston		
		PV	HP	Bat	PV	HP	Bat	PV	HP	Bat	PV	HP	Bat	PV	HP	Bat
Min Cost	3M	.50	.88	.50	.62	1	.50	.50	.88	.50	0	.88	.50	0	.62	.50
	M	.50	.62	.50	.62	.88	.62	.50	.75	.50	0	.75	.50	0	.62	.50
	W	.50	.50	0	.62	.62	.25	.50	.62	0	0	.75	0	0	.62	0
	d	.50	.50	.25	.62	.62	.62	.50	.62	.25	0	.75	.25	0	.62	.25
$-50\%CO_2$	3M	.38	.50	.25	.38	.50	.50	.38	.75	.25	.12	1	.12	0	1	0
	M	.25	.88	.25	.62	1	.12	.12	.75	.38	0	1	.38	.12	1	0
	W	.50	.75	.38	.62	.88	.38	.25	.88	.50	.38	.75	.12	.62	1	0
	d	.62	.88	.62	.75	1	.50	.75	.88	.75	.75	1	.75	1	1	.25
Net Zero	3M	.50	.50	1	1	1	1	1	.50	.50	1	1	1	.50	1	1
	M	.50	0	1	1	.50	1	1	.50	1	0	.50	1	.50	.50	1
	W	1	.50	1	1	.50	1	1	1	1	0	.50	1	.50	.50	1
	d	1	.50	1	1	.50	1	1	1	1	1	.50	1	1	.50	1

is only binding for one unique time step in each period. Daily profiles show the constraint is mostly binding between 5am and 7am. Moreover the maximum of these morning peaks necessitate the same heat-pump capacity with daily horizons and with the yearly reference (red margin in Figure 3.7).

Figure 3.7 shows a case where HWT is installed and further highlight the interplay between technologies. Daily profiles show a peak of heat-pump usage at 1pm resulting in an additional need of 7kW$_{th}$ compared to the yearly reference. This 1pm peak comes from the need to rapidly fill the HWT using carbon-free electricity from the PV in winter days.

We also observe a linear relation between the heat-pump capacity designed in the binding constraint and the peak of heating load in that period (Figure 3.8). Parallel usage of a boiler results in lower heat-pump capacities (e.g. min. Cost

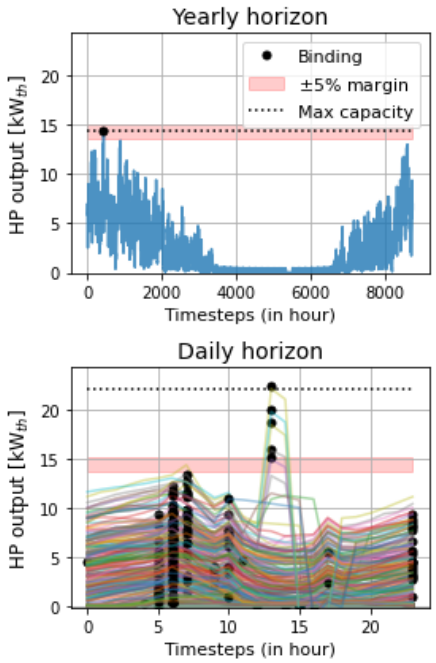


Figure 3.7: Location of binding constraints for HP capacity sizing in Baltimore for net-zero residential building.

and -50% CO_2 in Figure 3.8), while the use of a thermal storage may lead to higher heat-pump capacities (e.g. NZ in Figure 3.8) when the thermal storage must be loaded using clean solar power while still ensuring the heating demand (e.g. Daily horizon in Figure 3.7).

Our observations are different for the battery. Figure 3.9 shows the constraints related to storage capacities (eq.(3.3) & (3.4)) are binding multiple times in each period. The state of charge regularly reaches its maximum (Figure 3.9a). The ramping constraints are binding almost every day, regardless of the horizon, location, scenario or building type (Figure 3.9b).

Therefore, storage sizing is found to rely on the entirety of the information, including loads and other decisions taken within the optimization process. Moreover, no statistical relation between storage capacity sizing and features from the heating load, the electric load or irradiation was clearly identified.

3.3.4 Computation time

Figure 3.10 shows computation times required to transform code instructions into solvable problems ("*Create*") and to solve it at both the sizing and operation steps ("*Ope*"). The sizing part shows computational gains as soon as the optimization

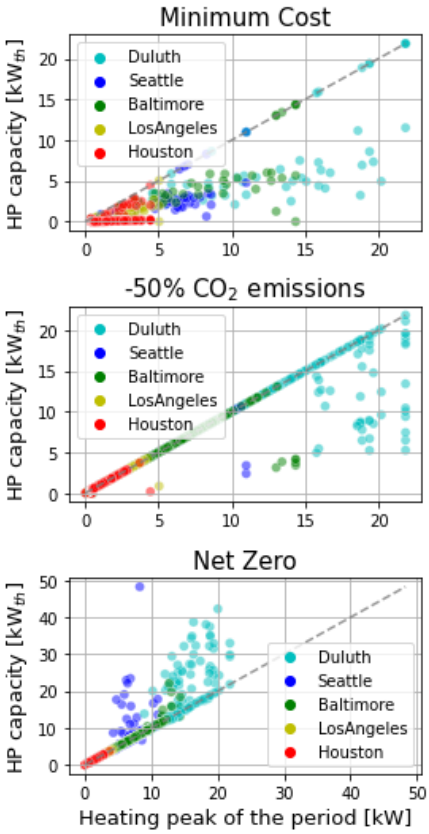


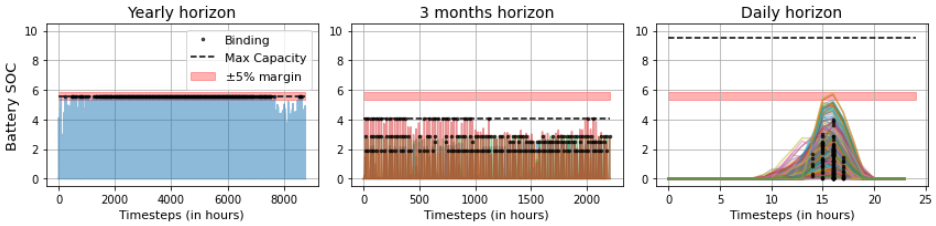
Figure 3.8: Relation between peak heating load and HP capacity in residential buildings.

is performed independently in periods (Figure 3.10). Computing 52 independent weeks show greater saving in computational resources while 365 days takes up to 60% of the reference time. Adding the operation, the daily horizon shows longest computation times compared to the full time series.

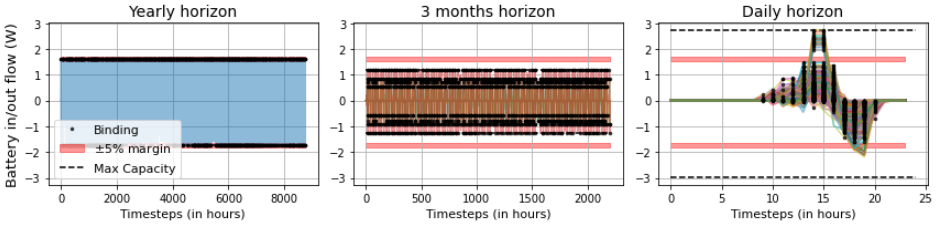
The computation time varies with the building type and systems for small office buildings take more time to compute in the present case. Figure 3.10 also shows higher computation times for the operation on the full time series, although most of the time is spent to create the problem, while the total solving time is larger for operations in time series subperiods.

3.4 Discussion

The same system is sized and operated for different time horizons. To ensure the validity of observations, the experiment is performed in different climate zones, for 2 building types, under different financial and carbon reduction scenarios. Still the procedure is not exhaustive and factors such as the use of simulated loads from



(a) Binding constraints for the storage level of an office building in Los Angeles



(b) Binding constraints for the storage ramping of an office building in Los Angeles

Figure 3.9: Decisive moments for the battery sizing: an example in Los Angeles at -50% CO_2 emissions. Both SOC (3.9a) and in/outflow limits (3.9b) occur daily. The maximum SOC at a daily horizon matches with the capacity designed with a yearly horizon, though a need to load and discharge faster results to an oversizing.

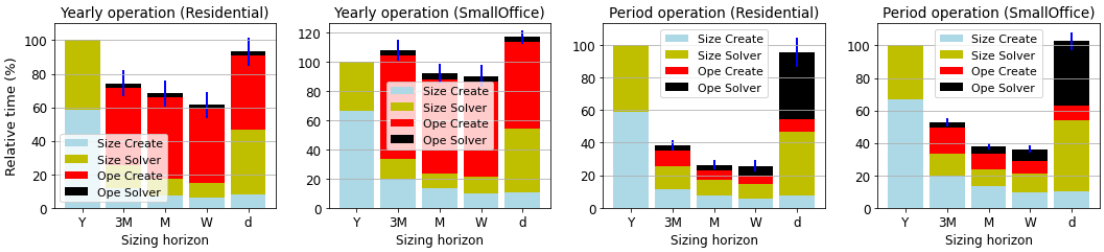


Figure 3.10: Average relative computation time per building type and operation horizon.

archetypes, the use of constant energy prices and efficiencies factors may influence the observations.

Considering the storage looping constraint (eq.(3.1)), the methodology suggests a capacity sizing adapted if a specific period (quarter, month, week or day) was repeated indefinitely. A subdivision of the yearly time series holds higher chances to obtain some periods holding accurate results with shorter horizons (Table 3.2). However shorter horizons also induce higher sensitivity to extreme events, increasing the range of suggested solutions. Therefore risk of oversizing is increased with the current methodology selecting the maximum capacity over all periods. Moreover the yearly cycle and long-term interactions play a dominant role in some cases, and nothing guarantees a-priori the possibility to capture these

and obtain accurate sizing using shorter horizons (Figure 3.6).

The methodology also highlights the importance of storage systems in the decarbonization of building energy demand. Although sizing the battery system independently in different periods and with different horizons yield a large diversity of suggested capacities, an installed battery is always used at its maximum during the operation step (Figure 3.5). This contributes in a reduction of expected carbon emissions (Figure 3.4), as well as a different predicted Heat-Pump and PV production.

Heat storage, mostly suggested for cold places at NZ in our experiment, also perturbs the sizing process, particularly the HP (Table 3.2). For residential buildings, HWT is designed as additional buffer to absorb peak of PV production, forcing the HP to be oversized to produce both to load the storage and meet the heat demand (Figures 3.7 and 3.8).

Even during the sizing stage, the storage devices are designed to support the entirety of the system (Figure 3.9), by being shaped to be used at their maximum at all times. This also influences the sizing of all system components, and highlights the main reason for the present methodology to fail at accurately address an energy hub sizing problem.

On the one hand, the storage system is the difficult piece in models reducing the complexity of energy hub problems [14, 82]. On the other hand, the suggested methodology reduces computation time of the sizing step alone. Thus sizing results show a good potential for being used as features in the process of selecting representative days into standard complexity reduction methods.

3.5 Conclusion

This study investigates the effect of time horizon on the energy system design via optimization for single buildings. The sizing is performed using portions of the time series data (loads and irradiation) of different length (a day to 3 months).

We found no guarantee that any of the results obtained with shorter horizons are similar to those obtained with a full optimization. Though general trends remain similar. Conversely the system operation is similar regardless of the horizon, once the design and capacities are fixed.

Battery systems drastically influence the sizing results. Storage devices always cycle daily to support the single building loads. At minimum cost or -50% CO_2,

both ramping and level of charge are exploited at their maximum on a daily basis, regardless of the considered horizon. At net-zero, the ramping (charging and discharging power) is the decisive element to absorb peaks of solar production.

The interplay of the different technologies in the considered systems does not allow to identify relevant shorter periods to approximate the overall sizing. Thus considering a full year remains necessary to accurately address the sizing. Though the diversity of results obtained with shorter horizons give valuable information on choices made by the solver. Exploiting this information in data-driven processes to select representative periods is of future work's interest.

Chapter 4

A comparison of 4th and 5th generation thermal networks with energy hub

Abstract

State-of-the-art thermal networks are key to address decarbonization of the heating and cooling in buildings. Energy Hub considers synergy between elements and allows rapid comparison between district energy systems incorporating 4th generation (4G) and 5th generation (5G) thermal networks at early-stage designs. To better understand and quantify the differences between systems based on 4G and 5G, a mixed-integer linear model distinguishing specific features of both technologies is implemented. We find systems relying on 5G to perform environmentally and financially better than with 4G generation over a wide set of scenarios. This is due to the warm/cold coupling characterizing 5G technologies. Systems based on 5G can also reduce its carbon emissions more than those with 4G. However, performances with both technologies appear sensitive to the topology and location of central energy station.

*This chapter is related to the publication **[P3]***

4.1 Introduction

The development of distributed multi-energy systems (MES) has been identified as a key driver in energy transition [16, 85]. Most recent developments of district thermal network (DTN) technologies place thermal networks at the heart of the MES concept, by increasing the integrability of renewable energy sources and exploiting low temperature waste-heat.

Currently, two technologies of district thermal networks are considered state-of-the-art [86]: 4th generation (4G) district heating [55] and 5th generation (5G) district heating and cooling [42].

4G networks are centralized systems in line with the three previous generations [43]. Their low operating temperature (50-70^{o}C) allows thermal sources such as industrial waste heat, renewable energy and centralized heat-pumps to compete with fossil fuels. They are commonly designed to supply heat, while a cooling service may be planned independently [54] or in parallel [87].

5G networks consist of a warm and a cold pipe with temperatures close to ground temperature (5-30^{o}C) [42, 43, 56, 86, 88]. In heating mode, heat-pumps (HP) extract heat from the network, and the cooled water is returned to the cold pipe. Water temperature in the building is raised using the extracted heat [56]. The inverse principle is used to deliver cooling, injecting heated water back to the warm pipe. This strong coupling between warm and cold sides induced by the usage of decentralized HP is the essence of 5G networks. This supports electrification, allows the integration of decentralized renewable energy (RnE) and ultra-low temperature waste-heat [86]. Each customer becomes a prosumer [43].

As more technologies can be used to generate thermal energy, the complexity of designing and operating district thermal systems increases [20]. Advanced modeling, e.g. following the "Energy Hub" (EHub) concept [20], were found relevant from early design stages to identify potential synergies between diverse technologies and sources to be exploited [16, 20, 89]. The EHub concept gathers production, storage, conversion and consumption of different energy carriers within one model, in a unified and interacting fashion [19]. Such model is commonly described through Mixed Integer Linear Programming (MILP) [16] to determine financially or environmentally optimal integration of technologies into a MES [29, 33].

Some studies compare 4G and 5G networks, as detailed in the literature section. Those not using the EHub concept omit the synergies between technologies and sources [90–95]. Two of them use the EHub concept [87, 96], but fundamental aspects of the 5G network are missing [56] and place the technology at a critical

disadvantage. Moreover, current literature is lacking a deeper investigation of the influence of model characteristics on the results. The present study addresses the following questions: (1) What expected differences in design, environmental and financial performances of integrated energy systems result from the consideration of 4G or 5G DTN technology at early design stages? (2) How do key modeling parameters influence the design and performances?

Related literature is presented in Section 4.2. Section 4.3 describes the full model. Section 4.4 covers data, control assumptions, tested network layouts and scenario parameters. Section 4.5 discusses the study limits. A detailed comparison of base scenarios is presented in Section 4.6.1. Sensitivity to the financial and environmental strategy, use of local renewable electricity, and geographical aspect of the model are presented in Sections 4.6.2, 4.6.3 and 4.6.4 respectively. Finally, implications and future perspectives are discussed in Section 4.7.

4.2 Literature & Scope

4.2.1 Modeling 4G DTN with EHub

Numerous EHub based methodologies have been suggested in the literature to find optimal topologies and technological synergies for 4G networks.

Some methodologies exclusively focus on the topology, with little regards for the technologies to supply energy or the operation. The selection of clients to connect via the DTN is of prime interest and was explored using peak demand [97], including the location and thermal power for central plants [98] and in coupling with the topology of natural gas and power network [99]. The latter concludes on the importance to include network losses in the modeling.

Some methodologies exclusively focus on the central station of 4G networks. The selection, sizing and operation of sets of technologies to ensure the supply is addressed, while geographical spread and selection of customers is excluded. The synergies between dispatchable thermal generation units are evaluated by Kuriyan et al. [100], including neither RnE nor storage. Storing capabilities of the network by variation of the network temperature are assessed in [101], while the benefit of boreholes [102], hydrogen storage [103,104] and a mix long and short-term storage technologies [105] were also studied. The sector coupling is also investigated to evaluate the potential of DTN integration with local and utility power grids [85], the need for upgrades in electric and heating networks [9] and the benefits of

coupling heating and cooling generation in the central station of a 4GDTN [106].

Methods to compare centralized vs. decentralized while investigating synergies between technologies are also suggested. Unternährer et al. [107] select and size dispatchable units of a central station with an EHub model, after a separate optimization of the network topology. Morvaj et al. [8] address both topology and size optimization within one EHub formulation, including storage, RnE and hourly operation of the whole system. The operational aspect was validated against simulations in [108] and used in [7] to address multi-scale planning. Gas network, thermal network and transportation planning are coupled within an EHub by Pantaleo et al. [109, 110] to study the integration of biomass in a city.

4.2.2 Modeling 5G networks

Linear optimization and the EHub concept have been little used for design and operation of 5G systems. Prasanna et al. [111] validate MILP and EHub concepts against data of a real system including borehole. The model is further used to study the potential of additional decentralized storage systems on the self-sufficiency of the whole system. Wirtz et al. [56] suggest a MILP formulation to size and operate all centralized and decentralized supply and storage around an already-designed 5G network for a university campus. The pre-sizing of the network prevents investigating centralized-decentralized trade-offs, the effect of long-term storage and carbon transition.

Linear Programming has been used for modeling 5G networks to study optimal operation strategies as well [111–114]. Although MILP was proven sufficiently accurate for early-stage studies, it is widely considered that fundamentally non-linear characteristics of thermodynamics and fluid mechanics [18] make linear solutions non suitable. Thus, a vast majority of studies rely on simulation models such as TRNSYS [18, 115] or Modelica [116, 117]. However, these only assume a predefined layout and supply set [88, 118–122], compare predefined configurations [123, 124], conceive ad-hoc iterative processes [95, 125, 126] or integrate the simulation into evolutionary algorithms (e.g. PSO [90, 127]). In their review, Brown et al. [18] acknowledge the need for a unified framework encompassing all aspects of the 5G network. Taylor et al. [124] also conclude simulation-based approaches do limit the exploration of trade-offs, synergies between technologies and usage of storage, which justifies the usage of MILP and EHub for early-stage designs of 5G networks.

4.2.3 Comparison of DTN technologies

A recurring question to address at very early-stage design is the choice of network technology. Transiting from centralized generations 1 to 4, the main concern is the reduction of supply temperature [128], allowing to integrate new energy sources [129–131]. Interacting with a decentralized 5GDTN is fundamentally different [43], and requires two closely related loops (warm and cold). Low temperature sources (waste heat, ground, etc.) and specific dispatchable units regulate the network (central station, borehole, long-term storage, etc.). Therefore, models encompassing specificities of each DTN technologies are needed for comparisons.

Gudmundsson et al. [93] use simulation for an economic comparison of 4G and 5G. They conclude the cheaper operation of 5G does not compensate for the high investment costs. Their method is only applied to heating service for residential districts and does not consider energy generation. Gross et al. [92] use simulation in a district with diversified loads to show the impact of active prosumers on a 5GDTN. They validate their flow model against data. Their method focuses on energy savings and no diversity of energy generation is considered. Jebamalai et al. [94] evaluate financial gains of specific topologies of centralized networks over a 5G network. They find 5G is only preferable if industrial waste heat is available for free. Their method does not consider any choice or design of the supply, and long-term storage is only considered for the centralized networks. Barely any cooling is included in their case study. Calise et al. [91] compare the DTN technologies with TRNSYS models for a residential district. Heating and cooling do not occur simultaneously, which allows the same infrastructure to supply heat in winter and cold in summer with the 4G network. Both networks only rely on an already designed central ground source heat pump, partially fed with 3MWp of solar panels. They conclude on the economic and environmental superiority of mature 4G technology for purely residential districts with little load sharing potential.

Brumana et al. [90] use PSO to design the most cost-effective PV-Battery-Chiller system to support both DTN technologies. They find centralized networks financially advantageous and requiring less generation capacity. Their methodology neglects the heat-cold coupling of the 5GDTN and is only applied to residential buildings. Millar et al. [96] compare financial performances of the two DTN technologies with an EHub. For 4G and 5G, a single pipe at constant temperature connects buildings of different nature. Storage, HP (5G) and exchangers (4G) are sized. They find that with storage, 5G is cheaper than 4G. However, their 5G model neglects the fundamental heat-cold coupling, and only financial performances are assessed. Nérot et al. [87] compare multiple DTN technologies for

different objectives and climates with an EHub. The DTN and generators are designed and operated to supply heating and cooling. They find 4G is the best trade-off between financial, environmental and exergy efficiency. 5G performs consistently worse than 4G and is only comparable to buildings operated individually. Their methodology uses a single ambient loop for 5G, with little difference with other generations in the way the network operates. This also neglects the heat-cold coupling. The method is used on a fully residential district, where seasonal storage and RnE is only available for 4G. Zhang et al. [95] combine optimization and simulation for an economic comparison of 4G and 5G. Their method sizes small thermal storages and is used for a mixed set of buildings under multiple electricity price, weather and retrofit scenarios. They find that 5G is not economically viable compared to 4G, unless the cooling share is above 27% or unless the price of electricity drops drastically. Their method does not include selection and sizing of energy generation, does not consider local PV and the analysis does not include carbon emissions.

4.2.4 Contributions

In this paper, an EHub model is used to compare the economic and environmental performances of MES relying on 4G and 5G DTN in various contexts. The model includes the selection of customers to connect to the network, i.e. the trade-off between connected and stand-alone thermal supply. The comparison is conducted for 3 climate zones in the USA and across a variety of financial and technical scenarios. The main contributions of this paper are:

- comparing overall expected performances of a MES relying on 4G and 5G technologies including all the features mentioned in Table 4.1.
- investigating the sensitivity of the comparison to strategic model parameters.

4.3 Methodology

The energy system model is split into nodes (hosting the energy loads, imports, exports and transformations) and a network layout (handling energy transfer between nodes and transmission losses). All model variables, including capacities, operation and energy purchase, are determined collectively via optimization (cost and emissions minimizations) using Gurobi [84]. This section presents the optimization objectives, then the linear formulations for the network and node models.

Table 4.1: Studies comparing 4G and 5G DTN technologies via modeling

	Brumana et al., 2022 [90]	Calise et al., 2023 [91]	Gross et al., 2021 [92]	Gudmundsson et al., 2022 [93]	Jebamalai et al., 2022 [94]	Millar et al., 2021 [96]	Nérot et al., 2023 [87]	Zhang et al., 2022 [95]	Current study
Heat & Cold	✓	✓			✓	✓	✓	✓	✓
H&C coupling		✓						✓	✓
min \$ & CO_2							✓		✓
Long storage					✓*	✓	✓*		✓
Network	Σ	Σ	Σ	Σ	Σ		H	Σ	H
Load diversity			✓		✓	✓		✓	✓
Supply design	E					H		S	H
With RnE	✓	✓	✓					✓	✓

(*) Not available for 5G
(Σ) Simulation; (H) Energy Hub MILP
(E) Evolutionary algorithm; (S) Scenario;

4.3.1 Objective function of the optimization

The cost J^a_{tot} in eq.(4.1) combines OPEX (operational expenditure) and annualized CAPEX (capital expenditure). The CAPEX considers the purchase and installation of the whole system considering prices C and installed capacities P^{nom}, and is annualized (eq.(4.2)). The OPEX includes energy bills J_{fuel}, carbon taxes J_{CO_2} and maintenance costs $J_{O\&M}$ and is not annualized. Bills in eq.(4.3) and emissions in eq.(4.4) consider net purchases of fuel and power $\dot{Q}$, prices, and carbon intensities G for the latter.

$$J^a_{tot} = J^a_{CAPEX} + J_{fuel} + J_{CO_2} + J_{O\&M} \tag{4.1}$$

$$J^a_{CAPEX} = \sum_{\tau \in \mathcal{T}} NPV_\tau (C^{fix}_\tau + P^{nom}_\tau C^{lin}_\tau) \tag{4.2}$$

$$J_{fuel} = \sum_{f \in \mathcal{F}} C_f^{fuel} \sum_{n \in \mathcal{N}} (\dot{Q}_{n,\sigma_f}^{\mathcal{N},in} - \dot{Q}_{n,\sigma_f}^{\mathcal{N},out}) \tag{4.3}$$

$$J_{CO_2} = C^{CO_2} \sum_{f \in \mathcal{F}} G_f \sum_{n \in \mathcal{N}} (\dot{Q}_{n,\sigma_f}^{\mathcal{N},in} - \dot{Q}_{n,\sigma_f}^{\mathcal{N},out}) \tag{4.4}$$

A carbon-cost optimal front is computed using the epsilon-constraint method [67] to simulate a forced decarbonization.

$$\sum_{f \in \mathcal{F}} G_f \sum_{n \in \mathcal{N}} (\dot{Q}_{n,\sigma_f}^{\mathcal{N},in} - \dot{Q}_{n,\sigma_f}^{\mathcal{N},out}) \leq X_{CO_2} \tag{4.5}$$

4.3.2 Network model

The model is tailored for very early design explorations. For the model to remain linear and tractable [113], a purely energy-based approach is considered, inspired by [8]. The model assumes a predefined and homogeneous temperature difference ΔT, and assumes water massflow $\dot{m}$ to be only influenced by pressure control [56]. The temperature difference is the difference between warm and cold pipes in the 5G network and the difference between supply and return in the 4G models. This results in a direct link between transferred thermal power $\dot{Q}$ and mass flow control (eq.(4.6)), allowing to only consider the transferred power as model variable. The modeling of thermal and circulation losses is revised consequently.

$$\dot{Q} = \dot{m}\ c_p\ \Delta T \tag{4.6}$$

4.3.2.1 Interface between nodes and network

The network can transfer energy between nodes. A node may take ($\dot{Q}^{\mathcal{N},in}$) or feed energy ($\dot{Q}^{\mathcal{N},out}$) directly to the grid. This energy may reach ($\dot{Q}^{\mathcal{L},in}$) or leave ($\dot{Q}^{\mathcal{L},out}$) the node proximity by any connected link $l \in \mathcal{L}$. Each link may transfer energy in one (4G) or both (5G) directions. Thermal losses $\dot{q}^{th}$ are unrelated to the energy flow in links, while the circulation loss $\dot{q}^{el}$ is positive regardless of the direction of the flow in the pipe. Thus, the interface between each node n and the

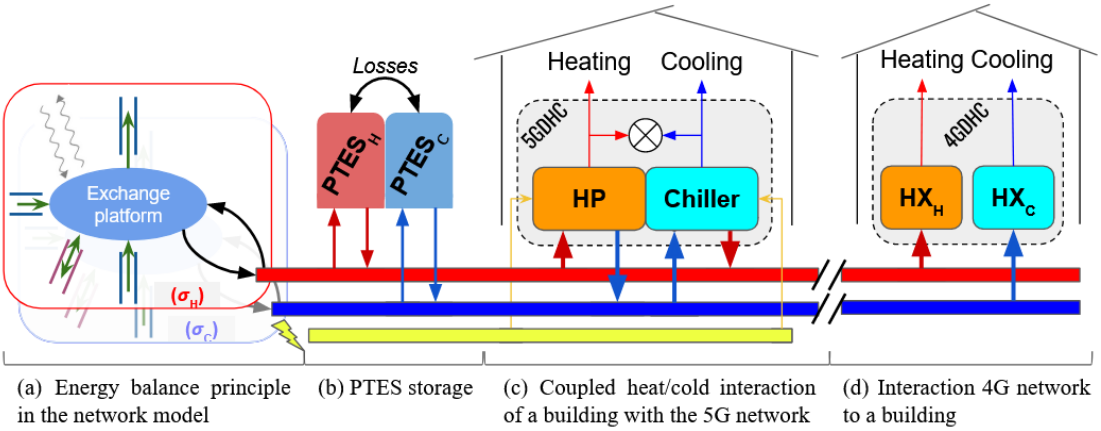


Figure 4.1: Visual principles of the models

network for stream σ at all time t is represented via eq.(4.7) and Figure 4.1a.

$$\dot{Q}^{\mathcal{N},out}_{n,\sigma,t} + \sum_{l \in \mathcal{L}(n)} \dot{Q}^{\mathcal{L},out}_{l,\sigma,t} = \dot{Q}^{\mathcal{N},in}_{n,\sigma,t} + \sum_{l \in \mathcal{L}(n)} \dot{Q}^{\mathcal{L},in}_{l,\sigma,t} + \sum_{l \in \mathcal{L}(n)} \dot{q}^{th}_{l,\sigma,t} \tag{4.7}$$

4.3.2.2 Thermal losses in links

5-15% of the transferred energy is dissipated in the ground due to the temperature difference with the network water [132]. These losses follow eq.(4.8) when the pipe l is installed (binary variable $\delta_l = 1$), even if no energy is transferred (unlike most related literature). Polyethylene pipes are uninsulated for 5GDTN ($\lambda = 0.4$ W/m.K) [56, 133] and insulated for the 4GDTN ($\lambda = 0.04$ W/m.K) [134]. The term $\dot{q}^{flex}_{l,t}$ refers to an additional thermal loss related to the thermal inertia of a 5G network (see eq.(4.21)).

$$\dot{q}^{th}_{l,\sigma,t} = L_l \, \delta_l \, \frac{2\pi\lambda}{ln(1 + \frac{2\epsilon}{D})}(T_w - T_s) + \dot{q}^{flex}_{l,\sigma,t} \tag{4.8}$$

4.3.2.3 Circulation losses in links

The formulation for the circulation losses in eq.(4.9) is adapted to the current energy transfer model and linearized. Since most related studies neglect pumping requirements, no equivalent formulation was found.

$$\dot{q}^{el}_{l,\sigma,t} = \dot{Q}_{l,\sigma,t} \times \left(\frac{8 f_D}{(\rho\pi)^2} \frac{L_l}{D_l^5} \frac{1}{c_p \ \Delta T \ \eta_{lin}} \right) \tag{4.9}$$

4.3.2.4 Flows in water pipes

If the network contains cycles[1], unrealistic loops of energy circulation can bias the model or be exploited to waste energy (e.g. excess of solar renewable) [9]. Thus, an additional variable $\delta p_{n,\sigma}$ is added for each node n and transferred stream σ, resulting in eq.(4.10) [8, 135]. It can be assimilated to a pressure level, forcing the flow to be unidirectional in pipes at any time t and to never close a circuit.

$$\dot{Q}^{\mathcal{L}}_{l,\sigma,t} = \frac{\tilde{P}^{nom}_l}{\Delta p_x \ L_l} (\delta p_{n_{in},\sigma,t} - \delta p_{n_{out},\sigma,t}) \tag{4.10}$$

4.3.3 Nodes

Nodes are where energy is purchased, consumed, produced, transformed and fed in, following the energy hub concept [20, 29].

4.3.3.1 Stream balance

The node is divided into energy streams, each one must be balanced between generation, consumption, storage interactions, imports and exports.

$$\dot{Q}^{load}_{n,\sigma,t} = \sum_{\tau_{cv} \in \mathcal{T}_{cv}} (H_{\tau_{cv},\sigma} \ \dot{Q}^{\mathcal{T},in}_{n,\tau_{cv},t}) + \dot{Q}^{\mathcal{N},in}_{n,\sigma,t} - \dot{Q}^{\mathcal{N},out}_{n,\sigma,t} + \sum_{\tau_s \ on \ \sigma} (\eta^{dch}_{\tau_s} \dot{Q}^{dch}_{n,\sigma,\tau_s,t} - \eta^{ch}_{\tau_s} \dot{Q}^{ch}_{n,\sigma,\tau_s,t}) \tag{4.11}$$

4.3.3.2 Converters

A converter τ_{cv} transforms power from one or multiple streams into one or multiple others. All converters are considered linear. This is common for most technologies

[1] https://en.wikipedia.org/wiki/Cycle_(graph_theory)

(e.g. Boilers, HP, PV, ST) [56, 61, 62, 67, 79, 80]. CHP and Absorption Chillers (AChill) [61, 62] are also assumed linear, as a minimum partial load does not influence the system design and partial load efficiency makes the model intractable [80].

A matrix-based formulation describes the operation (eq.(4.12)). The unitless coefficient $H_{(\tau_{cv},\sigma)}$ binds each converter τ_{cv} to each stream σ. For HP, solar thermal (ST) collectors and photovoltaic (PV) panels, this coefficient is precomputed for every time step t, it is constant for other converters (c.f. Appendix B). It is negative for inputs and positive for outputs. A reference input stream σ_{ref} is used to determine other input and output powers of each converter τ_{cv} (eq.(4.11)). Roof technologies (PV and ST) cannot be curtailed and must convert all received irradiation; a transformed factor $\xi^{roof}_{n,\tau_{cv},t}$ (in kW/kWp) is precomputed to account for direct and diffuse irradiation, air temperature, tilt and orientation [81].

$$\begin{cases} 0 \leq H_{(\tau_{cv},\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{n,\tau_{cv},t} \leq P^{nom}_{n,\tau_{cv}}, \\ \qquad\qquad\qquad \text{if } \tau_{cv} \text{ is dispatchable} \\ H_{(\tau_{cv},\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{n,\tau_{cv},t} = \xi^{sun}_{n,\tau_{cv},t}\ P^{nom}_{n,\tau_{cv}}, \\ \qquad\qquad\qquad \text{if } \tau_{cv} \in \{PV, ST\} \end{cases} \tag{4.12}$$

Every node has a limited rooftop area A^{max} to be shared between PV and ST. This further constrains the capacity of solar techs to be installed, with an assumption of $\alpha^{R}_{ST} = 2m^2/kWp_{th}$ for ST and $\alpha^{R}_{PV} = 8m^2/kWp_{el}$ for PV.

$$\sum_{\tau_{cv}\in\{ST,PV\}} \alpha^{R}_{\tau_{cv}}\ P^{nom}_{n,\tau_{cv}} \leq A^{max}_{n} \tag{4.13}$$

A pair HP-electric chiller is used to interact with a 5G network. We name electric chiller (EChill) a HP designed for cooling purposes. Both converters operate independently and interact with both the warm and cold network pipes (Figure 4.1c). A third converter helps feeding energy back to the network by canceling out one unit of active heat with one of active cold within the node (i.e. the demand remains unchanged). This ensures the coupling between the warm and cold sides of the 5G network [56, 92, 136, 137] and the connection between heat and cold within a node [56] (c.f. Appendix C).

4.3.3.3 Storage technologies

The storages link time steps and displace productions to better match the demand. Four storage technologies are considered: HWT, Solar Tank, Water in pipes (WiP) and Pit thermal energy storage (PTES).

The model describes a yearly cycle, thus the State Of Charge (SOC) at the start and end must be identical for every storage (eq.(4.14)).

$$SOC_{n,\sigma,\tau_s,\ t^{max}+1} = SOC_{n,\sigma,\tau_s,\ t=0} \tag{4.14}$$

For the sake of simplicity, the maximum charging and discharging power of every storage is set proportional to its capacity (eq.(4.15-4.16)).

$$0 \leq \dot{Q}^{ch}_{n,\sigma,\tau_s,t} \leq \beta^{ch}_{\tau_s} \times P^{nom}_{n,\tau_s} \tag{4.15}$$

$$0 \leq \dot{Q}^{dch}_{n,\sigma,\tau_s,t} \leq \beta^{dch}_{\tau_s} \times P^{nom}_{n,\tau_s} \tag{4.16}$$

The capacity of every storage binds the state of charge. For HWT, Solar Tank and PTES, these bounds are depicted by eq.(4.17). For the WiP, eq.(4.20) is used.

$$0 \leq SOC_{n,\sigma,\tau_s,t} \leq P^{nom}_{n,\tau_s} \tag{4.17}$$

The time coupling equations are expressed as follows (eq.(4.18)) for all but the PTES, and via eq.(4.22) for PTES. Decay κ, charging η^{ch} and discharging η^{dch} efficiencies are considered.

$$SOC_{n,\sigma,\tau_s,t+1} = (1-\kappa_{\tau_s})SOC_{n,\sigma,\tau_s,t} + \eta^{ch}_{\tau_s}\dot{Q}^{ch}_{n,\sigma,\tau_s,t} - \frac{1}{\eta^{dch}_{\tau_s}}\dot{Q}^{dch}_{n,\sigma,\tau_s,t} \forall \tau_s \in \{HWT, ST, WiP\} \tag{4.18}$$

The water in 5G networks has thermal inertia [118, 121, 137]. This inertia is modeled via another storage whose capacity is defined by the volume of water in pipes (eq.(4.19)). Its SOC represents the difference between the expected fixed temperature and the actual temperature in the pipes (eq.(4.20)). The resulting additional thermal losses are assessed in eq.(4.21) and added to the total thermal losses in eq.(4.8).

$$P^{nom}_{WiP} = \rho c_p \Delta T^{flex} \sum_{l \in \mathcal{L}} \delta_l L_l \frac{\pi D_l^2}{4} \tag{4.19}$$

$$-P_{WiP}^{nom} \leq SOC_{WiP,t} \leq +P_{WiP}^{nom} \tag{4.20}$$

$$\dot{q}_{\sigma,t}^{flex} = \frac{8\lambda}{\rho c_p \epsilon} SOC_{\sigma,WiP,t} \tag{4.21}$$

Long-term storage is commonly associated with 5G networks [86,102–104,138]. The PTES technology is chosen for its competitiveness, independence of geological conditions [138] and temperature adequacy [115]. Eq.(4.22) enforces the time coupling for PTES. Similar model formulations are used in related literature for aquifer and boreholes, coupled with additional HPs for the temperature adequacy [111]. With higher investment, operating and maintenance costs, other long-term storages would be less accepted. The PTES serves both the warm and cold network sides (Figure 4.1b), following principles described in [139,140].

$$\begin{aligned} SOC_{\tau_s,t+1} =& (1-\kappa_{\tau_s}) SOC_{\tau_s,t} + \kappa_{\tau_s'} SOC_{\tau_s',t} \\ &+ \eta_{\tau_s}^{ch} Q_{\tau_s,t}^{ch} - \frac{1}{\eta_{\tau_s}^{dch}} Q_{\tau_s,t}^{dch} \\ &\forall (\tau_s, \tau_s') \in \{PTES_H, PTES_C\} \ , \ \tau_s \neq \tau_s' \end{aligned} \tag{4.22}$$

4.3.3.4 Sets of technologies

This work compares DTNs, but nodes can choose to disconnect from it. Specific technologies tailored for collective systems cannot be used by disconnected nodes [107]. In the model, sets of technologies are allowed Ξ_n^{Net} or disallowed Ξ_n^{Node} if a node n connects to any potential link $l \in \mathcal{L}^{(>n)}$ of the DTN. Eq.(4.23–4.24) links the connection status (binary variable δ_l) with the technology installation status δ_{tau} via the Big$M$ method[2], M being an arbitrarily large number. The technology sets are described in section 4.4.

$$\sum_{\tau \in \Xi_n^{Node}} \delta_\tau \leq M \sum_{l \in \mathcal{L}^{(>n)}} \delta_l \tag{4.23}$$

$$\sum_{\tau \in \Xi_n^{Net}} (1-\delta_\tau) \leq M \sum_{l \in \mathcal{L}^{(>n)}} \delta_l \tag{4.24}$$

[2] https://en.wikipedia.org/wiki/Big_M_method

4.4 Case Study

4.4.1 Scenarios and sensitivity

We name *base cases* the scenarios using exclusively default **bolded** values in Table 4.2. Alternative scenarios always only use a single non-default value at a time.

Table 4.2: Scenarios and associated values

	Field	Values	
Always	**Location**	{LA, OR, WI}	
	Generation	{4, 5}	
	CO_2 Reduc	{-0,-25,-50,-75}	%
Strategy	Price Elec	{**0.2**,0.3,0.4}	\$/kWh
	Carbon Elec	{**0.25**,0.4,0.7}	$kgCO_2$/kWh
	Carbon Tax	{0,**65**,170}	$\$/tCO_2$
	Lifetime	{7,15,**tech**}	y
Resource	Storage	**y**/n	
	MicroGrid	y/**n**	
	WasteHeat	**y**/n	
	Roof Fraction	{ **50** }	%
Layout	Shape	{**Star**, Line, Loop} (Figure 4.2)	
	Station	{**Hospital**, SuperMarket, LargeOffice}	

Three US locations are selected: Louisiana (LA), Oregon (OR) and Wisconsin (WI) respectively for their hot, temperate and cold climate (Table 4.3). The carbon reduction is relative to what is achievable. It is not an absolute emission reduction, as outlined in the results section. The value *tech* for the lifetime means the annualization factor is considered for the lifetime of each technology. The roof fraction is the usable fraction of roof surface to install ST or PV. Three of the four load types (hospital, offices, supermarket, residential) are considered to host the central station (c.f. Figure 4.3).

4.4.2 Resources and data

The study uses aggregated cooling and heating loads, where the latter merges demands for hot water and space heating together. Load data of 4 building types

over 3 locations (Table 4.4, Appendix D) from the dataset [77] is used. The dataset is built from EnergyPlus simulations of archetypes in the US.

The EnergyPlus website[3] hosts the TMY3 weather files used for the simulations. These files are combined with Gsee[4] (PV) and the model from [141] (ST) to precompute solar resources. They are combined with a model from [142] for the soil temperature and with a model from [143] to precompute the COP of HPs and chillers.

Table 4.3: Degree days (base 18.3^oC)

	LA	OR	WI
HDD	754	2326	4082
CDD	1547	204	303

Table 4.4: Total thermal load for each node

	LA		OR		WI	
(GWh/y)	Heat	Cold	Heat	Cold	Heat	Cold
Apartments	1.83	5.57	3.39	1.20	6.24	1.53
Hospital	1.38	7.36	1.89	5.26	2.10	5.03
Offices	0.39	6.89	0.68	2.12	1.72	2.37
Superstore	0.44	0.52	1.37	0.04	1.96	0.09

4.4.3 Temperature setpoints assumptions and waste heat

The water in 5G networks is maintained at 25^oC and 10^oC for the warm and cold sides [43, 144]. For the 4G, heat supply and return are at 70^oC and 40^oC [8], cold supply and return are at 6^oC and 12^oC [54]. Demands for hot water and space heating are merged. Heating service is at 40^oC, cooling service is at 12^oC. HP and chiller efficiencies are precomputed for a refrigerant 15^oC above heating service and at 1.7^oC for cooling service. Indoor setpoints are specified in the dataset manual [77].

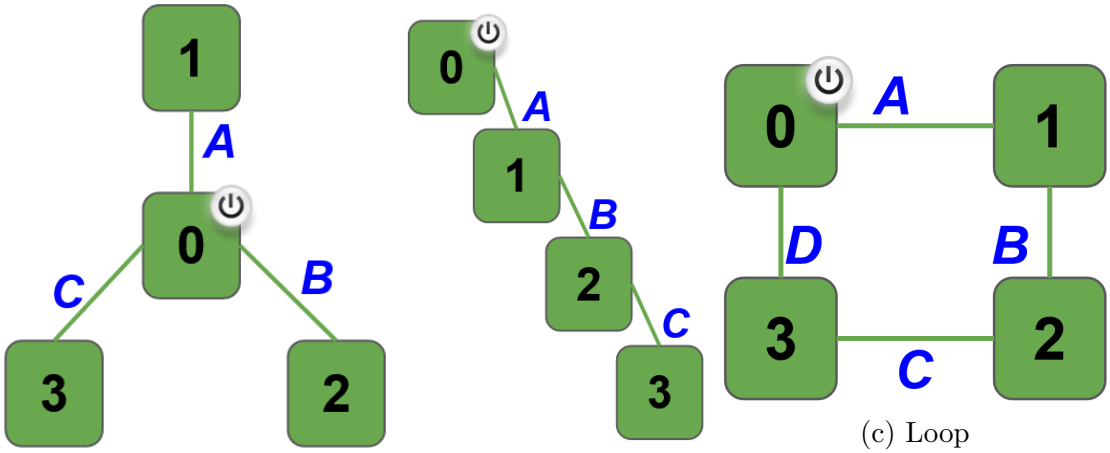
Waste heat (WH) is modeled by installing additional ST at the central station, inspired by the prosumer model in [92]. The ST capacity used for WH equals half the peak load of the station node. WH only represents a heat production surplus between 60^oC and 80^oC, it has neither investment costs nor roof area, and no associated solar tank.

[3] https://energyplus.net/weather
[4] https://pypi.org/project/gsee/

4.4.4 System potential designs

4.4.4.1 Network layouts

Nodes may connect following the "Star" configuration in the base cases (Figure 4.2). The interest of "Line" and "Loop" configurations are investigated for the assessment of geographical features in section 4.6.4.



(a) Star (b) Line (c) Loop

Figure 4.2: Network configuration layouts

4.4.4.2 Nodes layouts

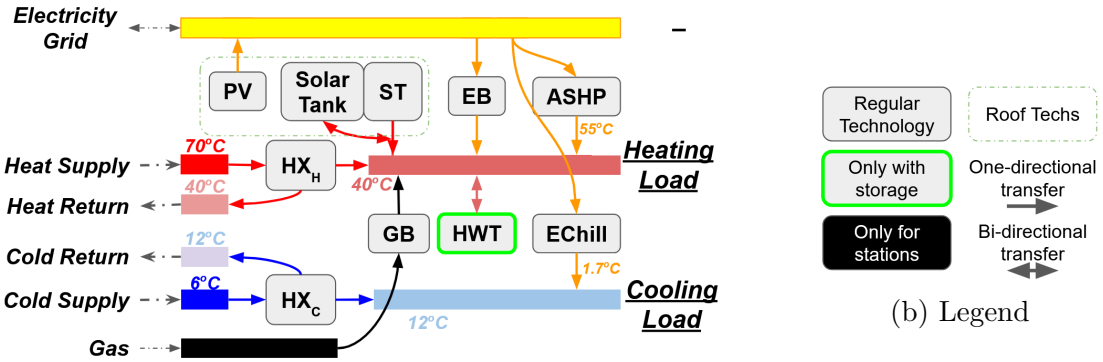
The potential configurations of nodes are shown in Figures 4.3a-4.3c for 4G nodes and in Figure 4.3d for 5G nodes. Technologies in black cells are only available for the central station. If a central station disconnects from the network, then:

- 4G: the potential layout of the central node becomes identical to those of a regular node (Figure 4.3a)
- 5G: elements in black cells in Figure 4.3d are no longer available.

4.5 Limitations

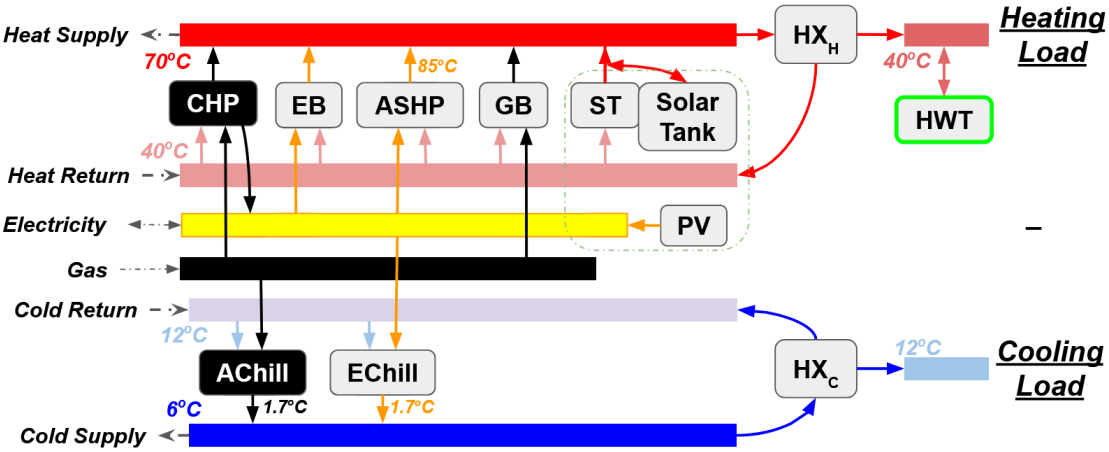
4.5.1 Model limitations

We acknowledge the following necessary simplification in the model:

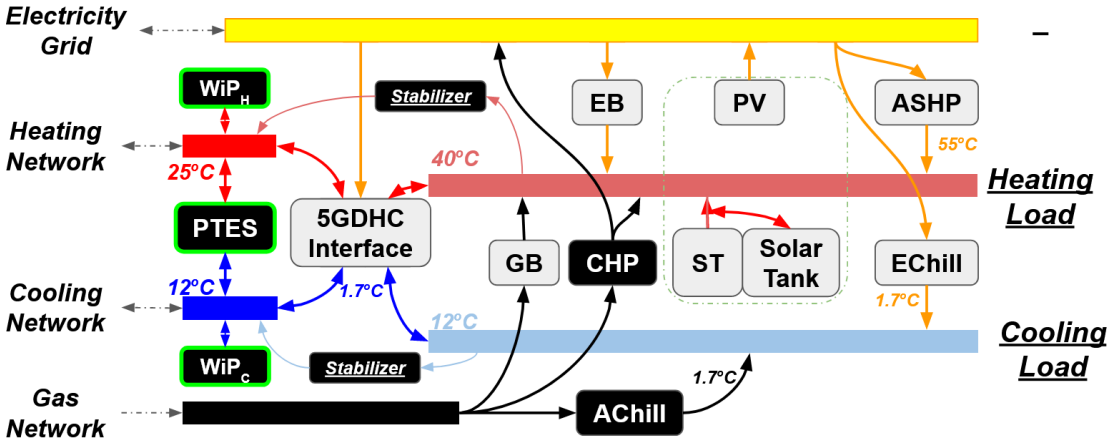


(a) Potential node connected to a 4G network

(b) Legend



(c) Potential station of a 4G network



(d) Potential node & station connected to a 5G network

Figure 4.3: Potential layouts of nodes

- A MILP formulation is used. It allows simultaneous sizing and operation of all nodes and network. However, it is designed for early stage assessment and is not a "ready to implement" model. The model is based on previously introduced literature. We believe its resolution matches the uncertainty on load and weather at very early design stage.

- Temperature setpoints are set constant prior to computations. No temperature control is included. Consistent temperature levels from related literature [8, 9, 43, 54, 56, 87, 90, 96, 144, 145] influenced this choice. These assumptions allow to precompute efficiencies and linearize the model (c.f. section 4.3.2). To maintain the scope of the model (e.g. sizing, central-decentral trade-off) and tractability [80], more advanced linearization techniques [113] were not considered.

- Some modeling elements (e.g. non-linear investment costs [80], time varying cost and carbon intensity of electricity) were not included in the study for the sake of simplicity and feasibility of all scenarios.

4.5.2 Scope limitations

The study breaks down results to gain understandability on how performances are reflected in the energy and investment strategies suggested by the solver. Three cases are selected to offer an overview of similarities and differences over different climates. A systematic in-depth investigation limits the possible number of cases (section 4.4.2). However, 3 cases are not sufficient to generalize observations. The impact of different load patterns, heating-cooling overlap (Appendix D), local weather and resource availability is left out of the scope to instead focus on decision mechanisms within the model.

4.6 Results

An in-depth analysis of base cases is conducted in Section 4.6.1. Sections 4.6.2-4.6.4 extend on the sensitivity of main results to key parameters.

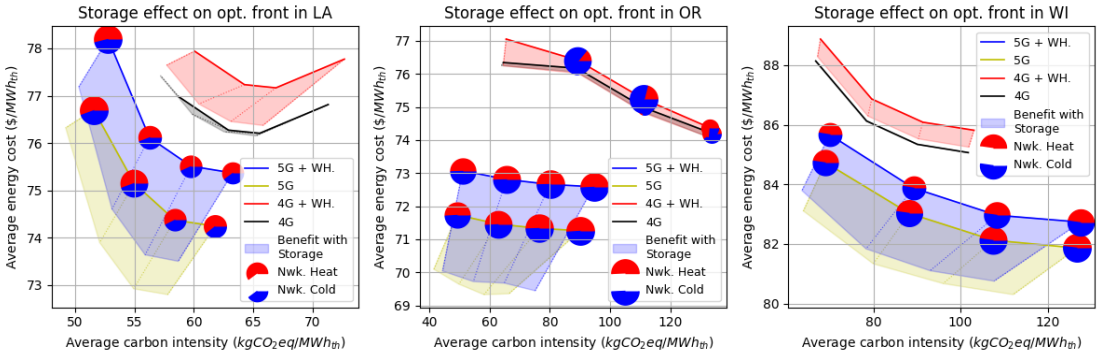


Figure 4.4: Pareto fronts of all base scenarios showing the improvements caused by allowing the usage of storage (colored area) and the overall relative connectivity of nodes in the network for each solution (pie charts)

4.6.1 Base cases: Storage and Waste Heat

4.6.1.1 Pareto Fronts and connectivity

Results for the base cases are shown in Figure 4.4, along with scenarios without storage and WH. The solid lines show cases without storage, the benefit of allowing storage is represented by the colored area. The pie charts show the share of heat and cold delivered by the network, their sizes depict the nodes connectivity.

Pareto fronts for 5G strictly dominate those for 4G in all locations. The 5G DTN is always used, with consistent share and connectivity throughout the decarbonization. The 4G DTN is only used in OR. For this case, an increasing share of cooling transits via the DTN to support decarbonization, except for the most restrictive carbon target. For all other 4G cases, the optimal solution excludes the DTN and focuses on decentralized systems.

The 5G network supplies equivalent shares of heat and cold in all places. This is more beneficial for places with balanced demands. For example, the DTN covers 46% of the heating and 38% of the cooling for a well-balanced demand in OR, but 77% and 14% respectively for the heating and cooling of a cooling dominated demand in LA.

The 4G network benefits from independent heating and cooling sides. When used in OR, it consistently covers 25% of the heating, and 10% to 62% of the cooling. Decentralized HPs are preferred, since high network temperatures limit the efficiency of central HPs. Thus only gas is used to supply heat via the DTN. Central and decentralized chillers have similar efficiencies due to comparable service temperatures, the centralized solution is preferred to reduce investment costs.

The storage increases the financial gap between 4G and 5G cases. Carbon emissions are also reduced, more for 5G cases than 4G. Waste heat always increases costs and emissions. This is compensated by storage in 4G. It is partly shared via the network in OR, which reduces the gap between 4G and 5G there.

The decarbonization cuts up to 75% of what is technically achievable, as loads cannot be reduced. With 5G, the kg of removed carbon emission costs on average ¢42 in LA, ¢1.3 in OR and ¢6.4 in WI. LA can only reduce 14.5% of absolute emissions, while a 41% cut is achieved in OR and WI. For the 4G cases, the transition costs ¢15 in LA, ¢2.7 in OR and ¢8.3 in WI, with respectively 12%, 54% and 34% absolute carbon reduction achieved.

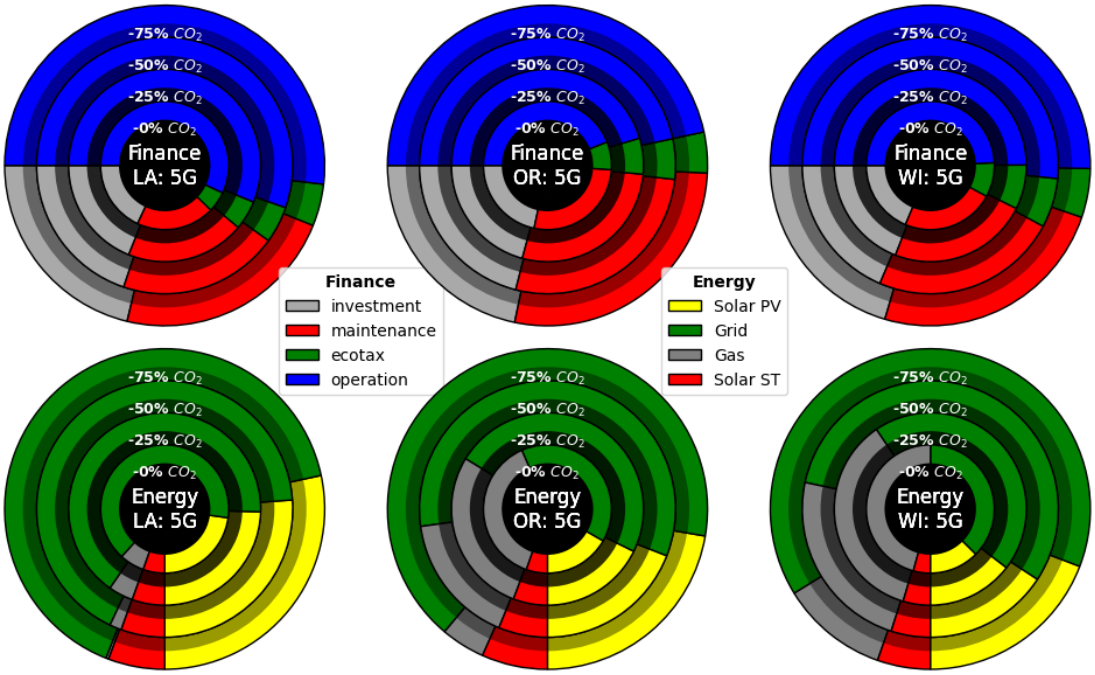
The cost increase is stronger for stricter carbon emission targets. This is particularly visible in LA, where half of the increase happens when going from -50% to -75% emissions. Reasons for this are highlighted in Figure 4.5.

4.6.1.2 Analysis of the energy strategy

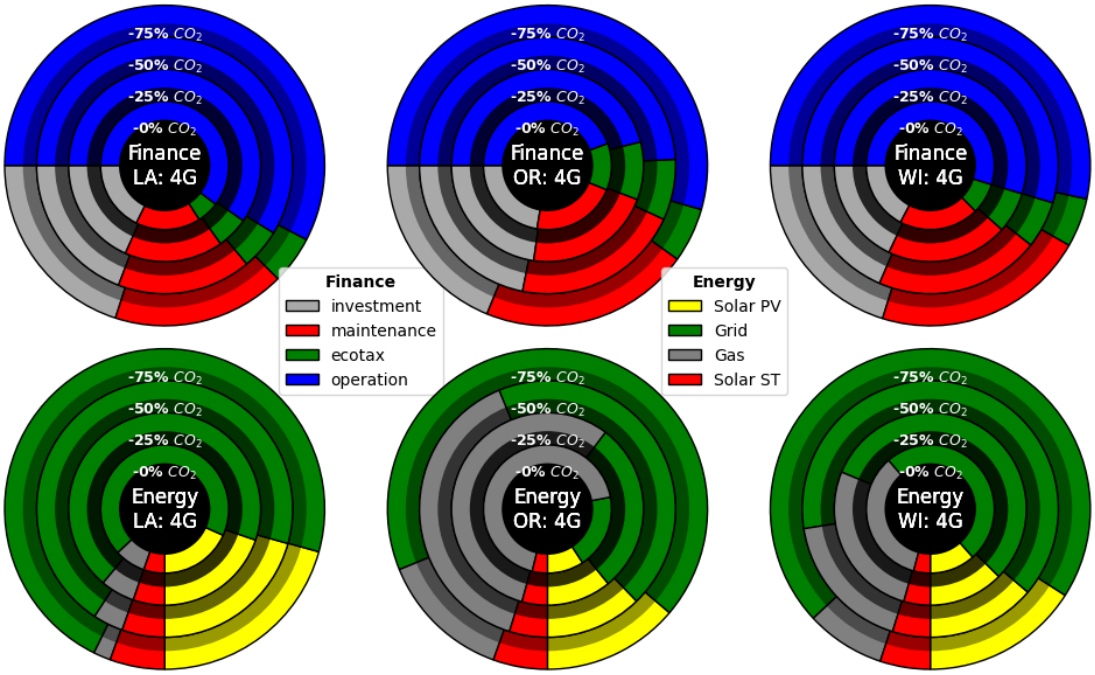
Electrification drives the decarbonization for all locations (Figure 4.5). It has minor influence on the financial strategy, consisting of 45-60% of operating costs, 35-45% of maintenance and annualized investment, and 5-10% of carbon tax. The most significant change is observed for the 4G case in OR: 10% of the cost is transferred from investment and maintenance to operation between -50% and -75% emissions (Figure 4.5b), i.e. when the 4G network is abandoned (Figure 4.4).

The share of gas in the energy mix consistently decreases throughout the decarbonization, replaced by grid and PV electricity. The share of PV electricity is greater in 5G cases than in 4G cases, and this share increases more throughout decarbonization. Inversely, 4G uses more gas and is less able to reduce its usage when decarbonizing.

The energy portfolio differs in every location and explains the different slopes observed in Figure 4.4. OR use the most gas at -0%, which is cheap to replace with electricity. LA is almost entirely electrified at -0%. Removing the remaining gas requires costly investments for PV and storage for minor reduction of the emissions. WI is an intermediate case where a trade-off of the above occurs.



(a) Decarbonization strategies with 5G



(b) Decarbonization strategies with 4G

Figure 4.5: Financial and energy strategies throughout decarbonization, with storage and waste heat included

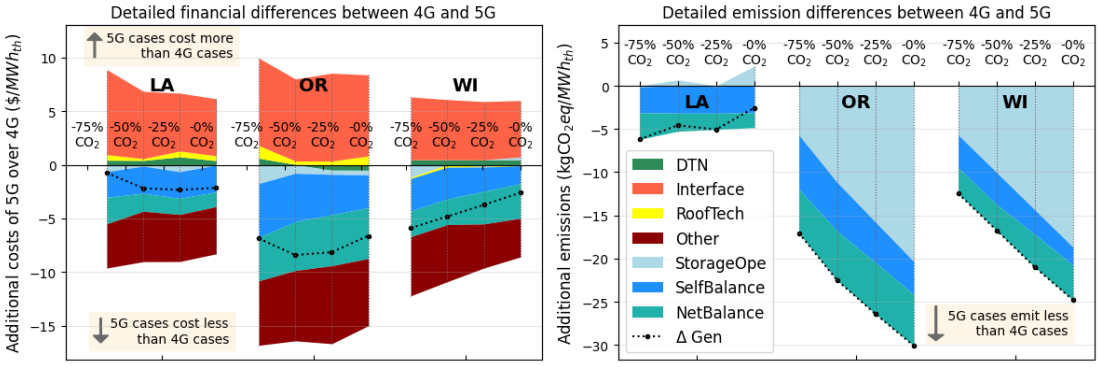


Figure 4.6: Detailed difference in costs and emissions between homologous 4G and 5G cases.

4.6.1.3 Difference between 4G and 5G cases

Figure 4.6 details the financial (left) and environmental (right) differences between 4G and 5G cases. Positive values mean that 5G cases are more expensive or polluting than their homologous 4G cases in the specified category. The dashed line ΔGen confirms the consistent overall lower cost and carbon intensity of 5G cases.

5G DTN can balance out simultaneous heating and cooling demands within nodes ("SelfBalance") and between nodes ("NetBalance"). This increases the overall system efficiency, with cost decreases of 5G cases over 4G cases valuated from $4\$/MWh_{th}$ in LA to $10\$/MWh_{th}$ in OR. This also consistently reduces emissions by 5-10 $kgCO_2eq/MWh_{th}$.

The interface network-buildings is more expensive and more often used in 5G cases (HP) than 4G cases (HX), representing 5-$10\$/MWh_{th}$ additional costs for 5G cases over 4G cases. However, it is mainly compensated by a reduction in investment for other decentralized generators ("Other"). The remaining investments (roof technologies, storage) weigh little on the financial difference between 4G and 5G, first because both cases benefit from it similarly, second due to the low associated maintenance costs and annualization of the investment. Their impact on the emissions is none as embodied emissions are not considered.

The storage operation ("StorageOpe") brings little financial difference between 4G and 5G cases, but is the main factor explaining differences in emissions in OR and WI. The lower total costs related to storage for 5G cases observed in Figure 4.4 are explained by an induced change in the system and in energy strategy, while returned energy itself has minor financial value against fixed grid and gas prices. However, the storage operation cuts carbon emissions up to $20kgCO_2eq/MWh_{th}$

more in 5G cases than in 4G cases. The energy stored and restituted is assumed carbon-free, can be stored in larger quantities in the PTES (5G) than in individual HWT (4G), and can be better redistributed through the 5G DTN. For OR and WI, decarbonization reduces the difference, as 4G cases increase their use of storage. For LA, storage is used little as renewable production is mainly directly used for air-conditioning.

4.6.2 Financial and environmental parameters

This section presents the sensitivity of results to strategic parameters (c.f. Table 4.2) for the cases in OR. OR is specifically selected for its representativeness and acceptance of the 4G DTN. Figures 4.7-4.10 have a similar structure and are purposefully dense in information. Subsections 4.6.2.1 to 4.6.2.4 analyse the overview and major trends from each figure, the reader is free to analyse the figures beyond the textual interpretation.

A block in Figures 4.7a to 4.10a shows the financial difference between two consecutive scenarios for a same network generation for one decarbonization level. The block filling details the contribution of each financial category in the difference. Each block couple covers all three studied scenarios for each case. Block couples of similar decarbonization levels are put side-by-side to facilitate comparison between 4G and 5G. The block frame helps visualizing the network generation. The solid lines show the carbon intensity for each scenario (right axis).

A bar in Figures 4.7b to 4.10b shows the energy strategy of one case under one scenario. The bar filling details the contribution of each energy source in the mix. Each bar triplet covers all three studied scenarios for each case. Bar triplets of similar decarbonization levels are put side-by-side to facilitate comparison between 4G and 5G. The bar frame helps visualizing the network generation. The solid line shows the share of renewable energy in the local production mix (right axis). The size of circles indicates the overall connectivity, i.e. the fraction of thermal energy covered by using the network.

4.6.2.1 Price of electricity

Both costs and emissions increase with the price of electricity (Figure 4.7a), making the decarbonization more expensive. An increasing price of electricity sees grid electricity being replaced by gas and RnE, increasing the operation and investment costs, while making the 4G network more attractive (Figure 4.7b). As the increase

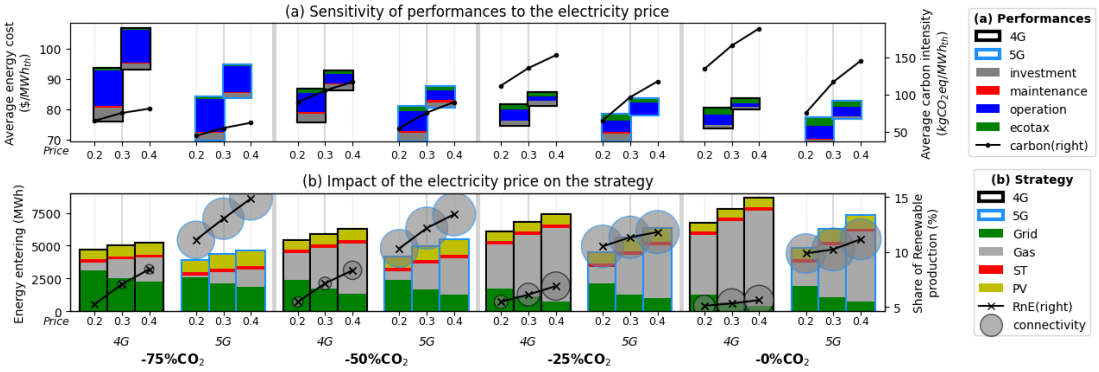


Figure 4.7: Effect of the electricity price on 4G and 5G cases in OR.

in electricity price affects the performances of all cases more than a change of DTN technology does, 5G cases remain better than 4G cases in costs, emissions, DTN usage and RnE penetration.

4.6.2.2 Carbon tax

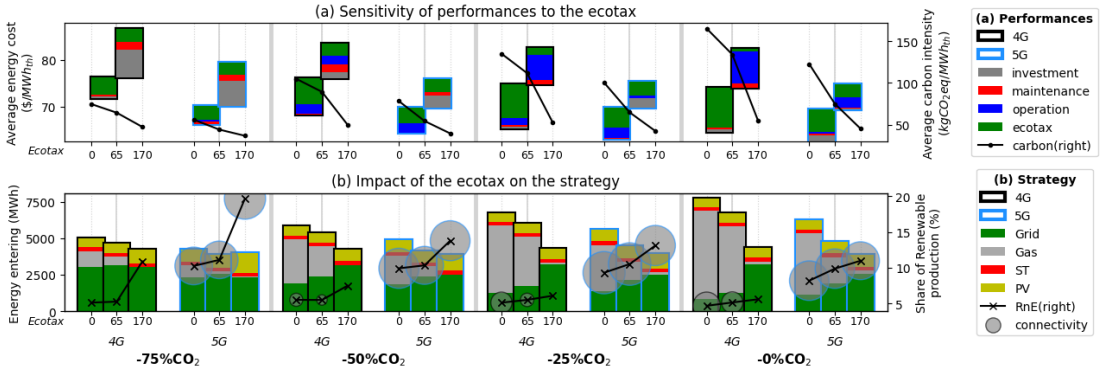


Figure 4.8: Effect of the carbon tax on 4G and 5G cases in OR.

The carbon tax affects the financial and environmental performances more than a change of DTN technology (Figure 4.8a), despite being a minor share of the investment strategy in base cases (Figure 4.5). Its increase significantly raises costs at all decarbonization levels, and reduces carbon emissions as much as the enforced decarbonization. When the ecotax goes from 0 to 65\$/t_{CO_2eq}, its share in the financial strategy explains the cost rise while little changes are observed in the design. From 65\$ to 170\$, the system design is impacted, and the share of operating and investment costs explains the cost increase. These changes come along with lower energy uses for 4G and 5G cases (Figure 4.8b), i.e. an overall increase in efficiency, in which gas is substituted by grid and renewable electricity. Thus the 4GDTN is less attractive while the 5GDTN remains highly used.

4.6.2.3 Carbon intensity of electricity

A higher carbon intensity of the electricity increases both the costs and carbon emissions. This effect is stronger with stricter carbon targets. Higher carbon intensities limit the decarbonization more in 4G cases than 5G, increasing the gap between them (Figure 4.9a). However, the increase in DTN usage is marginal. For -0% and -25% emissions, the ecotax and higher share of gas cause the mild cost increase (Figure 4.9b). For stricter carbon targets, the required cut in gas usage forces system changes, explained by higher shares of RnE and investment

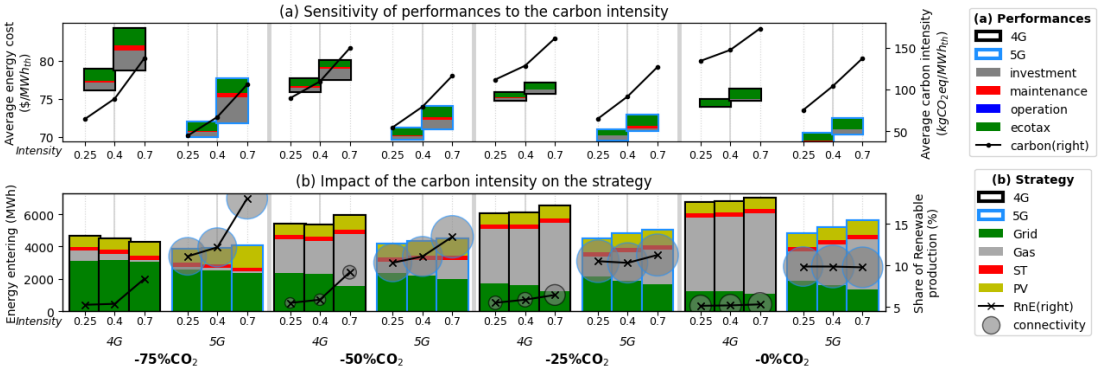


Figure 4.9: Effect of the electricity carbon intensity on 4G and 5G cases in OR.

costs. The penetration of RnE and its increase are significantly stronger with the 5G case.

4.6.2.4 Lifetime of equipment

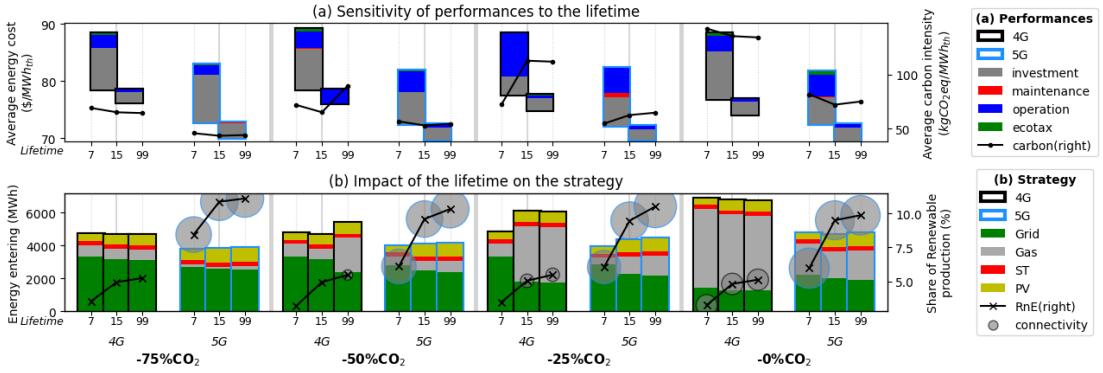


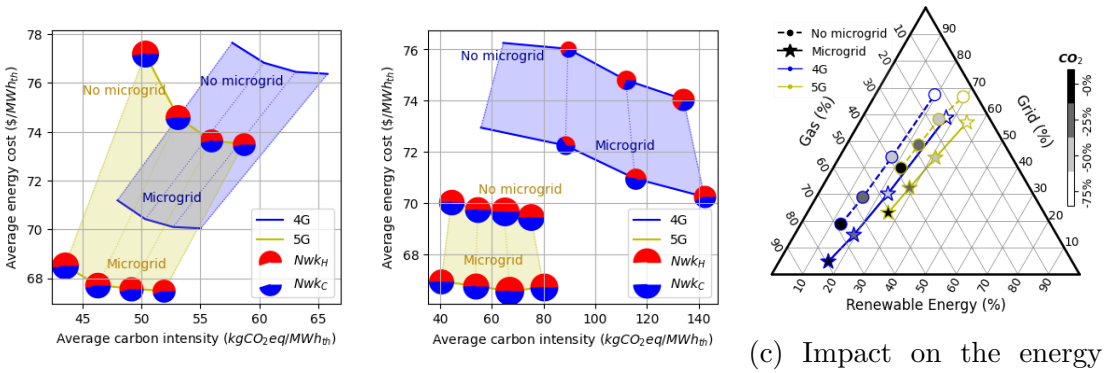
Figure 4.10: Effect of the lifetime of assets on 4G and 5G cases in OR.

The lifetime relates to the time span over which the annualized investments are spread. Shorter lifetimes have little impact on the overall emissions, while inducing a rise in costs mainly driven by higher investment (Figure 4.10a). The gap between 4G and 5G remains unchanged, while the impact of lifetime is greater than these of changing the DTN technology and is consistent throughout decarbonization. The higher investments only reflect the direct link between investment and lifetime, while stable energy strategies reveal that minor structural changes occur in the system (Figure 4.10b). Shorter lifetimes slightly reduce the share of gas and RnE, both relying on generation technologies, and the DTN usage is slightly reduced. With these changes the operation costs also contribute to increasing the overall costs.

4.6.3 Micro-grid

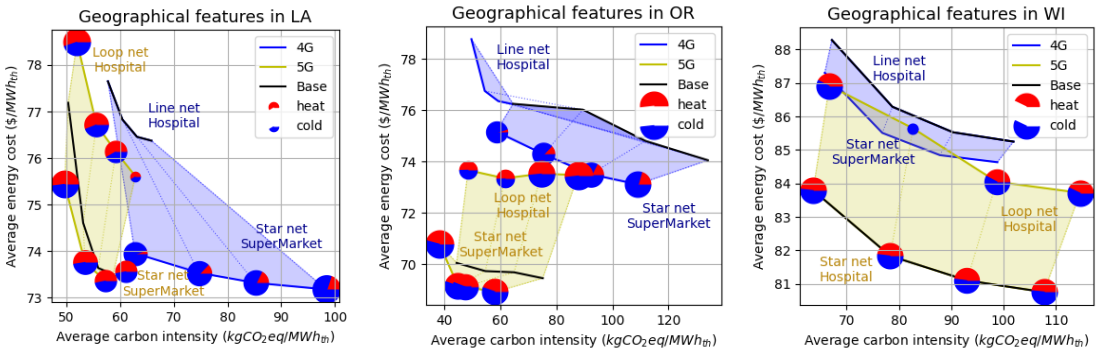
A micro-grid model was added to allow an electrical connection between buildings. The micro-grid simply considers DC lines and the cost of infrastructure is not considered. This is suggested to study the impact of accounting for locally generated electricity. Indeed, buildings generate PV electricity and feed surplus back into the grid. However, that electricity is likely to flow directly to a neighbour building in need.

Figures 4.11a and 4.11c show the effect on financial and environmental performances of accounting for local renewable generation. A drop in costs is observed



(a) Financial impact in LA (b) Financial impact in OR (c) Impact on the energy mix in OR

Figure 4.11: Impact of accounting for local electricity at the district level



(a) Performances in LA (b) Performances in OR (c) Performances in WI

Figure 4.12: Performances of different topologies and building types hosting the main energy station.

with local electricity. This decrease is twice as important as the financial gap between 4G and 5G in LA, while being equivalent to the gap in WA and half of it in OR. Additional cooling is supplied via the 4G DTN in OR, while the 4G DTN remains unused in other locations. The usage of 5G DTN remains unchanged. Accounting for local electricity reduces carbon emissions in LA and WI, where cooling and RnE potential align, however hardly impacts decarbonizing in OR.

The consequences on the energy strategy are shown in Figure 4.11c for OR, with similar trends in other locations. The share of gas increases with local electricity. This effect is stronger in OR, explaining the lack of carbon benefits there. With local electricity, the share of RnE systematically increases by 10%, and rises 10% more over the decarbonization instead of 7%. The rest of the strategy behind decarbonization is unchanged, with grid electricity replacing gas.

4.6.4 Topology and energy station

Different scenarios were tested for the topology and building hosting the main energy station.

Figure 4.12 shows the best and worst performing combination of topology and building type hosting the central energy station. With a favorable topology and station location, the usage of a 4G DTN can become viable (Figure 4.12a) and improve performances of 4G cases (Figure 4.12b). Performances of 5G cases can also be impacted, and the 5G DTN usage may be less consistent. A different topology and station location may result in close comparisons between 5G and standalone (Figure 4.12c), 5G and 4G DTN (Figure 4.12b) or even make 4G DTN more financially attractive than 5G (Figure 4.12a).

4.7 Discussion

The heat-cold balancing within a building allowed with 5G networks represents a significant advantage in this study. This is technically allowed by the usage of independent HP and EChill interacting with the network [56]. This effect could be reduced or canceled with buildings having no concurrent loads or using VRF and advanced HVAC systems. However, 4G networks only focus on delivering thermal energy at the supply temperature required by buildings and need the required temperature to be homogeneous across the network. 5G networks are rather associated with high quality buildings, thus the data and network interaction as implemented, seem reasonable.

The share of energy via the network represents another major part of the superiority of 5G networks over 4G networks, while most related literature concludes on the superiority of centralized systems. The current network model is based on energy transfer, as implemented by Wirtz et al. [56, 80, 140, 146] and used in estimates of 5G network potential [117, 145, 147, 148]. It also aligns with 4G models commonly used in literature [7–9, 108]. More generally, this energy transfer paradigm using MILP was found accurate when confronted to real data [111] for a 5G network, and even verified against simulation for a 4G case [108]. It differs from temperature-based networks in comparative studies using optimization and simulation [87, 90–96], which is supposed to be more accurate. For example, Gross et al. [92] verified the fluid-dynamic aspect of their simulation approach against real data.

An additional explanation may be found in the case study and share of cooling in the demand. Table 4.1 shows that most authors comparing both technologies do so with little load diversity, and when the diversity exists, chosen cases show a largely dominating heating demand. The study by Zhang et al. [148] shows conditions and types of buildings most suited to boost efficiency of a 5G network. It happens to match with the combinations of building types and climates we chose.

The observations are only valid within the reduced pool of cases. Other conditions (weather, resource availability, load patterns, concurrent heat and cooling demands, number of buildings, load density, temperature set points and controls) influence the performances of DTN based MES. Their importance needs further quantification through a significantly larger pool of cases.

Specifically, temperature integration and control have significant impacts on DTN flexibility and efficiency. This study focuses on early stage design assessment,

temperature approximations are assumed to level the high uncertainty of loads and weather conditions. Later and more informed stages of development would benefit from additional information (e.g. varying prices and carbon intensity of electricity) and control (e.g. temperature control) to better implement informed control strategy.

Including more information at early stage design is appealing but faces limitations. Limitations may reside in the need of drastic simplifications on other model aspects. Or burdensome iterative processes relying on advanced tools must be used [149]. For the latter, future work is needed to build fast and reliable surrogate models to replace burdensome advanced tools in such iterative process.

The current study also highlights the need for further investigation about 5G network modeling. The ability to balance energy is at the center of interest for 5G networks, however its implementation and results in the current study may be controversial. Future work should focus on assessing the importance and technicity of the energy sharing through a 5G network.

4.8 Conclusion

The selection network technology in the development of energy projects must be addressed at early-stage design. The present study implements a linear energy hub model encompassing specific characteristics of both 4th (4G) and 5th generation (5G) district heating and cooling networks, including the strong warm/cold coupling of 5G technologies. The model is challenged across a variety of scenarios for districts with a mix of building types.

The 5G is found to perform financially and environmentally better than 4G. This differs with results from related literature. The ability to balance out simultaneous heating and cooling demands across the 5G network explains most of the difference. It overtakes the higher investment for decentralized heat pumps. The current energy-based model uses a different paradigm than most literature comparing 4G and 5G, which is based on network temperature. Both paradigms were validated in publications and against simulation and real data. However, the latter does not seem to capture this balancing ability across the 5G network.

Results should be further nuanced due to the model sensitivity in critical parameters. Annualized models amplify the role of operation and minimize these of investments, which strongly differs from common practice accounting for shorter ROI periods. Moreover, potential topology and the spacial spread of loads within

the network is found to significantly impact the difference between the two network technologies, possibly making 4G more profitable than 5G in specific climate zones.

The difference between the temperature and energy paradigms in the modeling of 5G networks requires further investigation. The computationally intensive requirements of such models does not allow statistically rigorous sensitivity analysis and limits the integration of temperature control, which motives further development in surrogate modeling for district thermal network.

Chapter 5

Improved surrogate modeling for multi-energy system design: model architecture, sampling and scaling choices

Abstract

Multi-energy systems (MES) are a key concept to develop more sustainable energy systems, but optimizing their design is computationally burdensome. This paper explores the use of machine-learning (ML) based surrogate modeling to support the optimal design of MES. Surrogates are simple models, often ML-based, used to approximate detailed simulations, in this case MES design optimizations. These models give instant responses, enabling fast comparisons and explorations of trade-offs between design variables. While prior works have used surrogates to predict system cost and other objectives, few works have used them to directly predict the optimal system design, and those that do show poor performance.

This paper provides an extensive methodology to improve surrogate performance on a MES design problem with small datasets. Four approaches were found to significantly improve surrogate performance: a careful and objective-oriented selection of samples, the use of upsampling to balance datasets, the use of non-linear rescaling methods, and a specific neural-network architecture called Mixture-of-Experts. These work together to turn the original design variable distribution (i.e. of the output) into a Gaussian-like data distribution, that can be more easily learned by the neural-network.

The resulting surrogate model almost instantly predicts optimal energy system designs with high precision. This was tested across a wide variety of different climates, building types and decarbonisation goals. Such surrogate models will make it much easier to explore different MES design options.

*This chapter is related to the publication **[P4]***

5.1 Introduction

The transition towards sustainable energy systems requires innovative approaches to optimize the integration of diverse resources and build robust multi-energy systems (MES) [16]. Advanced modeling techniques, for example the "Energy Hub" (EHub) concept [20], were found to be relevant from early design stages to identify synergies between diverse technologies and energy sources to be exploited.

Traditionally, the optimization of MES models involves formulating complex mathematical models representing the interactions among different components and solving them using conventional optimization techniques [16]. However, the computational burden associated with solving these models grows exponentially with the system size [7] and level of detail [61, 62], thus limiting the possibility to scale models up.

In recent years, machine learning (ML) techniques and the concept of surrogate modeling have been explored to address challenges related to the design [150] and operation [151] of MES. Surrogate models aim at replacing a detailed and computationally burdensome task with a fast estimate, often ML based. Applications of interest lie in early stage design exploration and simulation intensive sensitivity analyzes, thus the acceptability of the time vs. accuracy trade-off. For instance, surrogates of building energy simulations are common [152–156].

The most common usage of surrogate models for MES design is called surrogate-assisted modeling. It is a bi-level optimization approach [157]: on the first level, an evolutionary algorithm (e.g. GA, PSO) suggests MES designs (i.e. selection and sizing of generators and storages); on the second level, a surrogate evaluates the potential performances of the design by estimating the costs [58, 158, 159], emissions [58, 160], and/or other objectives [161–163] related to the system operation.

Surrogates fundamentally specialize in linking a set of model parameters with a set of model responses [164], but do not offer guarantees of performance on

subregions of interest [165]. Where the indirect design suggested in [166] and [163] requires surrogates performing well for a diversity of designs (including absurd designs), nothing guarantees the required accuracy near the optimum or close to non-dominated solutions.

To the best of the author's knowledge, only four works attempted to surrogate the whole MES design process. Köhnen et al. [59] compare neural-network (NN) architectures to estimate capacities and total energy supply for an MES. Jahangiri et al. [167] build a surrogates with a deep NN for a grid-expansion planning model of Canada, then conduct a sensitivity analysis with the surrogate. Thrampoulidis et al. [44] build surrogates with NNs for a MES model addressing capacity sizing and selection of retrofit packages for residential buildings in Zurich. To define a cost-carbon pareto-front (PF) of optimal solutions, the NNs are chained and use both data and prediction from the previous NN to estimate the optimal design on the next point of the PF. The authors extends the surrogate capabilities to the whole of Switzerland in [60], by replacing each NN by a super-structure of 6 intertwined NNs. Each of the two regressors and four classifiers addresses a specific set of variables and benefits from the prediction of other NNs from the super-structure.

A common limitation found across [44, 59, 60] is the large difference in performance between the predictions of some variables. Where surrogates can predict the costs and emissions (i.e. the objective values) and some capacities with great precision ($R^2 \geq 0.95$), they systematically fail ($R^2 \leq 0.7$) at predicting other capacities (e.g. photovoltaics, heat pump and storage sizes). Results from [59] also reveal that the random weight-initialization before the training of NN may lead to significantly different performance between two repetitions of the same task, especially on the variables yielding insufficient performances. These limitations prevent the full deployment of a functional surrogate, but are neither addressed nor discussed.

This paper investigates the possible reasons for these discrepancies, and characterizes a set of methodological solutions to build more reliable surrogates of MES for design applications. Accurately predicting the optimal capacities of the components of an MES is the primary purpose of a design optimization; predicting just the objectives with no knowledge of the capacities, or predicting capacities with high errors, is not sufficient. The novel methodology developed in this paper particularly focuses on predicting all component capacities. To ensure its robustness and generalizability, the methodology has been tested across diverse climates, building types and problem constraints (e.g. desired carbon emissions limits).

The document is structured as follows. Section 5.2 introduces the case study

(core application, model, variables of interest and data used) and introduces the core hypothesis driving the study. The new process suggested to train surrogate models and address the hypothesis is presented in great details in Section 5.3. Results are presented and discussed in Section 5.4, investigating if the new implementation addresses the issue, how it addresses the issue and why it is successful.

5.2 Multi-Energy System design problem formulation

The methodology and results presented in this paper address the MES design problem for the system presented in Figure 5.1. The model formulation, variables of interest and data used are respectively described in sections 5.2.1, 5.2.2 and 5.2.3. A specific property of the application is highlighted in section 5.2.4, where the core hypothesis behind this research is formulated.

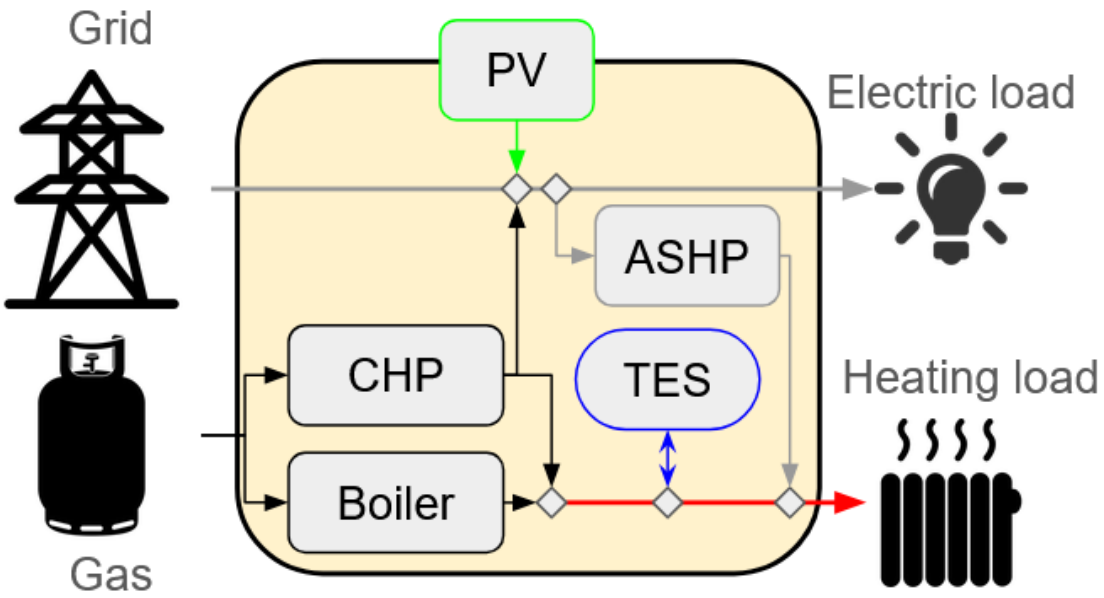


Figure 5.1: Potential layout of the studied energy system

5.2.1 Energy System model

The core MES is designed following the concept of the Energy Hub [20]. The design application is formulated as a mixed-integer linear programming (MILP) problem, as largely covered in related literature [16]. A mathematical optimizer optimally selects and sizes the asset presented in Figure 5.1, and suggests an

optimal hourly control strategy to operate the assets (purchase electricity and gas, store, restitute and transform energy) to cover the hourly heating and electric load of some building for a whole year (8760 time steps). The suggested sizing and operation strategy aims to minimize the annualized cost under some carbon emission reduction target. The few following equations are the main equations leading to the outputs of interest for this study.

The system design assumes an underlying hourly operation, during which energy can be bought ($\dot{Q}^{\mathcal{N},in}_{\sigma,t}$), sold ($\dot{Q}^{\mathcal{N},out}_{\sigma,t}$), transformed ($\dot{Q}^{\mathcal{T},in}_{\tau_{cv},t}$), stored ($\dot{Q}^{ch}_{\sigma,\tau_s,t}$) and consumed ($\dot{Q}^{load}_{\sigma,t}$) on different energy streams ($\sigma \in \mathcal{S}$) at every time step ($t \in \mathbb{T}$), following eq.(5.1).

$$\dot{Q}^{load}_{\sigma,t} = \sum_{\tau_{cv} \in \mathcal{T}_{cv}} (H_{\tau_{cv},\sigma}\ \dot{Q}^{\mathcal{T},in}_{\tau_{cv},t}) + \dot{Q}^{\mathcal{N},in}_{\sigma,t} - \dot{Q}^{\mathcal{N},out}_{\sigma,t} + \sum_{\tau_s\ on\ \sigma} (\eta^{dch}_{\tau_s} \dot{Q}^{dch}_{\sigma,\tau_s,t} - \eta^{ch}_{\tau_s} \dot{Q}^{ch}_{\sigma,\tau_s,t}) \quad (5.1)$$

Two outputs of interest are the annualized cost (eq.(5.2)) and yearly carbon emissions (eq.(5.3)) along 4 scenarios enforcing different carbon emission limits (eq.(5.4)) [67]. Both are related to the system operation, particularly the energy purchases ($\dot{Q}^{\mathcal{N},in}_{\sigma,t}$). The cost also involves the net-present value (NPV) for the anualization, fixed prices (C^{fix}_{τ}), linear prices (C^{lin}_{τ}) and capacities of devices (P^{nom}_{τ}), as well as prices for maintenance (C^{mnt}_{τ}) and carbon emissions (C^{CO_2}). The calculation of emissions involve the carbon factors G_f of the considered sources $f \in \mathcal{F}$.

$$J^{a}_{tot} = \left(\begin{array}{l} \sum_{f \in \mathcal{F}} C^{fuel}_{f}\ (\dot{Q}^{\mathcal{N},in}_{\sigma_f} - \dot{Q}^{\mathcal{N},out}_{\sigma_f}) \\ + \sum_{\tau \in \mathcal{T}} NPV_{\tau}\ (C^{fix}_{\tau} + P^{nom}_{\tau}\ C^{lin}_{\tau}) \\ + C^{CO_2}\ J_{CO_2} \\ + \sum_{\tau \in \mathcal{T}} C^{mnt}_{\tau}\ P^{nom}_{\tau} \end{array} \right) \quad (5.2)$$

$$J_{CO_2} = \sum_{f \in \mathcal{F}} G_f\ (\dot{Q}^{\mathcal{N},in}_{sigma_f} - \dot{Q}^{\mathcal{N},out}_{\sigma_f}) \quad (5.3)$$

The generation of solutions along the carbon limitation scenarios is done in two steps: (1) both solutions for minimum cost and minimum emissions are computed, respectively, by using the annualized cost (eq.(5.2)) and carbon emissions

(eq.(5.3)) as objective functions of 2 different problems. The carbon emission of both solutions are kept as bounds. (2) Then the points constrained to different carbon emission limitations are computed by minimizing the annualized cost while enforcing various upper limits X_{CO_2} on the carbon emissions (eq.(5.4)). The value of these limits are set between the two emission bounds from step 1 (c.f. List in section 5.2.3).

$$J_{CO_2} \leq X_{CO_2} \tag{5.4}$$

The capacity of a converter $\tau_{cv} \in \mathcal{T}_{cv}$ is defined by its maximum generation during hourly operation of the whole system (eq.(5.5)), conditioned by the conversion efficiency $H_{(\tau_{cv},\sigma_{ref})}$. Non-dispatchable converters like PV constantly operate at full capacity given the resource availabilities $\xi^{sun}_{PV,t}$. Regarding the storage, the state of charge (SOC) cannot exceed the device capacity (eq.(5.6)), and the charged/discharged energy within an hour is limited by a fraction $\beta^{ch/dch}_{\tau_s}$ of the storage capacity (eq.(5.7)). The state of charge is propagated throughout time steps following eq.(5.8) and forced to describe a yearly cycle (eq.(5.9)).

$$\begin{cases} 0 \leq H_{(\tau_{cv},\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{n,\tau_{cv},t} \leq P^{nom}_{n,\tau_{cv}}, \\ \qquad\qquad\qquad\qquad \text{if } \tau_{cv} \text{ is dispatchable} \\ H_{(PV,\sigma_{ref})}\ \dot{Q}^{\mathcal{T},in}_{PV,t} = \xi^{sun}_{PV,t}\ P^{nom}_{PV}, \end{cases} \tag{5.5}$$

$$0 \leq SOC_{\sigma,\tau_s,t} \leq P^{nom}_{\tau_s} \tag{5.6}$$

$$\left\{0 \leq \dot{Q}^{ch}_{\sigma,\tau_s,t} \leq \beta^{ch}_{\tau_s} \times P^{nom}_{\tau_s} 0 \leq \dot{Q}^{dch}_{\sigma,\tau_s,t} \leq \beta^{dch}_{\tau_s} \times P^{nom}_{\tau_s}\right. \tag{5.7}$$

$$SOC_{\sigma,\tau_s,t+1} = (\eta_{\tau_s})SOC_{\sigma,\tau_s,t} + \eta^{ch}_{\tau_s}\dot{Q}^{ch}_{\sigma,\tau_s,t} - \frac{1}{\eta^{dch}_{\tau_s}}\dot{Q}^{dch}_{\sigma,\tau_s,t} \tag{5.8}$$

$$SOC_{\sigma,\tau_s,\ t^{max}+1} = SOC_{\sigma,\tau_s,\ t=0} \tag{5.9}$$

5.2.2 Variables of interest and ranges

The system is kept simple for the following reasons; (a) it is generic enough to be suited for diverse building types and locations at early design stage; (b) it facilitates the exploration and understanding of design decisions; (c) it limits the number of design variables of interest, thus more focus can be put towards the surrogates. Despite its simplicity, such a system presents many design choices that may be optimal in some contexts or suboptimal in others. The surrogate modeling task is challenging, even for such a simple case. The choice between technologies may result in splits in the distribution of design variables, which is believed to limit the performances of the surrogate models.

Table 5.1: Surrogate model output space

Variable	Name	Unit	Additional Description
Total cost (J^a_{tot})	*Cost*	k$/y	Sum of expenses (eq.(5.2))
Emissions (J_{CO_2})	*Carbon*	tCO_2eq/y	Sum of emissions (eq.(5.3))
ASHP capacity (P^{nom}_{ASHP})	*ASHP*	kW_{th}	-
Boiler capacity (P^{nom}_{Boiler})	*Boiler*	kW_{th}	-
CHP capacity (P^{nom}_{CHP})	*CHP*	kW_{el}	-
TES capacity (P^{nom}_{TES})	*TES*	kWh	-
TES usage ($\sum_{t\in\mathbb{T}} \dot{Q}^{dch}_{TES,t}$)	*TesUse*	MWh/y	Yearly TES discharge
Electricity ($\sum_{t\in\mathbb{T}} \dot{Q}^{\mathcal{N},in}_{Elec,t}$)	*Elec*	MWh/y	One year of grid import
Gas ($\sum_{t\in\mathbb{T}} \dot{Q}^{\mathcal{N},in}_{Gas,t}$)	*Gas*	MWh/y	One year of gas import

The nine design variables of interest are introduced in Table 5.1. The cost and carbon are commonly used as objective values to define Pareto Fronts in the related literature, and are of primal interest to evaluate the overall system. The capacities of diverse technologies are often the main usage of EHub and indicate how to achieve least cost and emissions. The total energy purchases reveal the degree of electrification of the system, and is a main component of the design process and of the cost calculations. The total usage of storage informs about the overall need to shift the demand for better performances.

The core MES design problem has many parameters describing the context (e.g. costs of technology, price of fuels, conversion efficiencies). Five context variables, found most sensitive over reasonable ranges by previous research [168], are selected to define the input space: the carbon tax, the prices of electricity and gas, the carbon intensities of electricity and gas. The selected ranges of values are typical for a north-american context (Table 5.2). One unique value is considered throughout the whole year for the sake of simplicity. Values for other model parameters not included in the input space are presented in Table 5.3, where *pre* indicates a pre-calculation with hourly resolution.

Table 5.2: Ranges and bounds of the considered input space

Variable	Range	
Price of grid electricity C_{elec}^{fuel}	$[0.15, 0.35]$ $\$/kWh$	[169]
Price of gas C_{gas}^{fuel}	$[0.04, 0.08]$ $\$/kWh$	[170]
Carbon tax C^{CO_2}	$[50, 200]$ $\$/t_{CO_2}eq$	[167]
Carbon intensity of grid electricity G_{elec}	$[150, 700]$ $g_{CO_2}eq/kWh$	[171]
Carbon intensity of gas G_{gas}	$[219, 288]$ $g_{CO_2}eq/kWh$	[172, 173]

For the carbon intensity of gas, Venkatesh et al. [172] and Burnham et al. [173] give LCA-based estimates, ranging at 219-260 gCO2eq/kWh for the former, and either 252-288 gCO2eq/kWh or 324-414 gCO2eq/kWh depending on the time horizon considered (100 years and 20 years) for the latter.

Table 5.3: Fixed parameters of the MES design problem

Device	Life (y)	C^{fix}	C^{lin}	H/η	$\eta^{ch/dch}$	$\beta^{ch/dch}$
ASHP	20	2968	198	*pre*	-	-
Boiler	25	1492	47	97%	-	-
CHP	25	1283	246	35%/45%	-	-
PV	15	0	1020	*pre*	-	-
TES	25	400	99	98%	100%	40%

5.2.3 Data and Context

The energy loads used for this research are taken from a dataset of building energy simulations for different building types across the US [76]. A total of 9 cases are selected out of the original dataset, combining exhaustively the locations and building types in the list below. For each case, the system in Figure 5.1 is designed optimally according to eq.(5.4), with each relative decarbonization objective presented in the list below. With 9 contexts, 4 decarbonization objectives and 9 outputs of interest, a total of 324 independent datasets/experiments are considered.

- **Locations**: Helena, MT; San Francisco, CA; Miami, FL.
- **Building type**: Residential (SFH); Hospital; Office building.
- **Relative decarbonization goal**: 0% (minimum cost); -50%; -90%; -100% (minimum CO_2).

5.2.4 Motivational example

The base model selects technologies and sizes them to match energy demands. In many cases, a design variable of interest may exhibit a split distribution due to that selection. Figure 5.2 illustrates such a phenomenon in the design of an ASHP, through simplified problems using only one, two and three input variables. As shown in Figure 5.2, the split is non-trivial, i.e. it depends on all variables involved in the problem.

The methodology later presented in the current research is based on the hypothesis that ML-based surrogates learn better when the target values exhibit a Gaussian-like distribution, or at least a continuous and balanced distribution. Therefore, the suggested methodology focuses on ways to make the data perceived by the surrogate as balanced and continuous as possible. Mainly, the focus is drawn towards addressing potential splits in the data, with a possible imbalance in the number of samples on either side of a split.

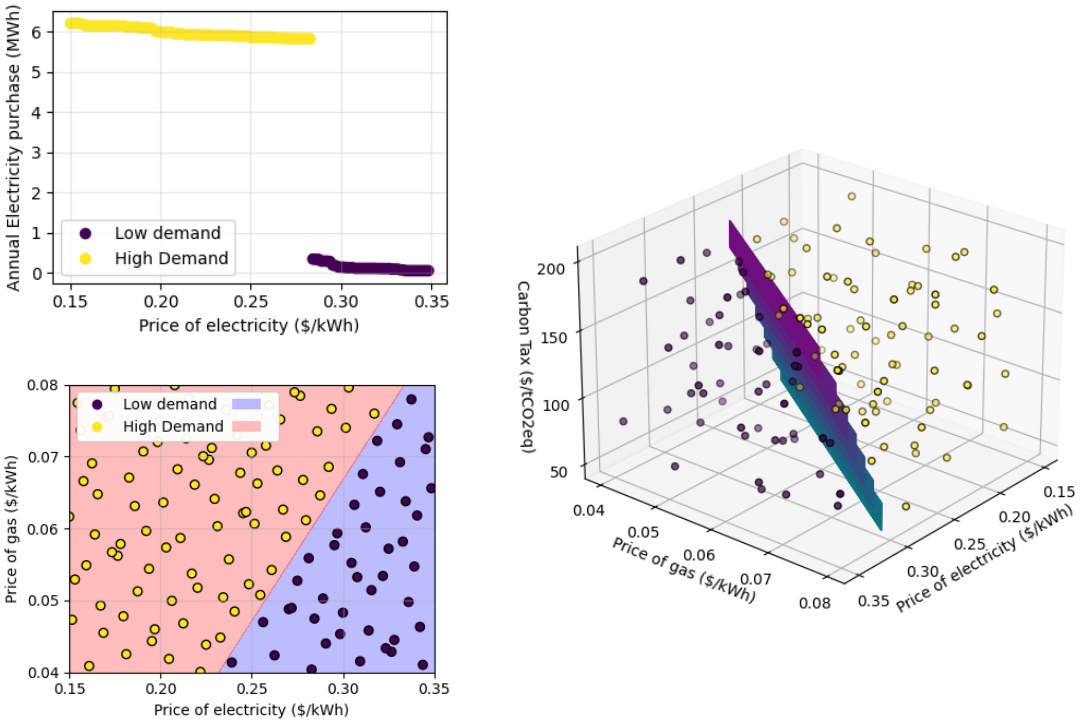


Figure 5.2: Illustration of splits occurring in the design variable on simplified problems with one (top left), two (bottom left) and three (right) variables.

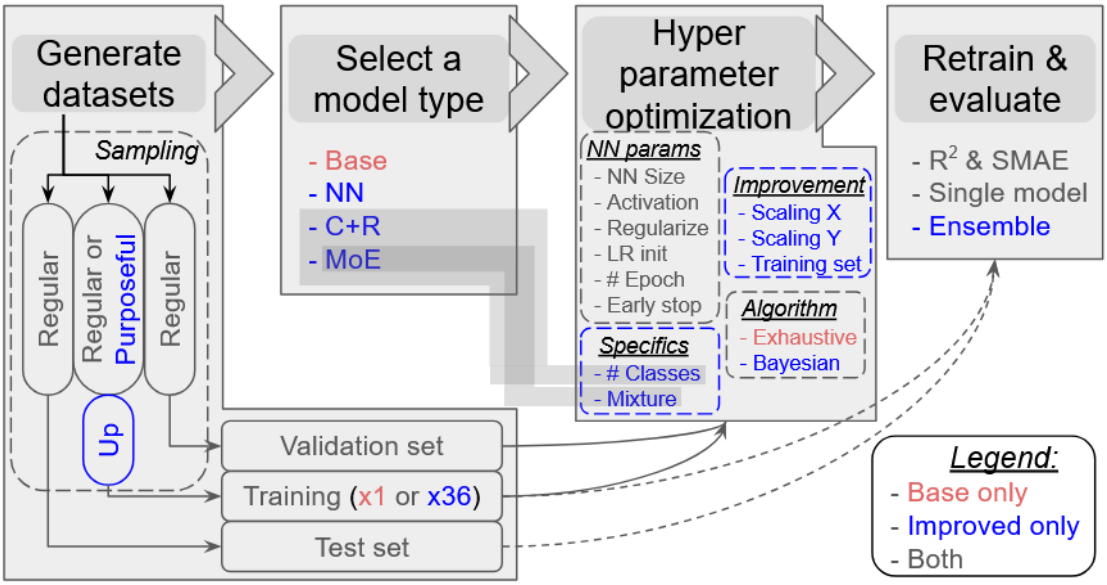


Figure 5.3: Overview of the procedure to train the surrogate models

5.3 Methodology

Figure 5.3 highlights the main steps in the process of building the surrogate models. The general pipeline is straight-forward: data is generated, a specific model type is selected, tuned via hyperparameter optimization (HPO), the final model is re-trained and evaluated. This section introduces a novel sampling step in the data generation (section 5.3.1), the different model types used (section 5.3.2), the range of parameters options during the HPO (section 5.3.3) and the evaluation procedure (section 5.3.4).

5.3.1 Data Generation

The selections of samples are based on a Halton sequence [174]. Each set of input is used to design the energy system via the EHub application. The resulting dataset is split evenly into a training set, a test set and a validation set, all three with 256 samples each. The training set may then be subject to additional upsampling. The choices of sampling sequence and of even size of training-testing-validation sets are discussed in subsection 5.3.1.1. Subsection 5.3.1.2 explains the sampling approach. Subsection 5.3.1.3 introduces the upsampling step.

5.3.1.1 Training, test and validation sets

All samples are drawn with a Halton sequence [174], a low-discrepancy sequence. Preparatory work revealed that the discrepancy of subsequences is more homogeneous than that of the Sobol sequence [175]. Therefore, a training, test and validation sets of similar size and issued from subsequences of the same series guarantee a similar coverage, with similar properties but no repeat.

Similar properties are needed for better training and evaluation. Splits and imbalances are expected in the distributions of design variables (c.f. Figure 5.2), however their occurrence or location is unknown prior to computation of the original application. If the density of samples near a split is insufficient:

- ... in the test set, the performance evaluation will not correctly incorporate the difficulty of the design space, despite modeling efforts.
- ... in the validation set, a suboptimal selection of hyperparameters is likely, and the resulting model may perform poorly on unseen samples near the difficulty.

5.3.1.2 Sampling

A pool of 16 blocks is drawn from a Halton sequence, each block containing 128 points. Blocks #15 and #16 are used as test set, blocks #13 and #14 as validation set. The other blocks may be used for the training set via one of the two following approaches.

The first approach merges blocks #1 and #2 to form the training set. It is the standard approach, used for example in all base cases.

The second approach, illustrated in Figure 5.4, is named "purposeful sampling" in this paper, and only available in the improved approach. It is inspired by the concept of "border patterns" [176] to adapt the purposeful sampling method common in qualitative surveying [177] to the current regression problem. Labels from block #1 are grouped into clusters (with Gaussian Mixture clustering). An artificial neural network (ANN) is then trained to classify samples from block #1 into their assigned cluster. The classifier is called "gating" due to similarities with a model introduced in section 5.3.2. The classifier is then used on blocks #2 to #12 to identify the 128 points least likely to be correctly classified (X_+). Distinctions between 2, 3 and 4 clusters are explored in this paper.

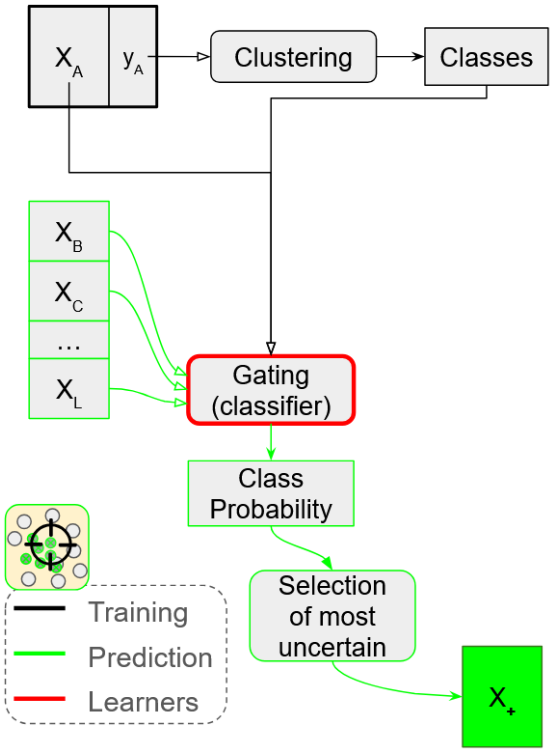


Figure 5.4: Methodology for the purposeful sampling

5.3.1.3 Up-sampling

The training set may be subject to upsampling, in the improved approach only. Upsampling adds synthetic information (i.e. not processed through the original simulation) in specific subregions of the dataset. This is commonly used to enhance the number of samples in underpopulated classes for classifications. The rare usages for regressions flag "exceptional" points, then interpolate between them with various methods.

The hyperparameter optimization chooses whether to perform upsampling, then selects one among eight interpolation methods. These methods are the available options from the Python package `ImbalancedLearningRegression` [178].

5.3.2 Surrogate model types

All regressors and classifiers used in this work rely on ANNs. The base neural network (NN) and fine NN are both ANNs (Section 5.3.2.1). A Classifier-Regressors enhance the input space of a regressor with predictions of a classifier (Section 5.3.2.2). A Mixture of Experts chains a classifier with multiple independent regressors (Section 5.3.2.3). The training and predicted data is rescaled independently

for each ANN composing the different models.

5.3.2.1 Base and Improved NN models

The experiment suggests a base and an improved NN models, each made of one single base ANN regressor. The former is trained via the base methodology, the latter via the improved methodology. Both choose among similar architectures (Table 5.4), but they differ in the number of options for the usable training sets, data preprocessing, a few discrete hyperparameters, and for the algorithm conducting the hyperparameter optimization (c.f. 5.3.3).

Table 5.4: Architecture and corresponding number of neurons on the hidden layers (hl)

	Number of neurons		
Architecture	hl1	hl2	hl3
Small	16		
Shallow	128		
Medium	6	36	6
Large	16	64	16

5.3.2.2 Classification-Regression

A classification-regression (C+R) model [60, 179, 180] chains together a classifier and a regressor, so that the prediction from the classifier is used as additional input information by the regressor (Figure 5.5). First, Gaussian Mixture clustering distinguishes 2, 3 or 4 groups within the training outputs y. A classifier is then trained to predict the class labels from the training inputs X. These inputs X and class labels are also merged to train the regressor to predict the training outputs y. In prediction mode, the class probabilities $\hat{y}$ predicted by the classifier are merged with the input data to inform the regressor, yielding the final prediction $\hat{y}$.

5.3.2.3 Mixture of Experts

A Mixture-of-Experts (MoE) is an ensemble model following a divide-and-conquer approach [181]. It subdivides the input space with a gating unit and delivers more localized predictions with expert units. Such architecture can address the possible abrupt jumps [182] observed in the design variables. The gating, experts, their interconnection and training take different forms in the literature. Any probabilistic model or supervised ML algorithm is suitable as gating and expert

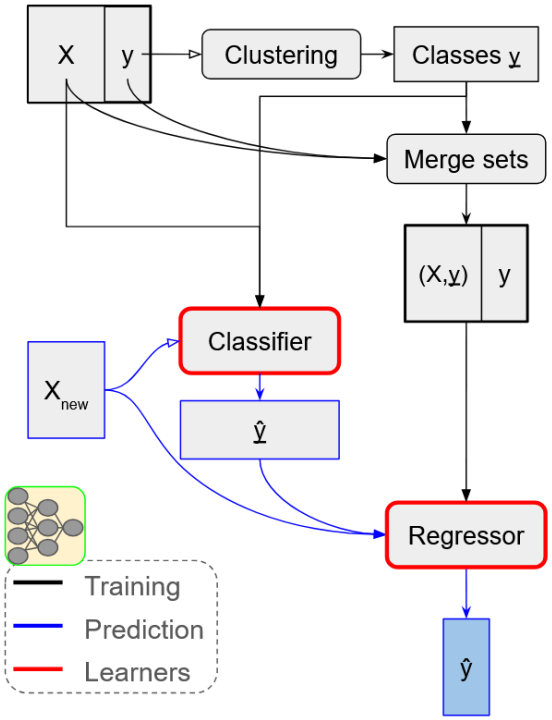


Figure 5.5: Methodology for the Classification-Regression

[181,183]; their training may be separated [183,184] or addressed all together [181], clustering may support the subdivision process [183,185], and the prediction may be handled as a mixture of all experts' opinion [183,184,186] or by the most trusted one in each subsection [187].

The proposed MoE implementation (Figure 5.6) clusters the training outputs y into 2, 3 or 4 groups using Gaussian Mixture [185]. A classifier (i.e. the gating) is trained to infer cluster labels from the training inputs X. The generated classes are also used to segment the training set into subsets, each subset being used to train independently a different expert regressor.

In prediction mode, the set of new inputs X_{new} is used through each expert to generate every expert's "opinion" $\hat{y}^k$, and through the gating to generate a level of trust in each expert (probability $\hat{\underline{y}}^k$ of a sample to belong to each class). As additional hyperparameter, the prediction for each sample may either consider only the most trusted expert, or consider an average of all experts' opinions, weighted by the relative probabilities.

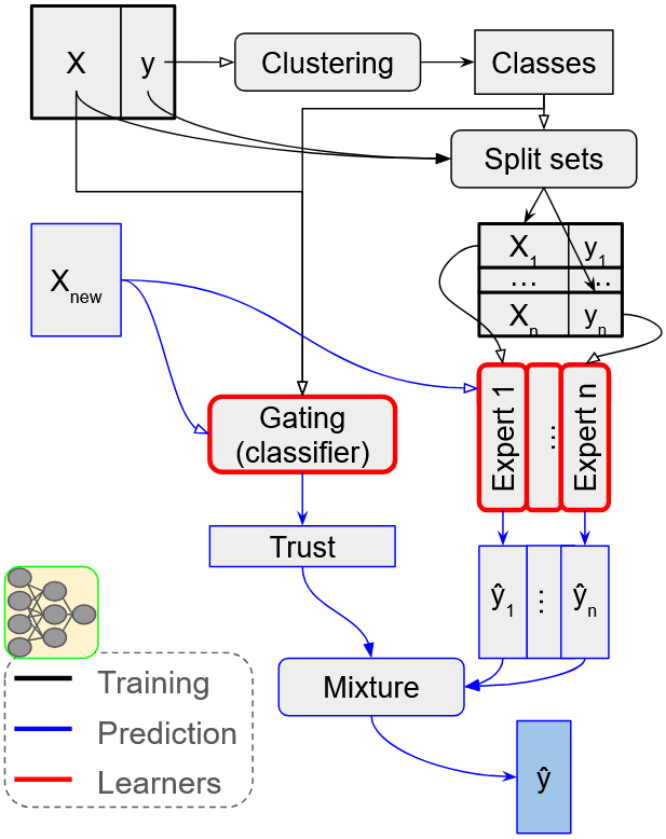


Figure 5.6: Methodology for the Mixture-of-Experts

5.3.2.4 Ensembles

Known as a robust solution in machine learning problems, training an ensemble of neural networks and averaging the prediction perform better on numerous machine learning problems [188, 189], particularly when dealing with small datasets.

Every model is trained and evaluated four times as part of the evaluation procedure, a last step of the procedure thus evaluates the averaged prediction of all four models. For regressors, the prediction of each sample is averaged (eq.(5.10)). For classifiers, the predicted probability of each class is averaged (eq.(5.10)) before the class estimated as most likely is considered as the final class prediction.

$$\hat{y}_\Omega = \frac{1}{4}\sum_{i=1}^{4}\hat{y}_i \tag{5.10}$$

5.3.2.5 Data scaling

Scaling data helps training ANNs effectively. Standardization, i.e. subtracting the mean and dividing by the standard deviation, is commonly used [190]. Logarithmic [191] or power transformations (e.g., box-cox [155,192]) may improve performances with skewed or imbalanced data, by reshaping it into a more Gaussian-like distribution. Quantile transformation [193] does the same by working on quantiles. This non-parametric approach is particularly suited to surrogate modeling, where ranges of values are fixed. It is therefore preferred to the power transformations in this study.

In this study, the improved methodology compares standardization with quantile transformations creating normal and uniform distributions. Inputs (X) and outputs (y) of each base learner can be scaled differently and independently. For example, a MoE can apply distinct scaling for the gating function and for each expert. Only standardization is used with the base methodology.

5.3.3 Hyperparameter optimization

Hyperparameter optimization (HPO) is a process that aims at selecting the combination of hyperparametes to maximize the model performances. Different methodologies exist to perform this task. An exhaustive search is used for the base model (i.e. all combinations of available options for each hyperparameter are compared). Bayesian optimization [194] is used for the improved models. For the non-common hyperparameters (e.g. architecture, scaling, sampling, upsampling) the HPO algorithm selects the value of an intermediate categorical variable, this value is later interpreted to refer to the correct action in the problem build up. The implementation from `scikit-optimize`[1] is used for the Bayesian HPO, and these from `scikit-learn`[2] for the exhaustive search.

Table 5.5 lists all sets and ranges of options to select from during the HPO[3]. The check mark ✓indicates that the option, set or range is considered entirely. A value indicates that other options are discarded. A ~~struck~~ value indicates that all but this option are considered. Specifically for MoE, the asterisk * indicates that the option, set or range is considered for each expert independently. The double-asterisk ** indicates that separate models were trained and evaluated for

[1] `https://scikit-optimize.github.io`

[2] `https://scikit-learn.org/`

[3] The classifier used for sampling (Fig.5.4) also undergoes a HPO, similar to these of the Base model, with only the *small* architecture.

each option. Similar classifiers (Clf) are used for the C+R and for the gating of MoE. They undergo a HPO separately from the regression part, but are presented together in Table 5.5.

Table 5.5: Hyperparameters and ranges

Name	Options/Range	Base	NN	C+R	MoE	
				Reg.	Clf.	Reg.
Neural Network hyperparameters						
Architecture	Small				✓	*
	Shallow	✓	✓	✓	✓	*
	Medium	✓	✓	✓	✓	*
	Large	✓	✓	✓		
Activation	{Linear, ReLu, TanH}	✓	✓	✓	✓	*
L2 regularization	$log(\alpha) \in [-5, -1]$	✓	✓	✓	✓	*
Learning Rate Init	$log(lr0) \in [-3, -1]$	✓	✓	✓	✓	*
Epochs	{200, 1500, 5000}	1500	✓	✓	✓	~~5000~~*
Early stopping	{8, 16, 32, 64}	32	~~8~~	~~8~~	~~8~~	~~64~~*
Tolerance	10^{-5}	✓	✓	✓	✓	✓
Solver	Adam	✓	✓	✓	✓	✓
Data scaling						
Input scaling	Standard	✓	✓	✓	✓	*
	Quantile $\mathcal{U}$		✓	✓		*
	Quantile $\mathcal{N}$		✓	✓		*
Output scaling	Standard	✓	✓	✓	✓	*
	Quantile $\mathcal{U}$		✓	✓		*
	Quantile $\mathcal{N}$		✓	✓		*
Sampling and data generation						
Number of classes	{2, 3, 4}			✓	**	**
Purposeful sampling	{No, 2, 3, 4}	No	✓	✓	✓	*
Upsampling	{Yes, No}	No	✓	✓	✓	*
Specific to the Mixture-of-Experts						
Weighted opinions	{Yes, No}					✓
Hyperparameter Optimization algorithm						
HPO	{Exhaustive, Bayesian}	Ex	B	B	B	B

5.3.4 Evaluation

5.3.4.1 Metrics

The evaluation of classifiers uses the accuracy metric, i.e. the ratio of correctly classified samples in the set. The evaluation of regressions use two unitless metrics: a standardized mean absolute error (SMAE) (eq.(5.11)) and the coefficient of determination (R^2) (eq.(5.12)), used as reference in all related literature [44, 59, 60, 167]. The error metrics involve the predicted value $\hat{y}_i$, the true value y_i and the average of the true values in a dataset $\bar{y}$.

$$SMAE = \frac{\frac{1}{n}\sum_i |y_i - \hat{y}_i|}{max(y) - min(y)} \tag{5.11}$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{5.12}$$

The SMAE is standardized with the range of target values (eq.(5.11)). The standardization was not done with the average, as the average of some strongly imbalanced datasets approach zero, magnifying the error. The MAPE (percentage error for each sample) is discarded, as most cases contain samples with zero values. The SMAE as described in eq.5.11 appeared as the best metric to ensure comparability between all cases (c.f. List in section 5.2.3). Moreover, the SMAE is easily interpretable, and simply converted into the MAE metric, interpretable in the physical unit of the predicted variable.

5.3.4.2 Evaluation procedure

Each model is trained and evaluated four times at each iteration of the HPO and during the final evaluation. This aims to limit the variability induced by the random initializations [59].

During the HPO, the robust performance measure sums the average and standard deviation of the SMAE for regressions (eq.(5.13)) and of the accuracy for classifications (eq.(5.14)) on the validation set.

The final reported performance are the average values of each metric on the prediction of all four models on the test set. Additionally, the prediction of the ensemble (c.f. Section 5.3.2.4) is evaluated similarly.

$$score_{HPO,reg} = average(SMAE) + std(SMAE) \tag{5.13}$$

$$score_{HPO,clf} = average(Accuracy) + std(Accuracy) \tag{5.14}$$

5.4 Results

The benefits of the novel approach are presented in three parts. Section 5.4.1 covers the performance across all cases. Section 5.4.2 illustrates the effect of the approach via specific examples. Section 5.4.3 validates the approach in regards to the initial hypothesis.

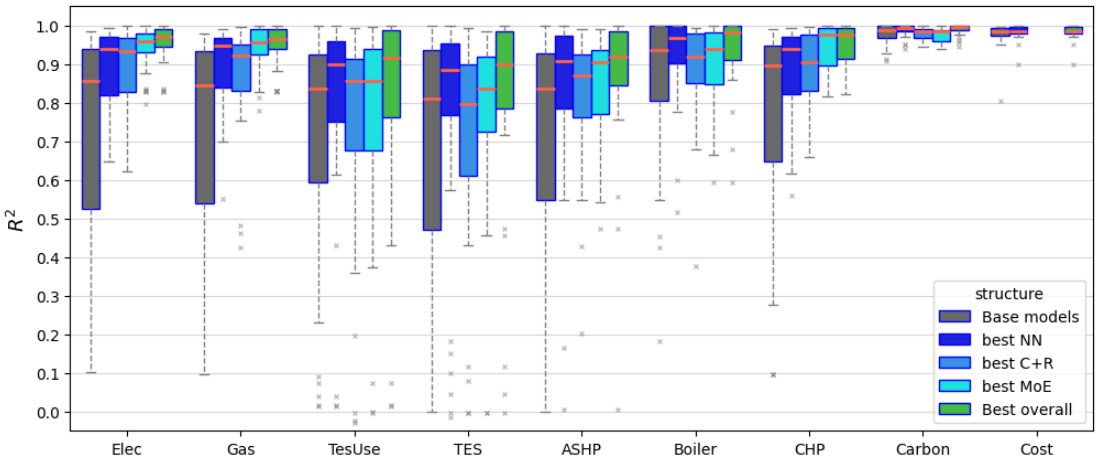
5.4.1 Does it work?

Figures 5.7a and 5.7b compare the performance of the base model, each of the three architectures and the overall best improved model for all considered cases and variables. Note that only the NN structure was considered for the cost during the improved process, due to already strong performance of the base model.

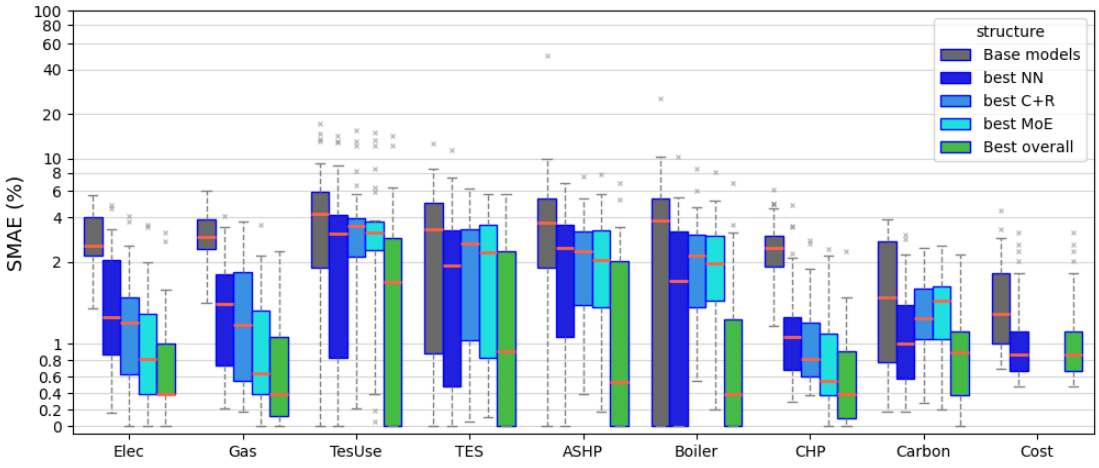
Overall, surrogates of all three structures consistently outperform base models, with only a few cases showing no improvement. Selecting the best among the three model types for each case shows significant performance improvements over each individual model type over the pool of cases. There is no single best architecture across all cases.

The R^2 metric highlights the strong performance of base models on Cost and Carbon (median ≥ 0.99). It also shows lower medians (between 0.8 and 0.93), with lower quartiles around 0.5 for other variables. This confirms the large variability of performance observed in related literature, since in 25% of cases, the model is only able to explain 50% of the variance in the data, leaving 50% unexplained. The improved approach significantly increases performance; for the best among all model types, medians rise above 0.9 for ASHP and storage-related variables, above 0.96 for others, and the 75^{th} percentiles rise above a R^2 of 0.75 (Figure 5.7a). The MoE seems to perform overall better than other model types for Elec, Gas and CHP, and the NN better for the Boiler, Carbon, Cost and storage related variables. However, the best among three model types is significantly better than each individual type.

The SMAE further increases the impression of clear superiority of the best-



(a) Coefficient of determination R^2



(b) Standardized Mean Absolute Error SMAE

Figure 5.7: Performance improvement of each architecture and with the improved methodology

among three over each individual model type. The base models show a median between 2% and 4% SMAE for most variables. It slightly exceeds 4% for TesUse and lies between 1% and 2% for Cost and Carbon. The improved process drops the median SMAE below 1% for all variables but TesUse (1.7%), with the 75^{th} percentile only exceeding 2% SMAE for the ASHP and storage-related variables (Figure 5.7b).

Despite clear improvements with the novel approach, performance variability remains, particularly for the ASHP capacity and storage-related variables. Moreover, the metrics show similar trends but greatly differ in intensity. For example, the performance on Cost and Carbon stand out with the R^2 metric, while they do not among improved models with the SMAE.

Figure 5.8 shows the frequency of usage of data preprocessing, alternative

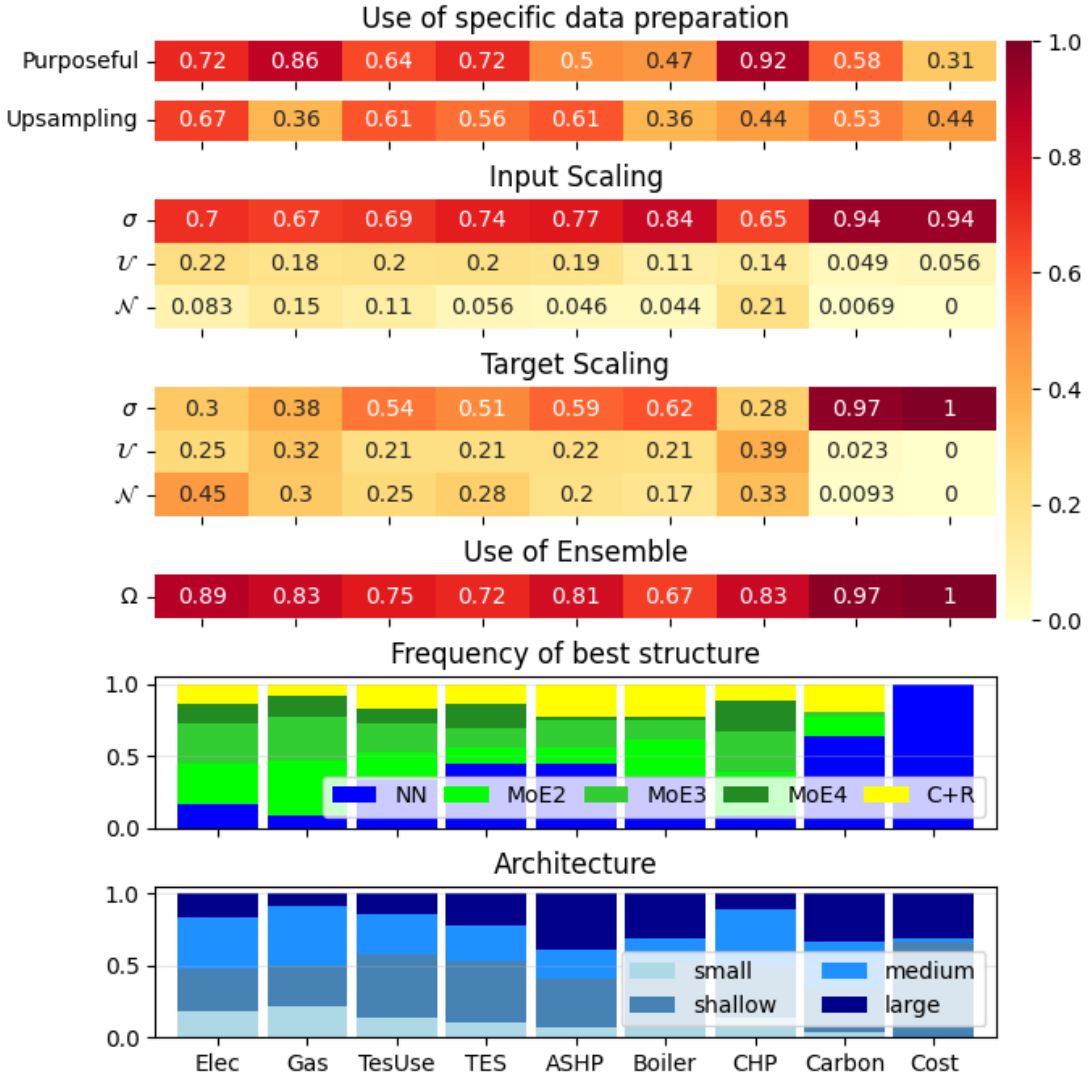


Figure 5.8: Frequency of usage of data sampling, processing and model types among the best performing improved models.

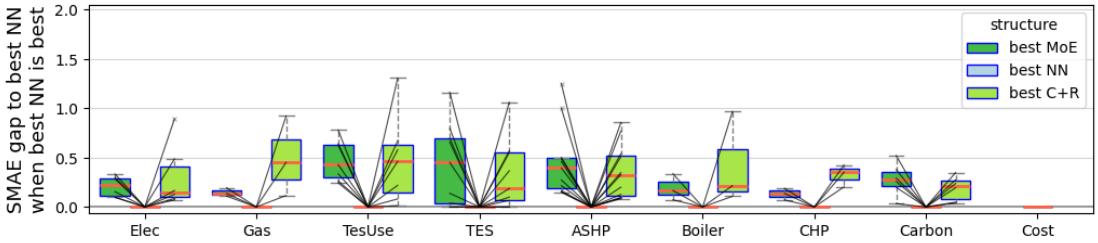
sampling and model structures among the best performing improved models.

Standard scaling (ϵ) is largely preferred to uniform ($\mathcal{U}$) and normal ($\mathcal{N}$) scaling through quantile transform for the inputs and for Cost and Carbon target variables. Frequent usage of all three approaches for other target variables suggest that the choice of scaling depends on each specific dataset.

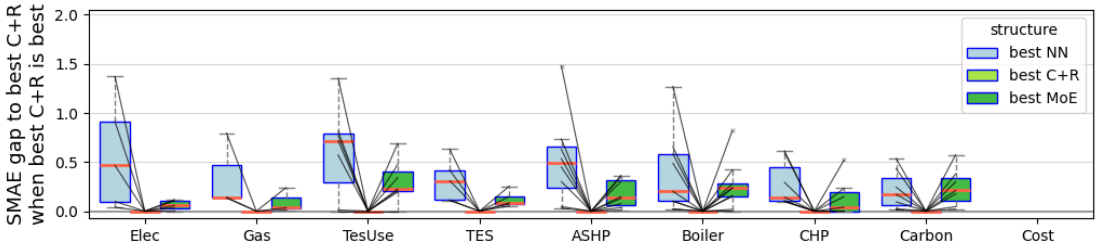
Purposeful sampling is used for all variables, the least frequent being for the Cost. Upsampling is also frequently used, although less than purposeful sampling. Both appear as mostly beneficial, however without guaranteeing optimal performance for all datasets.

The model preference is highly variable specific. MoE is preferred for energy

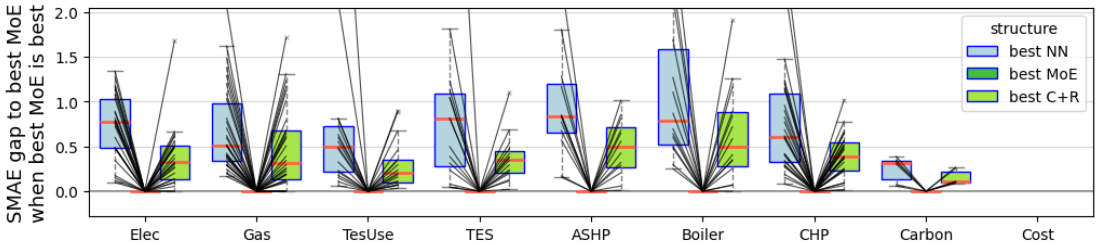
purchases and CHP capacity, while the fine NN is preferred for the already well-predicted Cost and Carbon variables. Both models are used with comparable frequencies for the ASHP, Boiler and storage variables. The C+R is never preferred for the Cost and Gas, and is otherwise used in 10%-20% of cases, but never as the predominant model.



(a) Performance gap to NN when NN is the preferred structure (128 cases)



(b) Performance gap to C+R when C+R is the preferred structure (51 cases)



(c) Performance gap to MoE when MoE is the preferred structure (145 cases)

Figure 5.9: Difference in performance between the overall best and the best of each architecture.

Figure 5.9 quantifies how superior each model type is on each case, detailed per model type and variable. Where NN is the overall best (Figure 5.9a), the difference with MoE and C+R remains low ($< 0.5\%$ SMAE median, 25^{th} precentile close to zero). When C+R is the overall best (Figure 5.9b) the median difference with MoE is below 0.25% SMAE, larger variability and a median difference up to 0.75% SMAE is found with the NN. When MoE performs best (Figure 5.9c), the median difference with the best NN exceeds 0.5% SMAE, with 75^{th} percentiles consistently between 1% and 1.5% SMAE. The difference is less with the best C+R but remains significant. Overall, the MoE structure appears as the most reliable, showing significant benefits over peers when performing better (Figure 5.9c) and a moderate disadvantage otherwise (Figures 5.9a&5.9b).

Overall, all of the suggested architectures and methodological improvements significantly outperform the base case on capacity variables needed for system design. Selecting the best-of-three improved model gives the highest performances. If only one architecture is to be chosen, the Mixture-of-Experts (MoE) model performs the best, achieving median results of within 0.5% SMAE of the other architectures.

5.4.2 How does it work?

The overall experiment evaluates the base model, three architectures and the best overall models through 324 cases. Figures 5.10-5.14 show predictions on five cases where the novel methodology, and particularly MoEs, shows clear improvements. Each figure compares true and predicted values. The left-hand graph shows the sorted true values and their corresponding predictions. The right-hand graph shows the true versus predicted values. Colored halos indicate the MoE-based clustering; samples misclassified by the MoE are marked and corrected in red. Performance measurements are included in a table, including the average relative error induced by misclassified samples (ErrAvg%) and performance if all samples were correctly classified (Perfect).

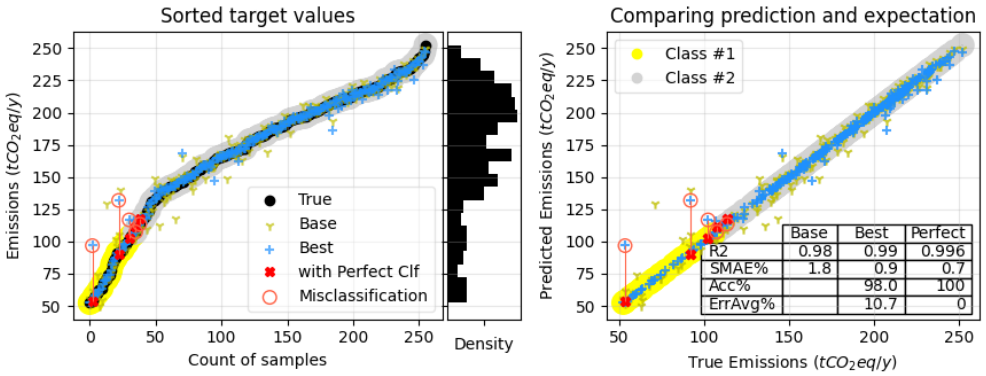


Figure 5.10: Carbon emissions output for a MediumOffice in SanFrancisco at -50%CO_2

For variables exhibiting continuous distributions (Figure 5.10), like the Cost and Carbon, the base NN already gives accurate predictions. In this example, the MoE structure slightly outperforms the base NN. It further halves the already low SMAE, despite misclassifications with minor consequences due to the continuity of the distribution.

Figure 5.11 shows a case exhibiting a clear discontinuity, which is well captured by the MoE structure. The improved model mainly addresses the errors of the

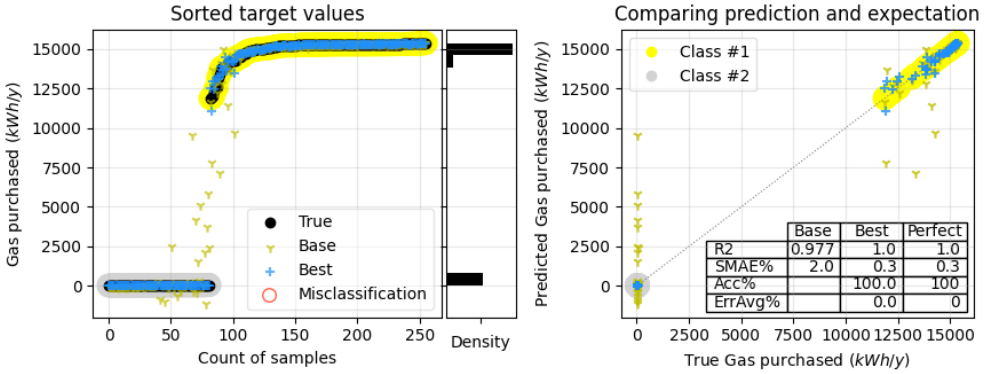


Figure 5.11: Gas output for a Hospital in Miami at minimum cost

base NN near the transition between the two data segments. Implausible negative predictions from the base NN are also addressed.

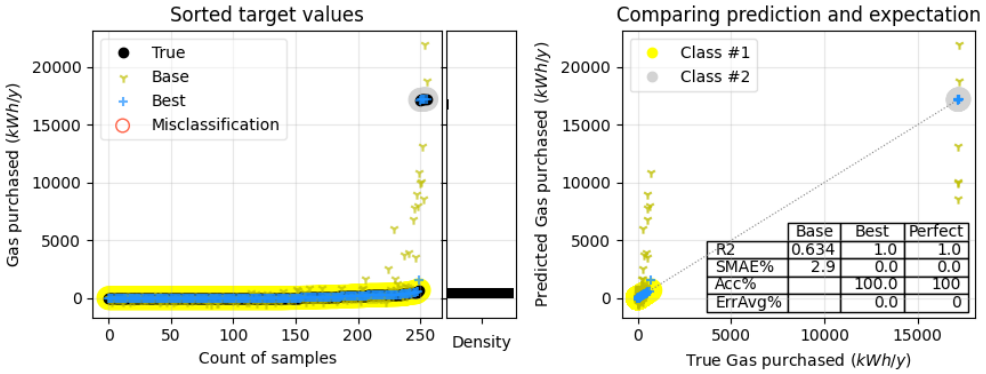


Figure 5.12: Gas output for a hospital in Helena at minimum emissions

The case in Figure 5.12 shows a clear discontinuity with strong imbalance between classes. A MoE, via the improved methodology, perfectly predicts the variable where the base NN shows errors similar to the case in Figure 5.11. Through these two cases, the R^2 metric is inconsistent; two drastically different scores are obtained for the base NN despite similar behaviors and SMAEs in both cases.

Figure 5.13 combines both a split and some marginal behaviors. A MoE is best suited, halving the SMAE achieved by the base NN. The MoE dedicates two experts to capture marginal behaviors, and two more to distinguish the two main segments. Misclassifications remain between the two main segments, but their relative proximity limits the influence on the performance.

Figure 5.14 shows two major splits in the data distribution, for which the base NN handles the transitions poorly. The MoE halves the SMAE, however few misclassified samples strongly impact the performance. While the structure of the data suggests using 3 experts, a model with 2 experts was preferred, performing

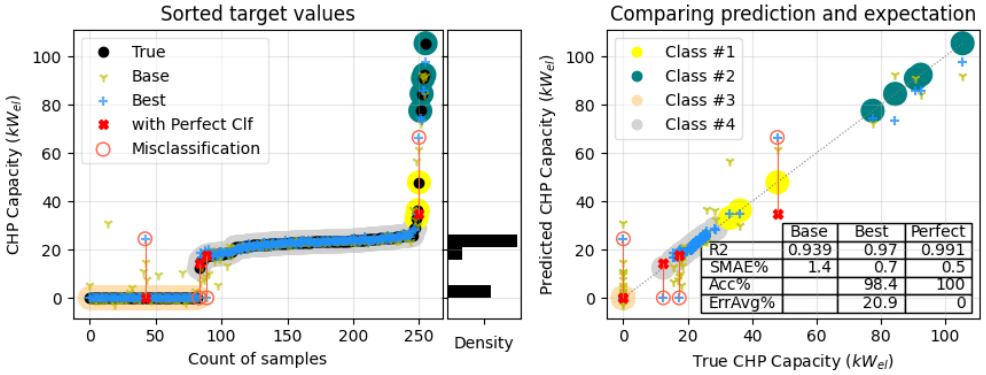


Figure 5.13: CHP capacity for a Medium Office in SanFrancisco at -50%CO_2

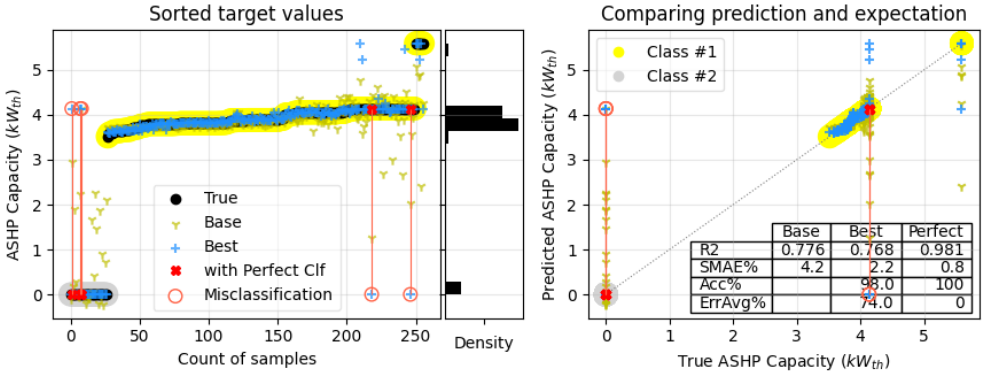


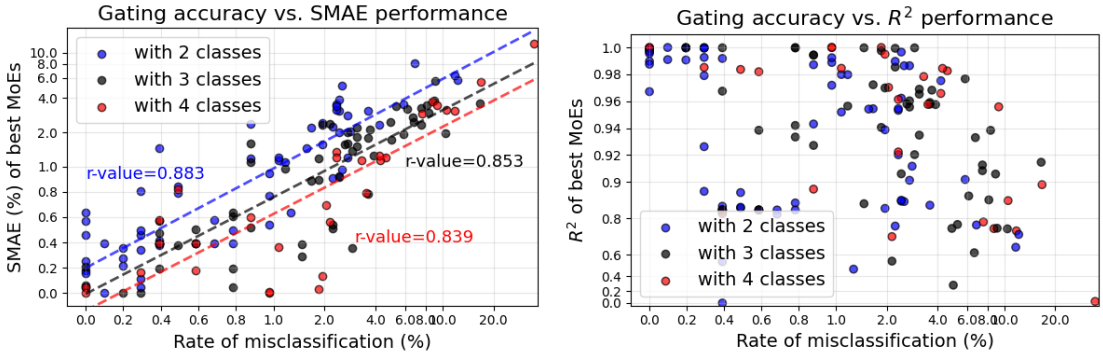
Figure 5.14: ASHP capacity for a single-family house in SanFrancisco at minimum cost

better particularly due to a better classification. Figure 5.14 also shows how R^2 can mask poor predictions on split data distributions, as 2% misclassification on the MoE is valued equally to a base NN capturing the data structure incorrectly.

These examples illustrate the superiority of MoE structures and the improved methodology over the base NN, particularly in handling splits and marginal behaviors. The MoE's ability to segment the data and suggest split predictions addresses the trends of base NNs to smooth out the transition between classes.

The examples also suggest that MoEs are limited by misclassifications, as each sample may yield large errors strongly impacting performance. Figure 5.15a confirms this trend globally, revealing a correlation with the SMAE. Interestingly, identical correlation slopes reveal slightly better MoE performance with more experts, while simultaneously reducing the strength of the correlation slightly.

These examples also illustrate how the R^2 metric often inadequately reflects model performance on data exhibiting splits. The absence of trends between the R^2 performance and misclassifications observed in Figure 5.15b, while clear trends



(a) SMAE performance

(b) R^2 performance

Figure 5.15: Link between global performance and misclassification

are obtained with the SMAE, confirm this.

5.4.3 Why does it work?

The hypothesis behind this work suggests that ANN surrogates may perform better with normally or smoothly distributed target values. Figure 5.16 displays results of a clustering process grouping cases by similarity of distribution. The clustering is applied to the sorted target values of base cases, then improved cases are assigned to these clusters after being transformed as during the training phase. The full clustering procedure is described in Appendix E.

The left column of Figure 5.16 shows the distribution of SMAE achieved by the base and the best improved models on each case of a cluster. Overlaying points show the SMAE achieved for each, distinguishing base and improved models, sorted vertically by decreasing performance, and aligned horizontally on the SMAE.

Nine clusters were identified. Cluster 0 stands out, by capturing one-third of all base cases but two-thirds of the improved cases. It shows a smooth distribution of values, with a peak of density between 0% and 1% SMAE, and rapid decline as the error grows. It shows that the novel approach pushes cases towards this cluster, increasing the chances of good performance.

Clusters 2, 3 and 4 represent cases with split distributions. Two-thirds of the base cases are reassigned to other clusters. These distributions are challenging for surrogates, as the SMAE spreads densely up to 5% in clusters 3 and 4, and peaks at 3% in cluster 2. The remaining improved cases show the best performance in their clusters.

Clusters 5 and 6 represent cases displaying marginal behaviors, and show SMAEs spreading densely up to 6%. Of the 68 base cases, the 20 remaining after improvement are mostly MoE models able to capture marginal behaviors adequately and perform well.

Cluster 1 represents cases with continuous yet heavily skewed distributions and a large spread of error. 50% fewer cases are assigned to Cluster 1 after improvement, most of them with an SMAE below 1%, although some remain with errors within 3%-5%. This confirms that the heavy skewness makes cases challenging but that the novel approach may correct it.

Table 5.6 shows the distribution of base and improved cases across clusters for each target variable. Cost and Carbon are entirely in Cluster 0 from the base cases, explaining their strong initial performance. For all other variables, the novel approach increases the number of cases assigned to Cluster 0. This supports both the hypothesis of this research and the increase in performance observed in Section 4.1.

For base cases, most variables are concentrated in specific clusters: ASHP and TesUse in Cluster 1, Boiler in Clusters 0 and 7, and CHP in Cluster 6. TES spreads across Clusters 0, 1, 6, and 7. This shows most variables have their own specific distribution.

Interestingly, at least a third of improved cases assigned in Cluster 1 are newly assigned, coming from clusters with marginal behaviors (Cluster 5 for Elec, Cluster 6 for CHP). This indicates that the novel approach may result in heavily skewed distributions for cases containing a few marginal behaviors.

Table 5.6: Count of cluster assignments per variable.

Cluster	0		1		2		3		4		5		6		7		8	
Method	Base	Improved	Base	Improved	Base	Improved	Base	Improved	Base	Improved	Base	Improved	Base	Improved	Base	Improved	Base	Improved
Elec	4	25		2	12		3	2		4	17	1						2
Gas	1	25	2	2			1		14	2		2	18	5				
TesUse	12	24	8	1	1		1	1	1	2	2		2	1	5	6	4	1
TES	10	20	8			1			3	2			7	6	7	7	1	
ASHP	16	25	6	2	1		1	1	1		1		3	2	5	5	2	1
Boiler	16	21	2	2		2	3		3						10	10	2	1
CHP	3	24	3	6			2	2	6	1			18	3			4	
Carbon	36	36																
Cost	36	36																

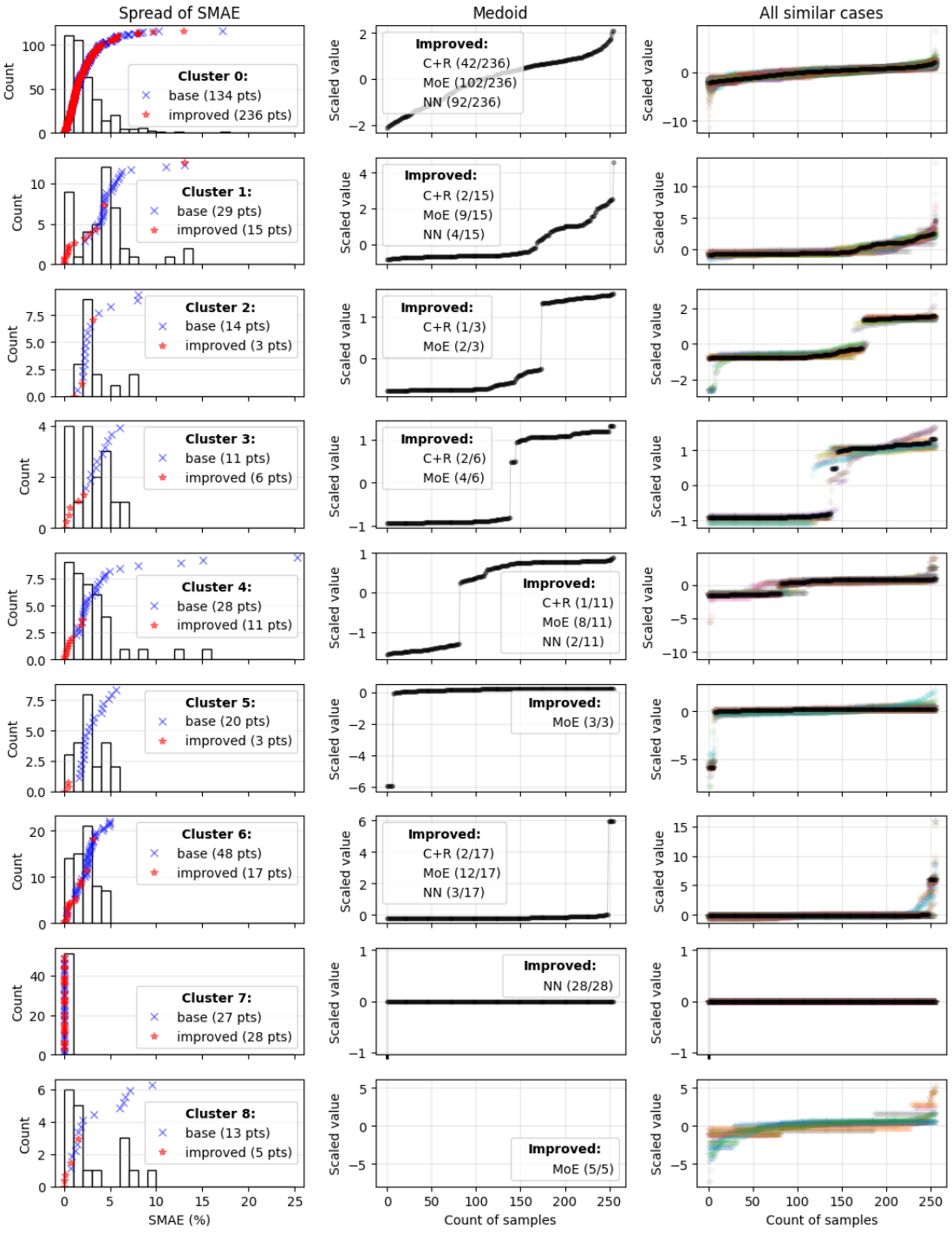


Figure 5.16: Sorted values of design variables and scores of all cases, clustered by the distribution of the scaled output value across the cases. The left column shows histograms of performance for assigned cases. The cluster medoids (i.e. most typical case of the cluster) are shown separately in the middle column and with all cases in the same cluster in the right column.

5.5 Conclusions

This study proposes a methodology to increase performance of surrogate models to support the design of multi energy systems (MES). The methodology combines different data preprocessing, sampling and model structures. It is tailored to address the main current difficulties, i.e. design variables often exhibiting split, skewed or extremely imbalanced distributions. These arise from the underlying physical problem: minor variations of the input may cause abrupt operational changes engendered by a sudden selection, exclusion, or combination of specific technologies.

The improved methodology is evaluated over 324 cases across diverse design variables and compared to the standard approach from literature. Results show that the improved methodology consistently outperforms the standard approach, particularly in cases with challenging data distributions. Supported by a specific sampling, upsamping, data scaling and architecture, the improved methodology turns the data distribution perceived by learning components into smoother and better balanced distributions. The best choice of surrogate model architecture is case dependent, but a mixture-of-experts (MoE) is most reliable overall. Particularly, the MoE successfully identifies data splits and marginal behaviors, where the standard architecture struggles. This is due to its particular architecture: a gating unit subdividing the datasets, possibly on irregularities, for expert units to learn the smoother data distribution of each segment separately.

A key limitation of the improved methodology lies in the possible misclassifications from the gating unit of MoE structures, which may strongly impact the overall reported performance. The experiment also highlights the inadequacy of the coefficient of determination R^2 to reflect model performance on data exhibiting split and imbalanced distributions.

Future efforts may improve this methodology and extend the application of surrogate modeling for MES to broader contexts. First, the gating of MoEs needs more attention to ensure the highest reliability and robustness of surrogates. Other ensemble methods could be investigated, including expanding the MoE approach to encompass the larger neural network and classifier-regressor architectures used here that sometimes outperform the current MoE. Next, future efforts should focus on expanding the application of each surrogate to more variables and contexts (e.g. across multiple building types, environmental and policy conditions). Similarly, the Mixture-of-Experts and other ensemble methods could be applied to other areas of surrogate modelling that exhibit similar issues with split or imbalanced data distributions. Finally, investigating the interconnection between surrogate models

could also extend the modeling capabilities to community-level and beyond. Such improvements could play a crucial role for energy planning and decarbonisation strategies on a broader scale.

Chapter 6

Conclusion

6.1 Synthesis

This thesis explores the following question: *How can methods for time-series aggregation and surrogate modeling be developed and evaluated to improve the accuracy, efficiency, and reliability of multi-energy system design optimization for buildings across diverse contexts?*

This dissertation investigates the possible use of representative days and surrogate modeling for design of multi-energy systems (MES). Particularly, it focuses on enhancing the efficiency and reliability of methods, while maintaining design accuracy across various contexts. The work is structured into four independent chapters, each developing different aspects of this challenge.

Chapter 2 evaluates the robustness of various methods for selecting representative days for the complexity reduction MES design problems. The evaluation is conducted methodically through diverse contexts, including the system complexity, geographical and techo-economic conditions, to evaluate the transferability of findings across case studies. The analysis reveals several key insights. The performance of methods sensitive to contexts, meaning that method demonstrating strong performances in a context may fail in others. This emphasizes the need to validate methods in context-specific manners and clearly indicate their range of application. In addition, the study shows that the quality of representation of the MES design objective function (often the cost) does not necessarily correlate with the quality of representation of other design aspects, like the system configuration. These findings highlight the need for systematic and broader assessments of the capabilities of selection methods across diverse contexts.

Chapter 3 examines the effect of considering shorter time horizons on the MES design decisions at the building scale. The results show that the operational control remains consistent when considering shorter time periods. On the contrary, the system design is strongly affected, and particularly the design of storage units. Shorter time horizons increase the sensitivity of designs to extreme events, often resulting in component oversizing. In addition, the study of decisions over shorter horizons offers novel perspectives in the understanding of design decision mechanisms in the application. The approach also provides new perspectives regarding the understanding of design decision mechanisms and the influence of temporal dimensions on system configurations.

Chapter 4 compares energy systems designed around two state-of-the-art district thermal network (DTN) technologies: 4th and 5th generation heating and cooling networks (4G and 5G). The study demonstrates the superiority of systems relying on 5G DTN for both financial and environmental aspects. This difference is mostly explained by the specific ability of the 5G DTN to balance simultaneous heating and cooling demands. The performance gap remains favorable to the 5G DTN through a wide set of scenarios. However, both the performance gap (and even the choice of technology) shows a sensitivity to the topology and spatial distribution of the thermal demand. Both DTN models are formulated using simple energy transfer principles, based on state-of-the-art methods in the literature. The formulation allows to incorporate the DTN at the heart of a broader energy system design. The formulation differs from most literature comparing 4G and 5G DTNs, preferring model based on flow and temperature control.

Chapter 5 explores the use of surrogate models to accurately predict the MES design. The key contribution is the development of a novel methodology tailored to address specific difficulties inherent to the optimization-based design of MES. The methodology combines data preprocessing techniques, an innovative sampling strategy and tailored neural network architectures. The Mixture-of-Experts architecture is found particularly effective to handle skewed, imbalanced and irregular data distributions. The methodology is rigorously tested across a wide set of contexts, systematically demonstrating strong performance improvements over traditional approaches. This chapter demonstrates the potential of surrogate models to address the design of MES. It also demonstrates the importance of jointly understanding of the decision process and complexity reduction techniques to develop efficient and context-specific solutions.

Together, the chapters converge on the three following points. First, they highlight the importance of validating proposed methodologies across diverse contexts to ensure their generalizability and facilitate broader adoptions. Second,

they emphasize the need to assess the sensitivity of suggested methods to better characterize their reliability and facilitate the understanding of their underlying decision processes. Finally, they underscore the need of a joint understanding of both the decision mechanisms for the core energy system design application and for the suggested complexity reduction methodology to effectively reduce computational burden while maintaining accuracy.

Each chapter contributes to addressing the central objectives of the thesis. Together, they propose evaluations of reliability of model results, understanding of decision processes and development and implementation of time reduction approaches.

- *Objective 1: Understanding the impact of context for the evaluation of accuracy and efficiency.* Chapter 2 demonstrates the context-dependent nature of methods selecting representative days. It contributes to an overall understanding of the transferability of findings across case studies. Chapter 3 illustrates novel perspectives about design decision mechanisms across multiple climates and temporal frames. Chapter 4 compares systems across various scenarios, informing about the context-induced diversity of responses. Chapter 5 develops a new context-resilient methodology by exploiting commonalities of the core application across contexts, and highlights the importance of defining adequate metrics to measure performance.

- *Objective 2: Providing a deeper understanding of decision processes.* Chapter 3 provides new insights about mechanisms underlying MES design decisions, specifically on the the influence of temporal dimensions on system configurations. Chapter 4 illustrates and details a process for the selection between two thermal network technologies. Chapter 5 demonstrates the effectiveness of a methodology issued from deeper understandings of the decision processes.

- *Objective 3: Assessing the importance of modeling details.* Chapter 4 shows the importance of the accurate representation of specific properties of technologies in their comparison and in the study of synergies within energy systems. Chapter 2 shows the possible sensitivity of complexity reduction methods to even minor changes in the system modeling.

- *Objective 4: Leveraging the use of machine-learning.* Chapter 5 Suggests an effective methodology, and demonstrates its reliability across contexts.

The thesis corpus highlights intricacy between design decisions, their contexts and solutions for effective reduction of computation burden. The context of use

is of foremost importance and needs careful specification. Thus the thesis calls for the development of more systematic frameworks to support extrapolations between contexts, as diversity is a main concept driving the decarbonization and energy transition.

6.2 Reflections and future research directions

This thesis makes significant advancements in the field of energy system design optimization, but several limitations must be acknowledged. Addressing these limitations represents opportunities for future research.

While the thesis emphasizes the importance of conducting extended validations over large contexts, the number of degrees of freedom for the design of multi-source energy systems in general is extremely large. As Chapter 2 demonstrates the need for validation across the expected range of usage of a model, the findings show no guarantee of generalizability to strongly different case studies. Specifically, strong choices were made regarding the selection of technologies, data, and parameters used in the validation. Justified by the building scale on which most of the work is focused, wind turbines are excluded desipte their important place in current and future energy landscapes [3, 195]. More specific applications often focus on the integration of precise technologies [103, 196] (as done in Chapter 2 with GSHPs and Chapter 4 with thermal networks). These emphasize naturally more on the modeling of these specific components, while a relatively generic model formulation was retained in this thesis for broader applicability. The use of a year of information to address system design is the most common practice [46], but design choices have long-term implications. These choices must lead to resilient systems to last through long-term changes and some trends argue in favor of the consideration of the systematic use of multiple years of data to address design choices [82]. Instead, we made the choice of using archetype data to enable a larger scope of exploration while retaining a generic framework.

The use of reduced complexity approaches, such as archetype data, linearization, simplified parameters, representative days (Chapter 2 and 3) or surrogate models (Chapter 5) necessarily result in a trade-off between errors and accuracy [46]. Suited for early-design stages, the proposed solutions present nonetheless the risk to alter the feasibility of systems in real contexts [71, 82]. While these considerations arise from a purely technical aspect, future efforts should also focus on the integration of more interdisciplinary considerations to encompass the impact of future climates, political, behavioral or broader financial implications

on design decisions.

An example of integration into reality is the hybrid usage of archetype, simulated data to explore possibilities during early design phases, and real data, better reflecting true usages. While this work solely relies on archetype building demands, applications to real-world cases may bring additional challenges [46].

Another example of integration into reality is the existence of two different model paradigms highlighted in Chapter 4 about thermal network modeling. One enables a realistic feature of temperature control, offering broader network flexibility. The other, based on simple energy transfer principles, allows to better represent the essential coupling of heating and cooling of 5G networks and address the broader design of the energy system surrounding the thermal network. While merging both seems essential for the future of modeling district thermal networks and more globally nested energy systems at community level, a deeper reflection on how these paradigms coexist and complete each other must be conducted.

All learning from Chapters 2, 3 and 4 lead to the novel elaboration of surrogate models. Due to its novelty, the proposed methodology remains closely tied to reduced scopes and strongly dependent on data created by computationally burdensome processes. The specific field would grow from additional usage over larger contexts, for which the exploration of transfer learning appears as an interesting potential direction. The extension of surrogate modeling to the operational aspect could further broaden the modeling capabilities and leverage the power of machine learning to the development of larger scale, component based models of energy systems. In this context, an important and unavoidable step is the deeper study of interactions between components in a nested energy model, as initiated through Chapter 4.

Appendices

Appendix A Ground storage (Chapter 2)

The interactions with the ground are designed with two ground-sourced heat pumps (GSHP), one for heating purposes and one for cooling purposes. The capacities of the ground to retain and deliver thermal energy is modeled using two coupled storage units (Figure Appendix A.1) The GSHP will consume electricity to generate heat on the heating stream, but it will also absorb heat from the hot segment of soil and generate cold to be stored in the other storage segment. An additional balancing unit allows the annihilation of one unit of active heat with one unit of active cold. This allows to actively store thermal energy. For example, an active unit of excess heat can be annihilated with one unit of cold generated by the GSCP, which will result in adding a unit of heat in the hot segment of the soil. The balancing unit is only available for system 3.

A decay induces a loss-of-charge proportional to the state-of-charge. The loss of of heat storage is considered a gain for the cold storage, as temperature naturally converge towards the "temperature at infinity". Equation (2.14) is modified to eq.(Appendix A.1) only for the soil model.

$$SOC^{A}_{\sigma,\tau_s,t+1} = (1-\kappa_{\tau_s})SOC^{A}_{\sigma,\tau_s,t} + \kappa_{\tau_s}\ SOC^{B}_{\sigma,\tau_s,t} + \eta^{ch}_{\tau_s}\dot{Q}^{ch}_{\sigma,\tau_s,t} - \frac{1}{\eta^{dch}_{\tau_s}}\dot{Q}^{dch}_{\sigma,\tau_s,t} \qquad \text{(Appendix A.1)}$$

Inspired by the the modeling from Fiorentini et al. [102], the capacity limit of the storage representing the ground are set to constrain the variations of ground temperature. Equation Appendix A.2 expresses the maximum capacity as function of the allowed temperature range ($T_{max} = 25^oC$, $T_{min} = 8^oC$, soil density $\rho = 1570kg/m^3$, soil thermal capacity c_p (input variable, c.f. Table 2.4) and maximum

depth of a borehole $L = 100m$.

$$P_{Soil}^{max} = \frac{(T_{max} - T_{min})}{\delta t} \frac{\pi}{4} \rho c_p L^3 \qquad \text{(Appendix A.2)}$$

From the same source, the decay κ is estimated in eq.(Appendix A.3), also involving the soil conductivity $\lambda = 2.4W/K.m$ and a constant $h = 21.2$ [102].

$$\kappa = \frac{2}{\pi} \frac{\lambda h}{\rho c_p L} \qquad \text{(Appendix A.3)}$$

Finally, a linear cost for the ground exchanger is estimated via eq.(Appendix A.4), involving the maximum number of boreholes n and drilling costs (input variable).

$$C_{Soil}^{lin} = C^{drill} \frac{L\ n}{P_{Soil}^{max}} \qquad \text{(Appendix A.4)}$$

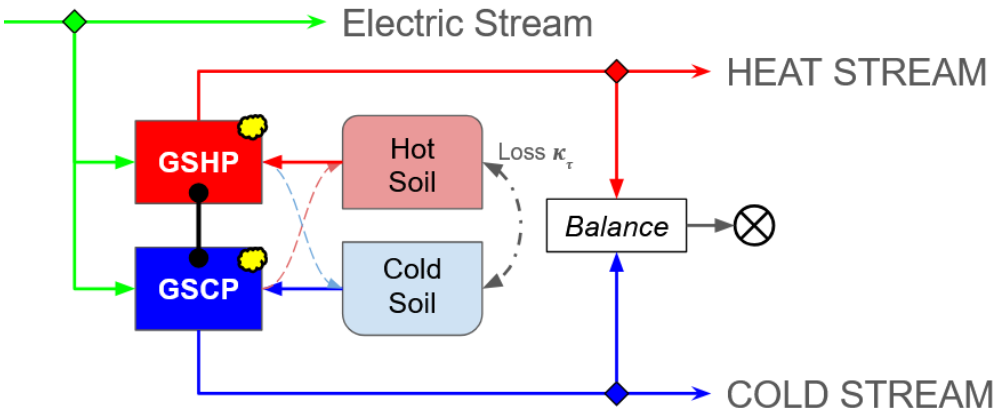


Figure Appendix A.1: Schematic model of the ground storage

Appendix B Parameters (Chapter 2 & 4)

The prices and parameters of the model are issued from commercial data, vendor websites, and related publications. Details about the price calculations can be found in the repository[1].

Costs, schedules ϵ and thermal capacities λ differ between 4G and 5G links due to the absence of insulation for the latter. Costs differ between Supply/Return and

[1] https://gitlab.com/fledee/detail_network

Table Appendix B.1: Techno-economic parameters for the network pipes

	C^{fix} ($/m)	C^{lin} ($/kW.m)	Δp_x (Pa/m)	D (mm)	ϵ (mm)	λ (W/m.K)	life (y)	L (m)
Supply 4G	387	0.03	200	250	44	0.04	30	200
Return 4G	102	0.03	200	250	44	0.04	30	200
Warm 5G	385	0.03	200	250	20	0.4	30	200
Cold 5G	100	0.03	200	250	20	0.4	30	200

Warm/Cold due to groundwork estimated as 285 $/m [56] and accounted for only once. Thus, we assume cold and warm 4G networks require separate groundwork.

Table Appendix B.2: Parameters of converters

	C^{fix} ($)	C^{lin} ($/kW)	$C^{O\&M}$ ($/kW)	$H_{\sigma,\tau}$	life (y)
PV	0	1020	9.7	pre	20
ST	-1366	1213	18.3	pre	25
GB	1492	47	15.6	0.925	25
EB	95	219	7.31	0.98	25
AChill	2460	664	11.6	0.72	20
CHP	3000	1150	42.5	0.37/0.46	20
ASHP	2968	198	32.5	pre	20
EChill	2968	198	32.5	pre	20
5G Interface	3876	115	83.2	pre	25
4G HX	0	100	1.1	1	20
Central GB	1268	40	13.3	0.98	25
Central EB	81	186	6.2	1	25
Central ASHP	2523	168	27.6	pre	20
Central EChill	2523	168	27.6	pre	20
GSHP / GSCP	2968	198	32.5	pre	30

Table Appendix B.3: Parameters for storage

	Solar Tank	HWT	WiP	PTES	Soil
C^{fix} ($)	0*	225	0**	554	1
C^{lin} ($/kWh)	0*	79	0**	0.4	C^{lin}_{Soil}
$C^{o\&m}$ ($/kWh)	-	-	-	-	-
β^{ch}/β^{dch}	0.3	0.5	0.2	**	-
κ	0.05	0.05	0.038	0.001	κ_{Soil}
η^{ch}/η^{dch}	1	1	1	1	1
life (y)	20	20	30	30	50

(*) Costs in ST, capacity is 5.7× ST
(**) Related to network links

The efficiencies "pre" are precomputed efficiencies for every time step t. The specific investment cost of ST results from a combination of factors: positive fixed costs, subsidies in most cases used to evaluate the cost, a minimum installation capacity to ensure a positive investment cost (see document available in the repository) A proportionality coefficient of 5.7 is enforced between the installed capacities of ST and Solar Tank. Detail the parameters. For the GSHP, several notes: the cost of "soil" correspond to the cost for a ground exchanger, based on drilling costs (Appendix A), while its decay is based on the soil properties. Both are variables of the problem, thus no value is specified. The costs of GSHP are estimated as equivalent to these of ASHP, since the ground exchanger is considered separately. Note that the investment costs only apply once for the GSHP, since the GSCP component is part of the reversible device.

The installable capacity of each device is limited. The limit is set as follows:

- **Electric generators**: maximum installable capacity is twice the hourly maximum electric power demand (hourly resolution) of the building over one year.
- **Heat generators**: maximum installable capacity is twice the hourly maximum heating power demand (hourly resolution) of the building over one year.
- **Cooling generators**: maximum installable capacity is twice the hourly maximum cooling power demand (hourly resolution) of the building over one year.
- **Electric storage**: maximum installable capacity is 12 times the worst hourly demand of the building over the year.
- **Heat storage**: maximum installable capacity is twice the worst hourly demand of the building over the year.
- **Soil**: The soil capacity is limited by P_{Soil}^{max} (eq.(Appendix A.2) in Appendix A), function of the soil capacity and a range of acceptable temperatures for the ground.

Appendix C 5G Interface (Chapter 4)

The unit is composed of three elements. A water sourced HP, a water sourced chiller and a balancing unit. The HP and chiller consume electricity to pump water

in one pipe, process it and reject the downgraded water in the other network pipe. The amount of thermal energy retrieved, fed back and transferred to the building equal the electricity input divided by the efficiency ($\dot{Q}_t^{el}/COP_t$). This principle represents the operation of 5G DTNs. For the building to also feed energy, the *balance* element is needed. It can cancel out 1 unit of active heat with one of active cold. Feeding an extra unit of heat to the network is equivalent to cooling the building down: (1) importing cold from the network, which generates the heat in the network, then (2) cancel out this imported cold with the excess heat.

This does not affect the heating and cooling demands, which must still be met. However, a simultaneous usage of the HP and chiller allows to cover simultaneous heating and cooling demands, exploiting the network locally with no generation of thermal energy, and at the only expense of the electricity bill to use both devices simultaneously.

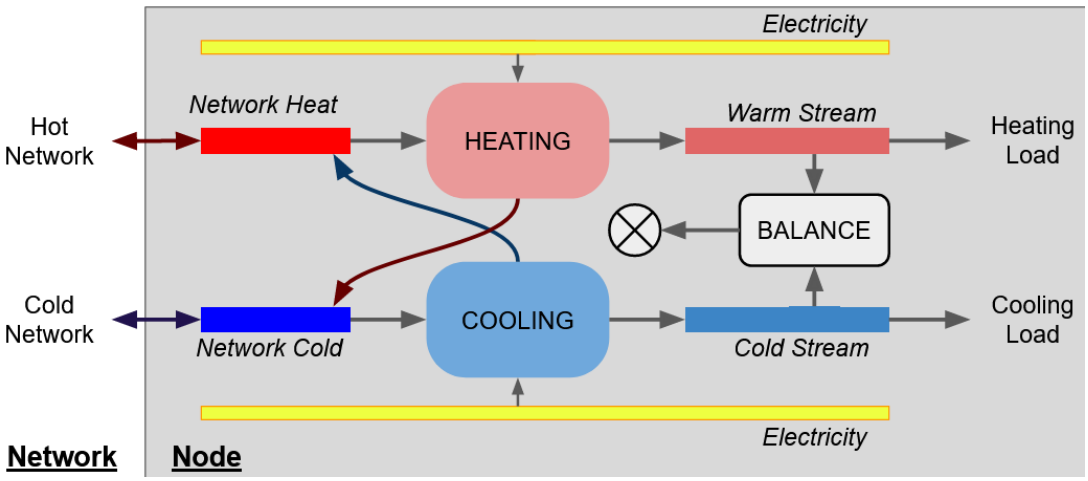


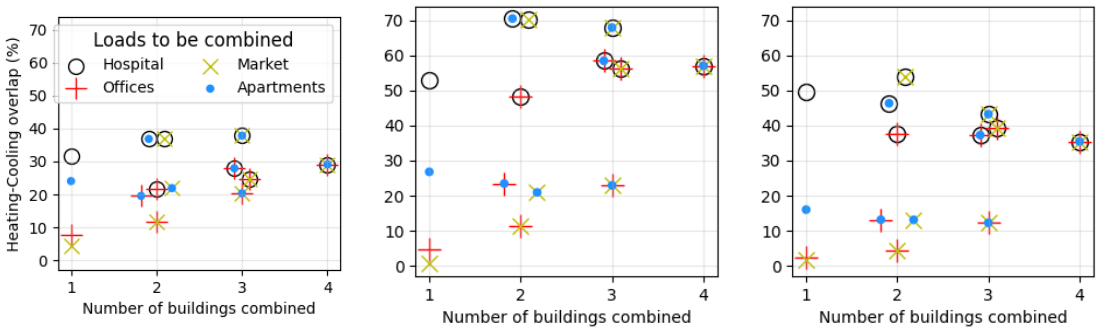
Figure Appendix C.1: Detailed principle of the interaction 5G DTN - Node

Appendix D Building data overlap (Chapter 4)

The total thermal load of every node type was described in Table 4.4 of section 4.4. The potential of 5G networks also relies on the simultaneous heating and cooling demands [56]. As a way to measure it, the Demand Overlap Coefficient (DOC) was suggested by Wirtz et al. [145]. The DOC Φ is expressed in eq.(Appendix D.5), where $\dot{Q}_{n,H,t}^{Load}$ and $\dot{Q}_{n,C,t}^{Load}$ are the heating and cooling loads at every time step t. Its value lies between 0% and 100%, with 100% meaning both loads are identical.

$$\Phi = \frac{2 \cdot \sum_{t \in \mathbb{T}} \min\{\dot{Q}_{n,H,t}^{Load}, \dot{Q}_{n,C,t}^{Load}\}}{\sum_{t \in \mathbb{T}} (\dot{Q}_{n,H,t}^{Load} + \dot{Q}_{n,C,t}^{Load}\})} \qquad \text{(Appendix D.5)}$$

The DOC Φ was calculated for every possible combination of node type considered in this study (Hospital, Apartments, SuperMarket, Offies). Figure Appendix D.1 displays the DOC values, detailed by number of combined nodes. Hospital and Apartments have the highest DOC for single nodes. Highest DOCs in all locations arise from combinations of two or three node types among Hospital, Apartments and SuperMarket. Figure Appendix D.1 also shows the DOC is overall lower in LA, at best 40%, and higher in OR, with values up to 70%.



(a) DOCs in LA (b) DOCs in OR (c) DOCs in WI

Figure Appendix D.1: Ratio of simultaneous heating and cooling for all possible combinations of loads

Appendix E Clustering methodology (Chapter 5)

This appendix explains the clustering algorithm used in Section 4.2 to group cases by similarity. The algorithm consists of three steps: data preprocessing, clustering, and reassignment of outlier samples.

The data preprocessing (Figure Appendix E.1) applies the preferred preprocessing steps of each case on the training target values. The base cases only go through standard scaling, while the refined cases may involve purposeful sampling, upsampling, a segmentation of the data and specific rescalings. The processed target values are then sorted by increasing order, and an interpolation using a random forest regressor may be used to reduce the number of samples down to 256 if upsampling was used.

The clustering procedure (Figure Appendix E.2), inspired by [197], applies a k-medoids clustering on the preprocessed data, where similarity is measured with the dynamic time warping (DTW) distance. A case i whose average distance $\bar{d}_{i,c}$ with the rest of its cluster c exceeds a threshold τ_1 are flagged as outliers. The

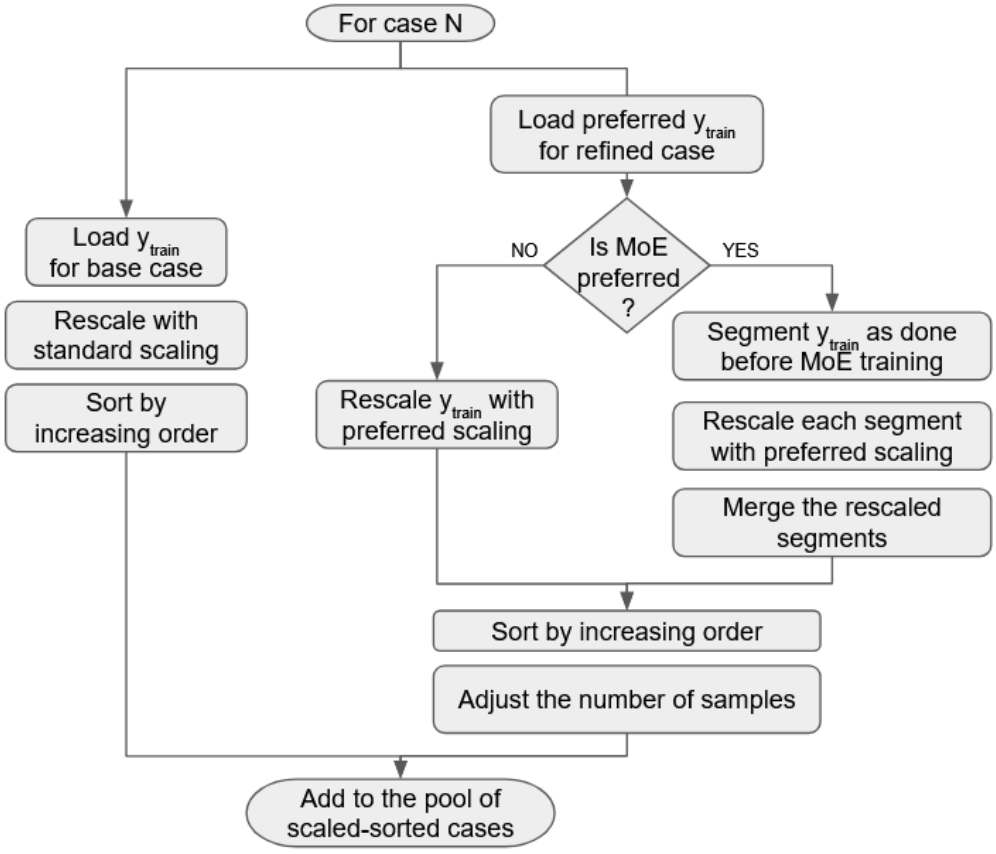


Figure Appendix E.1: Data preparation before clustering to organize cases by similarity

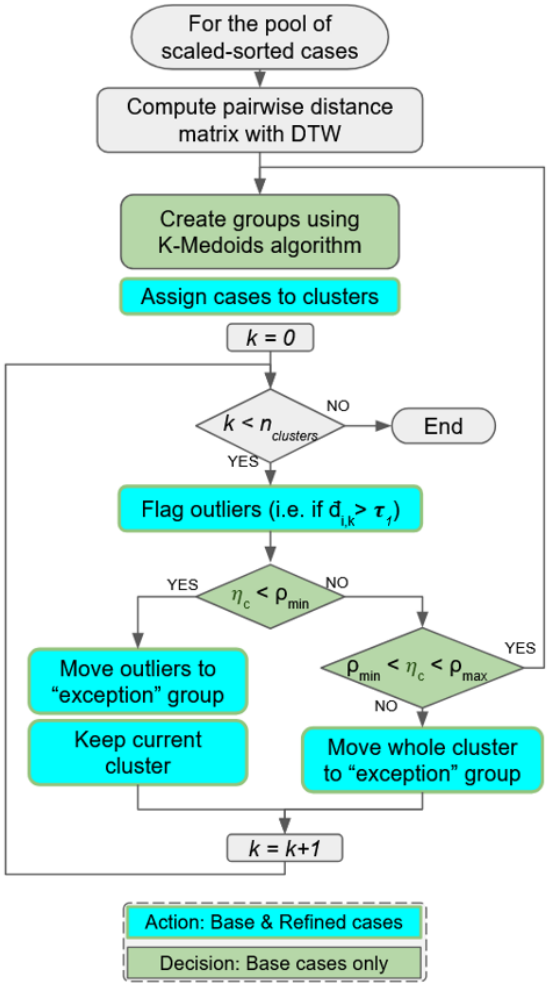


Figure Appendix E.2: Clustering process to group the cases by similarity

ratio of outliers is compared to critical values (ρ_{min} and ρ_{max}) to decide whether to reassign outliers or the whole cluster to an "exception" cluster, or to carry on a sub-clustering within the cluster. The formation of clusters and decisions with the ratio of outliers only involve base cases, then the refined cases are assigned to corresponding clusters.

Finally, samples in the exception cluster are reassessed following Algorithm Appendix E.1. This addresses the fact that the clustering step only divides the pool of cases. If the average distance of an outlier to any other cluster $đ_{i,c}$, or its distance to any other case in another cluster meets the distance threshold $d_{i,c}$, it is reassigned. This reassignment step is completed simultaneously for all cases of the exception cluster to prevent interference among reassigned cases.

Algorithm Appendix E.1: Clustering adjustment algorithm

for $i \in c_{exception}$ **do**
  Calculate $đ_{i,c} \forall c \in \mathcal{C}$;
  if $\exists c / đ_{i,c} < \tau_1$ **then**
    Assign i to cluster c with smallest $đ_{i,c}$
  else if $\exists c / (\tau_1 < đ_{i,c}) \& (\exists j \in c / d_{i,j} < \tau_1)$ **then**
    Assign i to cluster c complying with the condition with smallest $đ_{i,c}$
  else
    Leave sample i in cluster $c_{exception}$
  end
end

Appendix F Incompatible usage of decision trees (Chapter 5)

This appendix clarifies why decision trees or random forests are not suited to the methodology.

The approach uses Mixture of Experts, structures dividing the data into sub-regions to simplify the training and predictions processes. Although decision trees exhibit similar approaches, they partition the input space one variable at-a-time. This may be effective for problems with globally or locally dominant variables, i.e. explaining the variability of the design variable. In surrogate modeling however, and particularly for the design of MES, the design variables are expected to correlate with combinations of inputs, making single-variable splits insufficient (c.f. Figure 5.2 and Appendix F.1).

While random forests enhance decision trees through ensemble modeling, they simply aggregate models inherently inadequate to the current problem. They result in instability on the data segmentation, thus insufficient predictions on unseen data near a split.

Figure Appendix F.1 illustrates this argumentation with a simplified 2-variables MES design focusing on the installation of an ASHP. The graph axes represent linear ranges for electricity ($0.10-$0.35 per kWh) and gas ($0.05-$0.08 per kWh) prices. The 1st graph shows the sorted ASHP capacities for 64 samples. Graphs 2-8 show the segregation suggested by 7 different ML classifiers. Graph 9 shows the true segregation for 512 samples.

Figure Appendix F.1 mostly highlights the stairs-like boundary drawn by a decision tree, and the irregular boundary obtained from a random forest, while the expected boundary is a straight diagonal line.

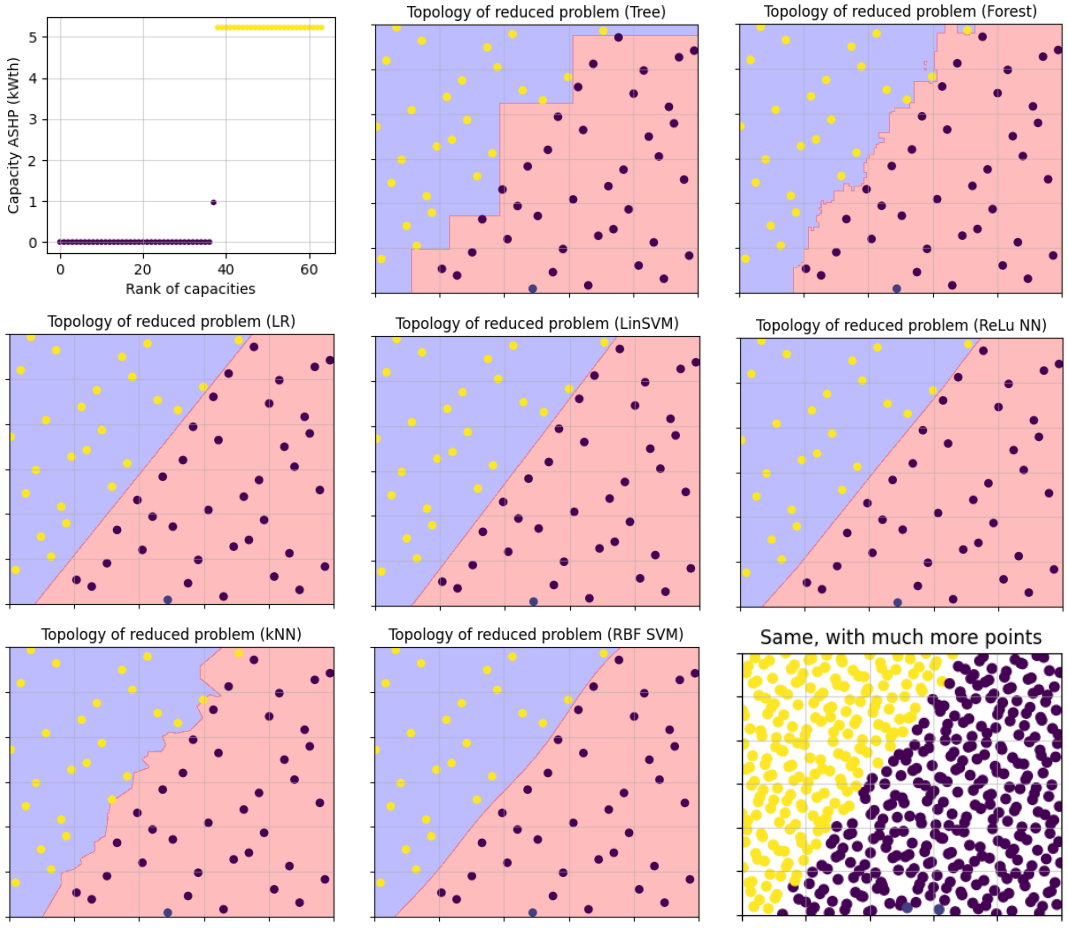


Figure Appendix F.1: Visual comparison of classification capabilities of different ML algorithms

Bibliography

[1] G. C. Aye, P. E. Edoja, Effect of economic growth on CO2 emission in developing countries: Evidence from a dynamic panel threshold model, Cogent Economics & Finance 5 (1) (2017) 1379239, publisher: Cogent OA _eprint: https://doi.org/10.1080/23322039.2017.1379239. doi:10.1080/23322039.2017.1379239.

[2] IEA, Global Energy Review 2020 - Analysis and key findings. A report by the International Energy Agency., Tech. rep., IEA, Paris, publication Title: IEA (Apr. 2020).

[3] K. Calvin, D. Dasgupta, G. Krinner, A. Mukherji, P. W. Thorne, C. Trisos, J. Romero, P. Aldunce, K. Barrett, G. Blanco, W. W. Cheung, S. Connors, F. Denton, A. Diongue-Niang, D. Dodman, M. Garschagen, O. Geden, B. Hayward, C. Jones, F. Jotzo, T. Krug, R. Lasco, Y.-Y. Lee, V. Masson-Delmotte, M. Meinshausen, K. Mintenbeck, A. Mokssit, F. E. Otto, M. Pathak, A. Pirani, E. Poloczanska, H.-O. Pörtner, A. Revi, D. C. Roberts, J. Roy, A. C. Ruane, J. Skea, P. R. Shukla, R. Slade, A. Slangen, Y. Sokona, A. A. Sörensson, M. Tignor, D. van Vuuren, Y.-M. Wei, H. Winkler, P. Zhai, Z. Zommers, J.-C. Hourcade, F. X. Johnson, S. Pachauri, N. P. Simpson, C. Singh, A. Thomas, E. Totin, P. Arias, M. Bustamante, I. Elgizouli, G. Flato, M. Howden, C. Méndez-Vallejo, J. J. Pereira, R. Pichs-Madruga, S. K. Rose, Y. Saheb, R. Sánchez Rodríguez, D. Ürge Vorsatz, C. Xiao, N. Yassaa, A. Alegría, K. Armour, B. Bednar-Friedl, K. Blok, G. Cissé, F. Dentener, S. Eriksen, E. Fischer, G. Garner, C. Guivarch, M. Haasnoot, G. Hansen, M. Hauser, E. Hawkins, T. Hermans, R. Kopp, N. Leprince-Ringuet, J. Lewis, D. Ley, C. Ludden, L. Niamir, Z. Nicholls, S. Some, S. Szopa, B. Trewin, K.-I. van der Wijst, G. Winter, M. Witting, A. Birt, M. Ha, J. Romero, J. Kim, E. F. Haites, Y. Jung, R. Stavins, A. Birt, M. Ha, D. J. A. Orendain, L. Ignon, S. Park, Y. Park, A. Reisinger, D. Cammaramo, A. Fischlin, J. S. Fuglestvedt, G. Hansen, C. Ludden, V. Masson-Delmotte, J. R. Matthews, K. Mintenbeck, A. Pirani, E. Poloczanska, N. Leprince-Ringuet, C. Péan, IPCC, 2023: Climate Change 2023: Syn-

thesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland., Tech. rep., Intergovernmental Panel on Climate Change (IPCC), edition: First (Jul. 2023). doi:10.59327/IPCC/AR6-9789291691647.

[4] B. Steffen, M. Beuse, P. Tautorat, T. S. Schmidt, Experience Curves for Operations and Maintenance Costs of Renewable Energy Technologies, Joule 4 (2) (2020) 359 – 375. doi:https://doi.org/10.1016/j.joule.2019.11.012.

[5] P. Jadun, C. McMillan, D. Steinberg, M. Muratori, L. Vimmerstedt, T. Mai, Electrification Futures Study: End-Use Electric Technology Cost and Performance Projections through 2050, Technical report 1416113, US Department of Energy - OSTI.GOV, United States (Dec. 2017).

[6] G. Mavromatidis, Model-based Design of Distributed Urban Energy Systems under Uncertainty, PhD Thesis, ETH Zurich, Zurich (2017).

[7] J. F. Marquant, R. Evins, L. A. Bollinger, J. Carmeliet, A holarchic approach for multi-scale distributed energy system optimisation, Applied Energy 208 (2017) 935–953. doi:10.1016/j.apenergy.2017.09.057.

[8] B. Morvaj, R. Evins, J. Carmeliet, Optimising urban energy systems: Simultaneous system sizing, operation and district heating network layout, Energy 116 (2016) 619 – 636. doi:https://doi.org/10.1016/j.energy.2016.09.139.

[9] B. Morvaj, R. Evins, J. Carmeliet, Decarbonizing the electricity grid: The impact on urban energy systems, distribution grids and district heating potential, Applied Energy 191 (2017) 125 – 140. doi:https://doi.org/10.1016/j.apenergy.2017.01.058.

[10] S. Moret, F. Babonneau, M. Bierlaire, F. Maréchal, Decision support for strategic energy planning: A robust optimization framework, European Journal of Operational Research 280 (2) (2020) 539–554. doi:10.1016/j.ejor.2019.06.015.

[11] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang, C. Josz, R. Korab, B. Lesieutre, J. Maeght, D. K. Molzahn, T. J. Overbye, P. Panciatici, B. Park, J. Snodgrass, R. Zimmerman, The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms, arXiv:1908.02788 [math] (Aug. 2019).

[12] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, S. McLaughlin, Multi-scale electrical load modelling for demand-side management, in: 2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), 2012, pp. 1–8, iSSN: 2165-4824. doi:10.1109/ISGTEurope.2012.6465866.

[13] F. Farzan, M. A. Jafari, J. Gong, F. Farzan, A. Stryker, A multi-scale adaptive model of residential energy demand, Applied Energy 150 (2015) 258–273. doi:10.1016/j.apenergy.2015.04.008.

[14] L. Kotzur, L. Nolting, M. Hoffmann, T. Groß, A. Smolenko, J. Priesmann, H. Büsing, R. Beer, F. Kullmann, B. Singh, A. Praktiknjo, D. Stolten, M. Robinius, A modeler's guide to handle complexity in energy systems optimization, Advances in Applied Energy 4 (2021) 100063. doi:10.1016/j.adapen.2021.100063.

[15] C. T. M. Clack, A. Choukulkar, B. Coté, S. A. McKee, Why Local Solar For All Costs Less: A new roadmap for the lowest cost grid., Technical report 1, Vibrant Clean Energy LLC, Boulder, Colorado (Dec. 2020).

[16] G. Mavromatidis, K. Orehounig, L. A. Bollinger, M. Hohmann, J. F. Marquant, S. Miglani, B. Morvaj, P. Murray, C. Waibel, D. Wang, J. Carmeliet, Ten questions concerning modeling of distributed multi-energy systems, Building and Environment 165 (2019) 106372. doi:10.1016/j.buildenv.2019.106372.

[17] M. McPherson, L. D. Harvey, B. Karney, System design and operation for integrating variable renewable energy resources through a comprehensive characterization framework, Renewable Energy 113 (2017) 1019–1032. doi:10.1016/j.renene.2017.06.071.

[18] A. Brown, A. Foley, D. Laverty, S. McLoone, P. Keatley, Heating and cooling networks: A comprehensive review of modelling approaches to map future directions, Energy 261 (2022) 125060. doi:10.1016/j.energy.2022.125060.

[19] M. Mohammadi, Y. Noorollahi, B. Mohammadi-Ivatloo, H. Yousefi, Energy hub: From a model to a concept – A review, Renewable and Sustainable Energy Reviews 80 (2017) 1512–1527. doi:10.1016/j.rser.2017.07.030.

[20] M. Geidl, G. Koeppel, P. Favre-Perrod, B. Klöckl, G. Andersson, K. Fröhlich, The energy hub-A powerful concept for future energy systems, in: Third Annual Carnegie Mellon Conference on the Electricity Industry, Carnegie Mellon, Pittsburg, 2007, pp. 13–14.

[21] P. Gabrielli, A. Poluzzi, G. J. Kramer, C. Spiers, M. Mazzotti, M. Gazzani, Seasonal energy storage for zero-emissions multi-energy systems via underground hydrogen storage, Renewable and Sustainable Energy Reviews 121 (2020) 109629. doi:10.1016/j.rser.2019.109629.

[22] G. Mavromatidis, K. Orehounig, J. Carmeliet, A review of uncertainty characterisation approaches for the optimal design of distributed energy systems, Renewable and Sustainable Energy Reviews 88 (2018) 258 – 277. doi:https://doi.org/10.1016/j.rser.2018.02.021.

[23] G. Mavromatidis, K. Orehounig, J. Carmeliet, Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems, Applied Energy 214 (2018) 219 – 238. doi:https://doi.org/10.1016/j.apenergy.2018.01.062.

[24] M. Faccio, M. Gamberi, M. Bortolini, M. Nedaei, State-of-art review of the optimization methods to design the configuration of hybrid renewable energy systems (HRESs), Frontiers in Energy 12 (4) (2018) 591–622. doi:10.1007/s11708-018-0567-x.

[25] M. Geidl, G. Andersson, Operational and structural optimization of multi-carrier energy systems, European Transactions on Electrical Power 16 (5) (2006) 463–477. doi:10.1002/etep.112.

[26] Y. Xia, Q. Xu, J. Fang, Y. Huang, L. Shi, F. Wu, Surrogate model enabled integrated energy system trading in buildings considering bidding characteristics, Energy and Buildings 306 (2024) 113939. doi:10.1016/j.enbuild.2024.113939.

[27] H. Sadeghi, M. Rashidinejad, M. Moeini-Aghtaie, A. Abdollahi, The energy hub: An extensive survey on the state-of-the-art, Applied Thermal Engineering 161 (2019) 114071. doi:https://doi.org/10.1016/j.applthermaleng.2019.114071.

[28] S. Pfenninger, Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability, Applied Energy 197 (2017) 1–13. doi:10.1016/j.apenergy.2017.03.051.

[29] R. Evins, K. Orehounig, V. Dorer, J. Carmeliet, New formulations of the 'energy hub' model to address operational constraints, Energy 73 (2014) 387–398. doi:10.1016/j.energy.2014.06.029.

[30] C. Shao, X. Wang, M. Shahidehpour, X. Wang, B. Wang, An MILP-Based Optimal Power Flow in Multicarrier Energy Systems, IEEE Transactions

on Sustainable Energy 8 (1) (2017) 239–248. doi:10.1109/TSTE.2016.2595486.

[31] K. Orehounig, R. Evins, V. Dorer, Integration of decentralized energy systems in neighbourhoods using the energy hub approach, Applied Energy 154 (2015) 277–289. doi:10.1016/j.apenergy.2015.04.114.

[32] M. Hoffmann, B. U. Schyska, J. Bartels, T. Pelser, J. Behrens, M. Wetzel, H. C. Gils, C.-F. Tang, M. Tillmanns, J. Stock, A. Xhonneux, L. Kotzur, A. Praktiknjo, T. Vogt, P. Jochem, J. Linßen, J. M. Weinand, D. Stolten, A review of mixed-integer linear formulations for framework-based energy system models, Advances in Applied Energy 16 (2024) 100190. doi:10.1016/j.adapen.2024.100190.

[33] T. Terlouw, P. Gabrielli, T. AlSkaif, C. Bauer, R. McKenna, M. Mazzotti, Optimal economic and environmental design of multi-energy systems, Applied Energy 347 (2023) 121374. doi:10.1016/j.apenergy.2023.121374.

[34] m. mohammadi, y. noorollahi, b. Mohammadi-ivatloo, h. yousefi, s. jalilinasrabady, Modeling the uncertainties in energy hubs for optimal scheduling - A review, Journal of Energy Management and Technology 1 (Online First) (2017) 1–17. doi:10.22109/jemt.2017.49432.

[35] A. J. Bogensperger, Y. Fabel, J. Ferstl, Accelerating Energy-Economic Simulation Models via Machine Learning-Based Emulation and Time Series Aggregation, Energies 15 (3) (2022) 1239, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en15031239.

[36] S. Risch, J. M. Weinand, K. Schulze, S. Vartak, M. Kleinebrahm, N. Pflugradt, F. Kullmann, L. Kotzur, R. McKenna, D. Stolten, Scaling energy system optimizations: Techno-economic assessment of energy autonomy in 11 000 German municipalities, Energy Conversion and Management 309 (2024) 118422. doi:10.1016/j.enconman.2024.118422.

[37] H. Kiani Rad, Z. Moravej, An approach for simultaneous distribution, subtransmission, and transmission networks expansion planning, International Journal of Electrical Power & Energy Systems 91 (2017) 166–182. doi:10.1016/j.ijepes.2017.03.010.

[38] M. Khodadadi Arpanahi, M.-E. Hamedani-Golshan, A competitive decentralized framework for Volt-VAr optimization of transmission and distribution systems with high penetration of distributed energy resources, Electric Power Systems Research 186 (2020) 106421. doi:10.1016/j.epsr.2020.106421.

[39] P. Du, Q. Zhao, H. Deng, Z. Ma, Z. Chen, M. Gong, Bi–level distributed day ahead economic dispatch model of transmission and distribution networks, in: 2020 International Conference on Electrical Engineering and Control Technologies (CEECT), 2020, pp. 1–9. doi:10.1109/CEECT50755.2020.9298604.

[40] M. Torabi, R. Evins, Towards net-zero carbon buildings: Investigating the impact of early-stage structure design on building embodied carbon, The International Journal of Life Cycle Assessment (Mar. 2024). doi:10.1007/s11367-024-02287-w.

[41] K. Lari, K. Cant, R. Evins, Developing tools for municipalities to meet carbon targets, Journal of Physics: Conference Series 2600 (8) (2023) 082014. doi:10.1088/1742-6596/2600/8/082014.

[42] S. Buffa, M. Cozzini, M. D'Antoni, M. Baratieri, R. Fedrizzi, 5th generation district heating and cooling systems: A review of existing cases in Europe, Renewable and Sustainable Energy Reviews 104 (2019) 504–522. doi:10.1016/j.rser.2018.12.059.

[43] H. Lund, P. A. Østergaard, T. B. Nielsen, S. Werner, J. E. Thorsen, O. Gudmundsson, A. Arabkoohsar, B. V. Mathiesen, Perspectives on fourth and fifth generation district heating, Energy 227 (2021) 120520. doi:10.1016/j.energy.2021.120520.

[44] E. Thrampoulidis, G. Mavromatidis, A. Lucchi, K. Orehounig, A machine learning-based surrogate model to approximate optimal building retrofit solutions, Applied Energy 281 (2021) 116024. doi:10.1016/j.apenergy.2020.116024.

[45] J. Marquant, Multiscale urban energy system optimization using spatio-temporal clustering, PhD Thesis, ETH Zurich, Zurich (2018).

[46] M. Hoffmann, L. Kotzur, D. Stolten, M. Robinius, A Review on Time Series Aggregation Methods for Energy System Models, Energies 13 (3) (2020) 1–61, publisher: MDPI, Open Access Journal.

[47] L. Kotzur, P. Markewitz, M. Robinius, D. Stolten, Impact of different time series aggregation methods on optimal energy system design, Renewable Energy 117 (2018) 474–487. doi:10.1016/j.renene.2017.10.017.

[48] T. Schütz, M. H. Schraven, M. Fuchs, P. Remmen, D. Müller, Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis, Renewable Energy 129 (2018) 570–582. doi:10.1016/j.renene.2018.06.028.

[49] K. Poncelet, E. Delarue, D. Six, J. Duerinck, W. D'haeseleer, Impact of the level of temporal and operational detail in energy-system planning models, Applied Energy 162 (2016) 631 – 643. doi:https://doi.org/10.1016/j.apenergy.2015.10.100.

[50] E. S. Pinto, L. M. Serra, A. Lázaro, Evaluation of methods to select representative days for the optimization of polygeneration systems, Renewable Energy 151 (2020) 488–502. doi:10.1016/j.renene.2019.11.048.

[51] H. Teichgraeber, C. P. Lindenmeyer, N. Baumgärtner, L. Kotzur, D. Stolten, M. Robinius, A. Bardow, A. R. Brandt, Extreme events in time series aggregation: A case study for optimal residential energy supply systems, Applied Energy 275 (2020) 115223. doi:10.1016/j.apenergy.2020.115223.

[52] D. Z. Fitiwi, F. de Cuadra, L. Olmos, M. Rivier, A new approach of clustering operational states for power network expansion planning problems dealing with RES (renewable energy source) generation operational variability and uncertainty, Energy 90 (2015) 1360–1376. doi:10.1016/j.energy.2015.06.078.

[53] B. Bahl, T. Söhler, M. Hennen, A. Bardow, Typical Periods for Two-Stage Synthesis by Time-Series Aggregation with Bounded Error in Objective Function, Frontiers in Energy Research 5, publisher: Frontiers (2018). doi:10.3389/fenrg.2017.00035.

[54] P. A. Østergaard, S. Werner, A. Dyrelund, H. Lund, A. Arabkoohsar, P. Sorknæs, O. Gudmundsson, J. E. Thorsen, B. V. Mathiesen, The four generations of district cooling - A categorization of the development in district cooling from origin to future prospect, Energy 253 (2022) 124098. doi:10.1016/j.energy.2022.124098.

[55] S. Fabozzi, G. De Luca, L. Vanoli, Chapter 9 - Fourth generation district heating and cooling, in: F. Calise, M. Dentice D'Accadia, L. Vanoli, M. Vicidomini (Eds.), Polygeneration Systems, Academic Press, 2022, pp. 323–350. doi:10.1016/B978-0-12-820625-6.00003-7.

[56] M. Wirtz, L. Kivilip, P. Remmen, D. Müller, 5th Generation District Heating: A novel design approach based on mathematical optimization, Applied Energy 260 (2020) 114158. doi:10.1016/j.apenergy.2019.114158.

[57] A. A. Eladl, M. I. El-Afifi, M. M. El-Saadawi, B. E. Sedhom, A review on energy hubs: Models, methods, classification, applications, and future trends, Alexandria Engineering Journal 68 (2023) 315–342. doi:10.1016/j.aej.2023.01.021.

[58] C. Vrionis, Generation Expansion Planning with high shares of Renewable Energy Sources: Single and Multi objective optimization based on Metamodel-assisted Evolutionary Algorithms, Doctoral thesis, National Technical University of Athens, Athens, accepted: 2020-07-28T08:14:17Z (Jul. 2020).

[59] C. S. Köhnen, J. Priesmann, L. Nolting, L. Kotzur, M. Robinius, A. Praktiknjo, The potential of deep learning to reduce complexity in energy system modeling, International Journal of Energy Research n/a (n/a), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/er.7448 (Nov. 2021). doi:10.1002/er.7448.

[60] E. Thrampoulidis, G. Hug, K. Orehounig, Approximating optimal building retrofit solutions for large-scale retrofit analysis, Applied Energy 333 (2023) 120566. doi:10.1016/j.apenergy.2022.120566.

[61] L. Kotzur, P. Markewitz, M. Robinius, D. Stolten, Time series aggregation for energy system design: Modeling seasonal storage, Applied Energy 213 (2018) 123–135. doi:10.1016/j.apenergy.2018.01.023.

[62] P. Gabrielli, M. Gazzani, E. Martelli, M. Mazzotti, Optimal design of multi-energy systems with seasonal storage, Applied Energy 219 (2018) 408–424. doi:10.1016/j.apenergy.2017.07.142.

[63] F. Domínguez-Muñoz, J. M. Cejudo-López, A. Carrillo-Andrés, M. Gallardo-Salazar, Selection of typical demand days for CHP optimization, Energy and Buildings 43 (11) (2011) 3036–3043. doi:10.1016/j.enbuild.2011.07.024.

[64] D. Bauer, R. Marx, J. Nußbicker-Lux, F. Ochs, W. Heidemann, H. Müller-Steinhagen, German central solar heating plants with seasonal heat storage, Solar Energy 84 (4) (2010) 612–623. doi:10.1016/j.solener.2009.05.013.

[65] R. Renaldi, D. Friedrich, Multiple time grids in operational optimisation of energy systems with short- and long-term thermal energy storage, Energy 133 (2017) 784–795. doi:10.1016/j.energy.2017.05.120.

[66] J. F. Marquant, A. O. Omu, K. Orehounig, R. Evins, J. Carmeliet, Application of spatial-temporal clustering to facilitate energy system modelling, in: BS2015, Hyderabad, India, 2015, pp. 550–558.

[67] P. Murray, J. Carmeliet, K. Orehounig, Multi-Objective Optimisation of Power-to-Mobility in Decentralised Multi-Energy Systems, Energy 205 (2020) 117792. doi:10.1016/j.energy.2020.117792.

[68] H. Teichgraeber, A. R. Brandt, Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison, Applied Energy 239 (2019) 1283–1293. doi:10.1016/j.apenergy.2019.02.012.

[69] M. Hoffmann, L. Kotzur, D. Stolten, The Pareto-optimal temporal aggregation of energy system models, Applied Energy 315 (2022) 119029. doi:10.1016/j.apenergy.2022.119029.

[70] K. Poncelet, H. Höschle, E. Delarue, A. Virag, W. D'haeseleer, Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems, IEEE Transactions on Power Systems 32 (3) (2017) 1936–1948, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2016.2596803.

[71] M. Fleschutz, M. Bohlayer, M. Braun, M. Murphy, The Hidden Cost of Using Time Series Aggregation for Modeling Low-Carbon Industrial Energy Systems: An Investors' Perspective (2024). doi:10.2139/ssrn.4796211. URL https://www.ssrn.com/abstract=4796211

[72] Z. Li, Y. Xia, Y. Bo, W. Wei, Optimal planning for electricity-hydrogen integrated energy system considering multiple timescale operations and representative time-period selection, Applied Energy 362 (2024) 122965. doi:10.1016/j.apenergy.2024.122965.

[73] A. Vecchi, D. Davis, M. Brear, L. Aye, Least-cost solutions to household energy supply decarbonisation in temperate and sub-tropical climates, Journal of Cleaner Production 448 (2024) 141465. doi:10.1016/j.jclepro.2024.141465.

[74] P. Thiran, H. Jeanmart, F. Contino, Validation of a Method to Select a Priori the Number of Typical Days for Energy System Optimisation Models, Energies 16 (6) (2023) 2772. doi:10.3390/en16062772.

[75] S. Pfenninger, Open code and data are not enough: understandability as design goal for energy system models, Progress in Energy 6 (3) (2024) 033002. doi:10.1088/2516-1083/ad371e.

[76] N. R. E. Laboratory, Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States [data set], type: dataset (2014). URL https://doi.org/10.25984/1788456

[77] E. Wilson, Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States (Nov. 2014). doi:10.25984/1788456. URL https://doi.org/10.25984/1788456

[78] N. Begum, L. Ulanova, J. Wang, E. Keogh, Accelerating Dynamic Time Warping Clustering with a Novel Admissible Pruning Strategy, in: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '15, Association for Computing Machinery, New York, NY, USA, 2015, pp. 49–58. doi:10.1145/2783258.2783286.

[79] B. van der Heijde, A. Vandermeulen, R. Salenbien, L. Helsen, Representative days selection for district energy system optimisation: a solar district heating system with seasonal storage, Applied Energy 248 (2019) 79–94. doi:10.1016/j.apenergy.2019.04.030.

[80] M. Wirtz, M. Hahn, T. Schreiber, D. Müller, Design optimization of multi-energy systems using mixed-integer linear programming: Which model complexity and level of detail is sufficient?, Energy Conversion and Management 240 (2021) 114249. doi:10.1016/j.enconman.2021.114249.

[81] S. Pfenninger, I. Staffell, Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data, Energy 114 (2016) 1251–1265. doi:10.1016/j.energy.2016.08.060.

[82] H. Teichgraeber, A. R. Brandt, Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities, Renewable and Sustainable Energy Reviews 157 (2022) 111984. doi:10.1016/j.rser.2021.111984.

[83] P. Westermann, T. Christiaanse, W. Beckett, P. Kovacs, R. Evins, besos: Building and Energy Simulation, Optimization and Surrogate Modelling, Journal of Open Source Software 6 (60) (2021) 2677. doi:10.21105/joss.02677.

[84] L. Gurobi Optimization, Gurobi Optimizer Reference Manual (2021). URL http://www.gurobi.com

[85] M. Aunedi, A. M. Pantaleo, K. Kuriyan, G. Strbac, N. Shah, Modelling of national and local interactions between heat and electricity networks in low-carbon energy systems, Applied Energy 276 (2020) 115522. doi:10.1016/j.apenergy.2020.115522.

[86] K. Gjoka, B. Rismanchi, R. H. Crawford, Fifth-generation district heating and cooling systems: A review of recent advancements and implementation barriers, Renewable and Sustainable Energy Reviews 171 (2023) 112997. doi:10.1016/j.rser.2022.112997.

[87] B. Nérot, N. Lamaison, M. Mabrouk, R. Bavière, B. Lacarrière, Optimization framework for evaluating urban thermal systems potential, Energy 270 (2023) 126851. doi:10.1016/j.energy.2023.126851.

[88] M. Bilardo, F. Sandrone, G. Zanzottera, E. Fabrizio, Modelling a fifth-generation bidirectional low temperature district heating and cooling (5GDHC) network for nearly Zero Energy District (nZED), Energy Reports 7 (2021) 8390–8405. doi:10.1016/j.egyr.2021.04.054.

[89] J. Lindhe, S. Javed, D. Johansson, H. Bagge, A review of the current status and development of 5GDHC and characterization of a novel shared energy system, Science and Technology for the Built Environment 28 (5) (2022) 595–609. doi:10.1080/23744731.2022.2057111.

[90] G. Brumana, G. Franchini, E. Ghirardi, S. Ravelli, Optimization of Solar District Heating & Cooling Systems, Journal of Physics: Conference Series 2385 (1) (2022) 012113. doi:10.1088/1742-6596/2385/1/012113.

[91] F. Calise, F. L. Cappiello, L. Cimmino, M. D. d'Accadia, M. Vicidomini, A comparative thermoeconomic analysis of fourth generation and fifth generation district heating and cooling networks, Energy (2023) 128561doi:10.1016/j.energy.2023.128561.

[92] M. Gross, B. Karbasi, T. Reiners, L. Altieri, H.-J. Wagner, V. Bertsch, Implementing prosumers into heating networks, Energy 230 (2021) 120844. doi:10.1016/j.energy.2021.120844.

[93] O. Gudmundsson, R.-R. Schmidt, A. Dyrelund, J. E. Thorsen, Economic comparison of 4GDH and 5GDH systems – Using a case study, Energy 238 (2022) 121613. doi:10.1016/j.energy.2021.121613.

[94] J. M. Jebamalai, K. Marlein, J. Laverge, Design and cost comparison of district heating and cooling (DHC) network configurations using ring topology – A case study, Energy 258 (2022) 124777. doi:10.1016/j.energy.2022.124777.

[95] Y. Zhang, P. Johansson, A. Sasic Kalagasidis, Assessment of district heating and cooling systems transition with respect to future changes in demand profiles and renewable energy supplies, Energy Conversion and Management 268 (2022) 116038. doi:10.1016/j.enconman.2022.116038.

[96] M.-A. Millar, Z. Yu, N. Burnside, G. Jones, B. Elrick, Identification of key performance indicators and complimentary load profiles for 5th generation district energy networks, Applied Energy 291 (2021) 116672. doi:10.1016/j.apenergy.2021.116672.

[97] C. Bordin, A. Gordini, D. Vigo, An optimization approach for district heating strategic network design, European Journal of Operational Research 252 (1) (2016) 296–307. doi:10.1016/j.ejor.2015.12.049.

[98] T. Résimont, Q. Louveaux, P. Dewallef, Optimization Tool for the Strategic Outline and Sizing of District Heating Networks Using a Geographic Information System, Energies 14 (17) (2021) 5575, number: 17 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en14175575.

[99] D. Chen, Z. Abbas, Y. Li, X. Hu, S. Zeng, Y. Liu, Optimal centralized integrated energy station site approach based on energy transmission loss analysis, International Journal of Energy Research 45 (1) (2021) 894–907. doi:10.1002/er.5980.

[100] K. Kuriyan, N. Shah, A combined spatial and technological model for the planning of district energy systems, International Journal of Sustainable Energy Planning and Management (2019) Vol. 21 (2019)Publisher: International Journal of Sustainable Energy Planning and Management. doi:10.5278/IJSEPM.2019.21.8.

[101] A. Lorestani, J. Chebeir, M. Narimani, J. S. Cotton, Multi-objective Optimization of Integrated Community Energy and Harvesting (ICE-Harvest) System Based on Marginal Emission Factor, in: 2021 IEEE International Smart Cities Conference (ISC2), 2021, pp. 1–7, iSSN: 2687-8860. doi:10.1109/ISC253183.2021.9562882.

[102] M. Fiorentini, P. Heer, L. Baldini, Design optimization of a district heating and cooling system with a borehole seasonal thermal energy storage, Energy 262 (2023) 125464. doi:10.1016/j.energy.2022.125464.

[103] I. Petkov, P. Gabrielli, Power-to-hydrogen as seasonal energy storage: an uncertainty analysis for optimal design of low-carbon multi-energy systems, Applied Energy 274 (2020) 115197. doi:10.1016/j.apenergy.2020.115197.

[104] I. Petkov, P. Gabrielli, M. Spokaite, The impact of urban district composition on storage technology reliance: trade-offs between thermal storage, batteries, and power-to-hydrogen, Energy 224 (2021) 120102. doi:10.1016/j.energy.2021.120102.

[105] P. Murray, K. Orehounig, D. Grosspietsch, J. Carmeliet, A comparison of storage systems in neighbourhood decentralized energy system applications from 2015 to 2050, Applied Energy 231 (2018) 1285–1306. doi:10.1016/j.apenergy.2018.08.106.

[106] H. Dorotić, T. Pukšec, N. Duić, Multi-objective optimization of district heating and cooling systems for a one-year time horizon, Energy 169 (2019) 319–328. doi:10.1016/j.energy.2018.11.149.

[107] J. Unternährer, S. Moret, S. Joost, F. Maréchal, Spatial clustering for district heating integration in urban energy systems: Application to geothermal energy, Applied Energy 190 (2017) 749–763. doi:10.1016/j.apenergy.2016.12.136.

[108] D. Wang, X. Li, J. Marquant, J. Carmeliet, K. Orehounig, Advancing the Thermal Network Representation for the Optimal Design of Distributed Multi-Energy Systems, Frontiers in Energy Research 9 (2021) 668124. doi:10.3389/fenrg.2021.668124.

[109] A. M. Pantaleo, S. Giarola, A. Bauen, N. Shah, Integration of biomass into urban energy systems for heat and power. Part II: Sensitivity assessment of main techno-economic factors, Energy Conversion and Management 83 (2014) 362–376. doi:10.1016/j.enconman.2014.03.051.

[110] A. M. Pantaleo, S. Giarola, A. Bauen, N. Shah, Integration of biomass into urban energy systems for heat and power. Part I: An MILP based spatial optimization methodology, Energy Conversion and Management 83 (2014) 347–361. doi:10.1016/j.enconman.2014.03.050.

[111] A. Prasanna, V. Dorer, N. Vetterli, Optimisation of a district energy system with a low temperature network, Energy 137 (2017) 632–648. doi:10.1016/j.energy.2017.03.137.

[112] M. Wirtz, L. Neumaier, P. Remmen, D. Müller, Temperature control in 5th generation district heating and cooling networks: An MILP-based operation optimization, Applied Energy 288 (2021) 116608. doi:10.1016/j.apenergy.2021.116608.

[113] P. Gabrielli, A. Acquilino, S. Siri, S. Bracco, G. Sansavini, M. Mazzotti, Optimization of low-carbon multi-energy systems with seasonal geothermal energy storage: The Anergy Grid of ETH Zurich, Energy Conversion and Management: X 8 (2020) 100052. doi:10.1016/j.ecmx.2020.100052.

[114] M. Taylor, S. Long, O. Marjanovic, A. Parisio, Model Predictive Control of Smart Districts With Fifth Generation Heating and Cooling Networks, IEEE Transactions on Energy Conversion 36 (4) (2021) 2659–2669, conference Name: IEEE Transactions on Energy Conversion. doi:10.1109/TEC.2021.3082405.

[115] A. Lyden, C. Brown, I. Kolo, G. Falcone, D. Friedrich, Seasonal thermal energy storage in smart energy systems: District-level applications and modelling approaches, Renewable and Sustainable Energy Reviews 167 (2022) 112760. doi:10.1016/j.rser.2022.112760.

[116] J. von Rhein, G. P. Henze, N. Long, Y. Fu, Development of a topology analysis tool for fifth-generation district heating and cooling networks, Energy Conversion and Management 196 (2019) 705–716. doi:10.1016/j.enconman.2019.05.066.

[117] M. Abugabbara, S. Javed, D. Johansson, A simulation model for the design and analysis of district systems with simultaneous heating and cooling demands, Energy 261 (2022) 125245. doi:10.1016/j.energy.2022.125245.

[118] G. Quirosa, M. Torres, R. Chacartegui, Analysis of the integration of photovoltaic excess into a 5th generation district heating and cooling system for network energy storage, Energy 239 (2022) 122202. doi:10.1016/j.energy.2021.122202.

[119] P. Saini, P. Huang, F. Fiedler, A. Volkova, X. Zhang, Techno-economic analysis of a 5th generation district heating system using thermo-hydraulic model: A multi-objective analysis for a case study in heating dominated climate, Energy and Buildings 296 (2023) 113347. doi:10.1016/j.enbuild.2023.113347.

[120] A. Maccarini, A. Sotnikov, T. Sommer, M. Wetter, M. Sulzer, A. Afshari, Influence of building heat distribution temperatures on the energy performance and sizing of 5th generation district heating and cooling networks, Energy 275 (2023) 127457. doi:10.1016/j.energy.2023.127457.

[121] T. Bu, R. Fan, B. Zheng, K. Sun, Y. Zhou, Design and operation investigation for the fifth-generation heating and cooling system based on load forecasting in business districts, Energy and Buildings 294 (2023) 113243. doi:10.1016/j.enbuild.2023.113243.

[122] F. Calise, F. L. Cappiello, M. Dentice d'Accadia, F. Petrakopoulou, M. Vicidomini, A solar-driven 5th generation district heating and cooling network with ground-source heat pumps: a thermo-economic analysis, Sustainable Cities and Society 76 (2022) 103438. doi:10.1016/j.scs.2021.103438.

[123] Y. J. Youn, Y. H. Im, Analysis of operating characteristics of interconnected operation of thermal grids with bidirectional heat trade, Applied Thermal Engineering 229 (2023) 120608. doi:10.1016/j.applthermaleng.2023.120608.

[124] M. Taylor, W. Gao, S. Masum, M. Qadrdan, Techno-economic assessment of Bi-directional Low Temperature Networks, Applied Energy 347 (2023) 121202. doi:10.1016/j.apenergy.2023.121202.

[125] M. Belliardi, P. Caputo, G. Ferla, N. Cereghetti, B. Antonioli Mantegazzini, An innovative application of 5GDHC: A techno-economic assessment of shallow geothermal systems potential in different European climates, Energy 280 (2023) 128104. doi:10.1016/j.energy.2023.128104.

[126] H. Hirsch, A. Nicolai, An efficient numerical solution method for detailed modelling of large 5th generation district heating and cooling networks, Energy 255 (2022) 124485. doi:10.1016/j.energy.2022.124485.

[127] A. Allen, G. Henze, K. Baker, G. Pavlak, M. Murphy, An optimization framework for the network design of advanced district thermal energy systems, Energy Conversion and Management 266 (2022) 115839. doi:10.1016/j.enconman.2022.115839.

[128] Q. Qin, L. Gosselin, Multiobjective optimization and analysis of low-temperature district heating systems coupled with distributed heat pumps, Applied Thermal Engineering 230 (2023) 120818. doi:10.1016/j.applthermaleng.2023.120818.

[129] I. Best, Economic comparison of low-temperature and ultra-low-temperature district heating for new building developments with low heat demand densities in Germany, International Journal of Sustainable Energy Planning and Management (2018) 45–60 PáginasArtwork Size: 45-60 Páginas Publisher: International Journal of Sustainable Energy Planning and Management. doi:10.5278/IJSEPM.2018.16.4.

[130] I. Best, J. Orozaliev, K. Vajen, Impact of Different Design Guidelines on the Total Distribution Costs of 4th Generation District Heating Networks, Energy Procedia 149 (2018) 151–160. doi:10.1016/j.egypro.2018.08.179.

[131] P. Sorknæs, P. A. Østergaard, J. Z. Thellufsen, H. Lund, S. Nielsen, S. Djørup, K. Sperling, The benefits of 4th generation district heating in a 100% renewable energy system, Energy 213 (2020) 119030. doi:10.1016/j.energy.2020.119030.

[132] O. Arnfalk, Modeling energy losses and gains in low temperature bi-directional heating and cooling grids, Thesis M.Sc and Engineering, Lund University, Lund (Jun. 2022).

[133] Y. Song, Y. Yao, W. Na, Impacts of Soil and Pipe Thermal Conductivity on Performance of Horizontal Pipe in a Ground-source Heat Pump, in: Renewable Energy Resources and a Greener Future, Vol. Vol. VIII-11-1, Texas A&M University, Shenzhen, China, 2006, pp. 1–6.

[134] A. Hayrullin, A. Haibullina, A. Sinyavin, Insulation thermal conductivity heating networks during transportation thermal energy under dry and moisturizing condition: a comparative study of the guarded hot plate and guarded hot pipe method, Transportation Research Procedia 63 (2022) 1074–1080. doi:10.1016/j.trpro.2022.06.109.

[135] M. McPherson, B. Karney, A scenario based approach to designing electricity grids with high variable renewable energy penetrations in Ontario, Canada: Development and application of the SILVER model, Energy 138 (2017) 185–196. doi:10.1016/j.energy.2017.07.027.

[136] J. Amirmadhi, M. Abeysekera, J. Wu, Modelling of Electrical-Thermal-Hydraulic System Interdependencies in 5th Generation District Heating and Cooling Networks, preprint, Volume 23: Sustainable Energy Solutions for a Post-COVID Recovery towards a Better Future: Part VI (Feb. 2022). doi:10.46855/energy-proceedings-9460.

[137] M. Abugabbara, J. Lindhe, A Novel Method for Designing Fifth-Generation District Heating and Cooling Systems, E3S Web of Conferences 246 (2021) 09001. doi:10.1051/e3sconf/202124609001.

[138] T. Yang, W. Liu, G. J. Kramer, Q. Sun, Seasonal thermal energy storage: A techno-economic literature review, Renewable and Sustainable Energy Reviews 139 (2021) 110732. doi:10.1016/j.rser.2021.110732.

[139] M. Wirtz, Storages in 5GDHC networks: Benefits, planning, calculation - nPro (2024).
URL https://www.npro.energy/main/en/5gdhc-networks/storages-5gdhc-networks

[140] M. Wirtz, nPro: A web-based planning tool for designing district energy systems and thermal networks, Energy 268 (2023) 126575. doi:10.1016/j.energy.2022.126575.

[141] C. Rockenbaugh, J. Dean, D. Lovullo, L. Lisell, G. Barker, E. Hanckock, P. Norton, High Performance Flat Plate Solar Thermal Collector Evaluation, Tech. Rep. NREL/TP-7A40-66215, National Renewable Energy Lab. (NREL), Golden, CO (United States) (Sep. 2016). doi:10.2172/1326887.

[142] D. Hillel, Introduction to soil physics., Academic Press, New York, NY, USA, 1982.

[143] O. Ruhnau, L. Hirth, A. Praktiknjo, Time series of heat demand and heat pump efficiency for energy system modeling, Scientific Data 6 (1) (2019) 189, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s41597-019-0199-y.

[144] S. Buffa, A. Soppelsa, M. Pipiciello, G. Henze, R. Fedrizzi, Fifth-Generation District Heating and Cooling Substations: Demand Response with Artificial Neural Network-Based Model Predictive Control, Energies 13 (17) (2020). doi:10.3390/en13174339.

[145] M. Wirtz, L. Kivilip, P. Remmen, D. Müller, Quantifying Demand Balancing in Bidirectional Low Temperature Networks, Energy and Buildings 224 (2020) 110245. doi:10.1016/j.enbuild.2020.110245.

[146] M. Wirtz, M. Heleno, H. Romberg, T. Schreiber, D. Müller, Multi-period design optimization for a 5th generation district heating and cooling network, Energy and Buildings 284 (2023) 112858. doi:10.1016/j.enbuild.2023.112858.

[147] R. Zarin Pass, M. Wetter, M. Piette, A thermodynamic analysis of a novel bidirectional district heating and cooling network, Energy 144 (2018) 20–30. doi:10.1016/j.energy.2017.11.122.

[148] Y. Zhang, P. Johansson, A. S. Kalagasidis, Quantification of overlapping heating and cooling demand for the feasibility assessment of bi-directional systems over Europe, Energy and Buildings 294 (2023) 113244. doi:10.1016/j.enbuild.2023.113244.

[149] M. Miri, M. Saffari, R. Arjmand, M. McPherson, Integrated models in action: Analyzing flexibility in the Canadian power system toward a zero-emission future, Energy 261 (2022) 125181. doi:10.1016/j.energy.2022.125181.

[150] A. T. D. Perera, P. U. Wickramasinghe, V. M. Nik, J.-L. Scartezzini, Introducing reinforcement learning to the energy system design process, Applied Energy 262 (2020) 114580. doi:10.1016/j.apenergy.2020.114580.

[151] Q. Cai, X. Luo, P. Wang, C. Gao, P. Zhao, Hybrid model-driven and data-driven control method based on machine learning algorithm in energy hub and application, Applied Energy 305 (2022) 117913. doi:10.1016/j.apenergy.2021.117913.

[152] P. Westermann, R. Evins, Surrogate modelling for sustainable building design – A review, Energy and Buildings 198 (2019) 170–186. doi:10.1016/j.enbuild.2019.05.057.

[153] P. Westermann, M. Welzel, R. Evins, Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones, Applied Energy 278 (2020) 115563. doi:10.1016/j.apenergy.2020.115563.

[154] P. Westermann, R. Evins, Using Bayesian deep learning approaches for uncertainty-aware building energy surrogate models, Energy and AI 3 (2021) 100039. doi:10.1016/j.egyai.2020.100039.

[155] D. Rulff, E. Ralph, C. Kevin, Development of a surrogate model for interactive early-stage net-zero building design, in: Proceedings of the 18th IBPSA Conference, Vol. 18 of Building Simulation, IBPSA, Shanghai, 2023, pp. 980–987. doi:10.26868/25222708.2023.1394.

[156] B. Birdsell, R. Evins, Surrogate Modeling performance for building design problems, in: Proceedings of the 18th IBPSA Conference, Vol. 18 of Building Simulation, IBPSA, Shanghai, 2023, pp. 2612–2619. doi:10.26868/25222708.2023.1376.

[157] R. Evins, Multi-level optimization of building design, energy system sizing and operation, Energy 90 (2015) 1775 – 1789. doi:https://doi.org/10.1016/j.energy.2015.07.007.

[158] B. Liu, N. Sun, Q. Zhang, V. Grout, G. Gielen, A surrogate model assisted evolutionary algorithm for computationally expensive design optimization problems with discrete variables, in: 2016 IEEE Congress on Evolutionary Computation (CEC), 2016, pp. 1650–1657. doi:10.1109/CEC.2016.7743986.

[159] G. Zeng, M. Liu, Z. Lei, S. Zhang, Z. Chen, Optimal configuration of hydrogen energy storage in an integrated energy system considering variable hydrogen production efficiency, Journal of Energy Storage 98 (2024) 113044. doi:10.1016/j.est.2024.113044.

[160] M. G. Prina, M. Dallapiccola, D. Moser, W. Sparber, Machine learning as a surrogate model for EnergyPLAN: Speeding up energy system optimization at the country level, Energy 307 (2024) 132735. doi:10.1016/j.energy.2024.132735.

[161] J. Jiang, H. Yu, G. Song, J. Zhao, K. Zhao, H. Ji, P. Li, Surrogate model assisted multi-criteria operation evaluation of community integrated energy systems, Sustainable Energy Technologies and Assessments 53 (2022) 102656. doi:10.1016/j.seta.2022.102656.

[162] P. Aghaei Pour, T. Rodemann, J. Hakanen, K. Miettinen, Surrogate assisted interactive multiobjective optimization in energy system design of buildings, Optimization and Engineering 23 (1) (2022) 303–327. doi:10.1007/s11081-020-09587-8.

[163] Y. Elomari, C. Mateu, M. Marín-Genescà, D. Boer, A data-driven framework for designing a renewable energy community based on the integration of machine learning model with life cycle assessment and life cycle cost parameters, Applied Energy 358 (2024) 122619. doi:10.1016/j.apenergy.2024.122619.

[164] A. Sciazko, Surrogate modeling techniques applied to energy systems, Master's thesis, The School for Renewable Energy Science, Akureyri (Mar. 2011).

[165] C.-H. Park, C. S. Park, Limitations and issues of conventional artificial neural network-based surrogate models for building energy retrofit, Journal of Building Performance Simulation 17 (3) (2024) 361–370, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/19401493.2023.2282078. doi:10.1080/19401493.2023.2282078.

[166] A. T. D. Perera, P. U. Wickramasinghe, V. M. Nik, J.-L. Scartezzini, Machine learning methods to assist energy system optimization, Applied Energy 243 (2019) 191 – 205. doi:https://doi.org/10.1016/j.apenergy.2019.03.202.

[167] Z. Jahangiri, M. Judson, K. M. Yi, M. McPherson, A Deep Learning Approach for Exploring the Design Space for the Decarbonization of the Canadian Electricity System, Energies 16 (3) (2023) 1352, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en16031352.

[168] F. Lédée, R. Evins, A comparison of 4th and 5th generation thermal networks with energy hub, Energy 311 (2024) 133336. doi:10.1016/j.energy.2024.133336.

[169] EIA, Electric Power Monthly - U.S. Energy Information Administration (EIA) (2024).
URL https://www.eia.gov/electricity/monthly/epm_table_grapher.php

[170] EIA, Average Residential Price of Natural Gas (2023).
URL https://www.eia.gov/dnav/ng/ng_pri_sum_a_epg0_prs_dmcf_a.htm

[171] Statista, U.S. power sector carbon index by state 2023 (2023).
URL https://www.statista.com/statistics/1133295/electric-sector-carbon-dioxide-emission-rate-by-state-united-states/

[172] A. Venkatesh, P. Jaramillo, W. M. Griffin, H. S. Matthews, Uncertainty in Life Cycle Greenhouse Gas Emissions from United States Natural Gas End-Uses and its Effects on Policy, Environmental Science & Technology 45 (19) (2011) 8182–8189. doi:10.1021/es200930h.

[173] A. Burnham, J. Han, C. E. Clark, M. Wang, J. B. Dunn, I. Palou-Rivera, Life-Cycle Greenhouse Gas Emissions of Shale Gas, Natural Gas, Coal, and Petroleum, Environmental Science & Technology 46 (2) (2012) 619–627. doi:10.1021/es201942m.

[174] A. B. Owen, A randomized Halton algorithm in R, number: arXiv:1706.02808 arXiv:1706.02808 [cs, math, stat] (Jun. 2017). doi:10.48550/arXiv.1706.02808.
URL http://arxiv.org/abs/1706.02808

[175] P. Bratley, B. L. Fox, Algorithm 659: Implementing Sobol's quasirandom sequence generator, ACM Trans. Math. Softw. 14 (1) (1988) 88–100. doi:10.1145/42288.214372.

[176] S. Ahmad, G. Tesauro, Scaling and Generalization in Neural Networks: A Case Study, in: Advances in Neural Information Processing Systems, Vol. 1, Morgan-Kaufmann, Denver, CO, 1988, pp. 160–168.

[177] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, K. Hoagwood, Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research, Administration and Policy in Mental Health and Mental Health Services Research 42 (5) (2015) 533–544. doi:10.1007/s10488-013-0528-y.

[178] W. Wu, N. Kunz, P. Branco, ImbalancedLearningRegression - A Python Package to Tackle the Imbalanced Regression Problem, in: M.-R. Amini, S. Canu, A. Fischer, T. Guns, P. Kralj Novak, G. Tsoumakas (Eds.), Machine Learning and Knowledge Discovery in Databases, Springer Nature Switzerland, Cham, 2023, pp. 645–648. doi:10.1007/978-3-031-26422-1_48.

[179] G. Saraswat, B. Lundstrom, M. V. Salapaka, Scalable Hybrid Classification-Regression Solution for High-Frequency Nonintrusive Load Monitoring, in:

2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), IEEE, Washington, DC, USA, 2023, pp. 1–5. doi:10.1109/ISGT51731.2023.10066447.

[180] N. Faris Ali, M. Atef, An efficient hybrid LSTM-ANN joint classification-regression model for PPG based blood pressure monitoring, Biomedical Signal Processing and Control 84 (2023) 104782. doi:10.1016/j.bspc.2023.104782.

[181] M. Jordan, R. Jacobs, Hierarchical mixtures of experts and the EM algorithm, in: Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), Vol. 2, 1993, pp. 1339–1344 vol.2. doi:10.1109/IJCNN.1993.716791.

[182] M. Koosha, G. Khodabandelou, M. M. Ebadzadeh, A hierarchical estimation of multi-modal distribution programming for regression problems, Knowledge-Based Systems 260 (2023) 110129. doi:10.1016/j.knosys.2022.110129.

[183] F. Giacometto Torres, Adaptive load consumption modelling on the user side: contributions to load forecasting modelling based on supervised mixture of experts and genetic programming, Doctoral thesis, Universitat Politècnica de Catalunya, Departament d'Enginyeria Electrònica (Jul. 2017).

[184] M. Mohammadzadeh, E. Akbari, A. A. Salameh, M. Ghadamyari, S. Pirouzi, T. Senjyu, Application of Mixture of Experts in Machine Learning-Based Controlling of DC-DC Power Electronics Converter, IEEE Access 10 (2022) 117157–117169. doi:10.1109/ACCESS.2022.3218667.

[185] H. Nakada, K. Takaba, T. Katayama, Identification of piecewise affine systems based on statistical clustering technique, Automatica 41 (5) (2005) 905–913. doi:10.1016/j.automatica.2004.12.005.

[186] M. Nejati, N. Amjady, A New Solar Power Prediction Method Based on Feature Clustering and Hybrid-Classification-Regression Forecasting, IEEE Transactions on Sustainable Energy 13 (2) (2022) 1188–1198. doi:10.1109/TSTE.2021.3138592.

[187] P. Petersen, T. Rudolf, E. Sax, A Data-driven Energy Estimation based on the Mixture of Experts Method for Battery Electric Vehicles:, in: Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems, SCITEPRESS - Science and Technology Publications, Online Streaming, — Select a Country —, 2022, pp. 384–390. doi:10.5220/0011081000003191.

[188] M. P. Perrone, L. N. Cooper, When networks disagree: Ensemble methods for hybrid neural networks, in: How We Learn; How We Remember: Toward an Understanding of Brain and Neural Systems, Vol. Volume 10 of World Scientific Series in 20th Century Physics, WORLD SCIENTIFIC, 1995, pp. 342–358. doi:10.1142/9789812795885_0025.

[189] S. Massoudi, C. Picard, J. Schiffmann, An Integrated Approach to Designing Robust Gas-Bearing Supported Turbocompressors Through Surrogate Modeling and Constrained All-At-Once Multi-Objective Optimization, Journal of Mechanical Design 146 (121706) (Jul. 2024). doi:10.1115/1.4065823.

[190] Y. LeCun, L. Bottou, G. B. Orr, K. R. Müller, Efficient BackProp, in: G. B. Orr, K.-R. Müller (Eds.), Neural Networks: Tricks of the Trade, Springer, Berlin, Heidelberg, 1998, pp. 9–50. doi:10.1007/3-540-49430-8_2.

[191] S. Sen, K. P. Singh, P. Chakraborty, Dealing with imbalanced regression problem for large dataset using scalable Artificial Neural Network, New Astronomy 99 (2023) 101959. doi:10.1016/j.newast.2022.101959.

[192] J. M. del Castillo, F. G. Benítez, Improving trip forecasting models by means of the Box–Cox transformation, Transportmetrica A: Transport Science 9 (7) (2013) 653–674, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/18128602.2011.643931. doi:10.1080/18128602.2011.643931.

[193] R. Krzysztofowicz, Transformation and normalization of variates with specified distributions, Journal of Hydrology 197 (1) (1997) 286–292. doi:10.1016/S0022-1694(96)03276-3.

[194] J. Snoek, H. Larochelle, R. P. Adams, Practical Bayesian Optimization of Machine Learning Algorithms, in: Advances in Neural Information Processing Systems, Vol. 25, Curran Associates, Inc., 2012, pp. 1–9.

[195] IEA, World Energy Outlook 2024, Tech. rep., IEA, Paris (Oct. 2024).

[196] T. Christiaanse, R. Loonen, R. Evins, Techno-economic optimization for grid-friendly rooftop PV systems – A case study of commercial buildings in British Columbia, Sustainable Energy Technologies and Assessments 47 (2021) 101320. doi:10.1016/j.seta.2021.101320.

[197] D. Tan, Time Series Clustering — Deriving Trends and Archetypes from Sequential Data (Aug. 2021).
URL https://towardsdatascience.com/time-series-clustering-deriving-trends-