Impact of renewable energy hubs configurations on the national infrastructure

Application of levelized infrastructure-connected in energy systems modelling

Master Thesis

Author:
Arthur CHUAT

Electrical and Electronic Section

Supervisors:
Prof. François MARÉCHAL
Jonas SCHNIDRIG
Cédric TERRIER

Industrial Processes and Energy Systems Engineering
Ecole Polytechnique Fédérale de Lausanne

Sion, 24th March 2023

Contents

Abstract

This research presents a framework to identify and integrate typical districts in a national energy system to assess the impact of decentralized energy hub configurations. The framework includes a two-step global sensitivity analysis to determine the model's most influential parameters and explore the districts' solution space. The application of the framework on 10 representative districts of Switzerland reveals a variety of configurations for the built environment. Integrating these solutions into the national model indicates a 15% increase in annual costs due to a higher space heating demand of the built environment. The decentralized approach focuses on the deployment of PV within urban areas with a third of the total capacity of the centralized approach, while the distribution of decentralized technologies is the same for both models. The impact of feed-in tariff variation allows for identifying a levelized cost of electricity within each district. Districts in urban areas have a lower levelized cost of electricity than those in rural areas due to higher PV penetration. Overall, this framework provides insights into the potential to deploy renewable energy sources at the district level and the required investments to reach energy independence and carbon neutrality in the country.

Acknowledgments

This Master Thesis concludes my Master in Energy Science and Technology accomplished in the Electrical and Electronic section of EPFL. This master's degree has been wonderful journey and I am extremely grateful to have been able to conclude this chapter within the IPESE group.

I would like to thank Professor François Maréchal for giving me the opportunity to work under his supervision and to cultivate my passion and understanding of the energy field.

I would like to thank the many people who are part of the lab, who even while chasing deadlines know how to keep good spirit and have a laugh.

Finally, I would like to thank my two supervisors Jonas and Cédric for your unconditional support and patience. It has been a true honor to work by your side and a fulfilling experience.

Arthur

List of Figures

List of Tables

Acronyms

AMPL a mathematical programming language

CAPEX capital expenditure

DBSCAN density-based spatial clustering of applications with noise

DES distributed energy system

EE elementary effect

EHV extra high voltage

ERA energy reference area

ES EnergyScope

EUD end use demand

GHG greenhouse gases

GSA global sensitivity analysis

HP heat pump

HV high voltage

HW hot water

IPCC Intergovernmental Panel on Climate Change

KPI key performance indicators

LCOE levelized cost of electricity

LSA local sensitivity analysis

LV low voltage

MILP mixed integer linear programming

MV medium voltage

NG natural gas

OPEX operational expenditure

PV photovoltaic

RE renewable energy

REHO Renewable Energy Hub Optimizer

SA sensitivity analysis

SFOE Swiss Federal Office of Energy

SH space heating

SIA Schweizerische Ingenieur- und Architektenverein

TD typical day

TOTEX total expenditure

Chapter 1

Introduction

In March 2023, the Intergovernmental Panel on Climate Change (IPCC) published the synthesis report of their Sixth Assessment Report (AR6) with the following press release head line: *Urgent climate action can secure a livable future for all.* One single sentence grasps the amplitude of the situation. Throughout their report the IPCC provides a panel of possible solutions to reduce greenhouse gases (GHG) emissions which are the primary causes of global warming. Among those solutions, the energy supply strategy has a significant emissions reduction potential and illustrates itself as one of the most feasible mitigation options. Indeed, accounting for 73% of global emissions in 2020 [1], the energy sector has great potential to reduce its impact through the development of renewable energy (RE) and energy efficiency improvement. Within the sector, the built environment represented more than a third of the final energy demand in 2018 and represents an excellent opportunity for the deployment of RE [2].

Incorporating renewable energy sources into the building stock is a change of paradigm for the current energy system, which is predominantly operated from a centralized perspective. This centralized approach relies on a limited number of large energy conversion units supplying the majority of the energy demand. The widespread adoption of renewable energy sources in buildings significantly disrupts this system by enabling a decentralized energy supply. Therefore, a coordination between distributed investments and centralized actors is needed to promote a feasible energy transition.

Due to the complexity of this coordination, the development of large-scale models to investigate the impact and feasibility of such transformation are a key-point to enable a successful transition. They can support decision makers by informing on the impact of new technologies or policies on the current system. Moreover, this approach promotes fair solutions since the interests of local stakeholders are included in the design of the whole system.

In conclusion, a transition of the energy sector in place is crucial to mitigate the effects of global warming. A decentralized energy supply strategy based on the penetration of renewable energy within the built environment is proposed as a major solution. Such transformation of the system

must be assessed by models to provide optimal pathways towards carbon neutrality.

Chapter 2

State of the art

This chapter assesses the current state-of-the-art regarding energy system modelling (section 2.1). Secondly, the application of global sensitivity analysis (GSA) to energy system model is presented in section 2.2. Additionally, diverse clustering techniques are addressed in section 2.3. And finally, the gaps and contributions of the project is presented (section 2.4).

2.1 Energy system modelling

Energy models are mathematical formulations of energy systems. They help characterising the behaviour of a given energy system under different inputs and constraints. And play a crucial role in guiding the ongoing transition towards carbon-free energy systems as they allow decision-makers to assess the impact of their choices on the system in place. The solutions provided by energy models can be optimal or not, it depends on their mathematical formulation. There exists two main family of models: simulation, and optimization model. A simulation model is transformed into an optimization model when a degree of freedom is introduced, eg. an input parameter is not fixed, it then become a variable. The problem's purpose is to identify which numerical value of the variable will fulfill the optimality criterion. The optimality criterion of a problem is defined by its objective function, f.e., minimizing the GHG emissions or maximizing the self-sufficiency of a system.

Both of these objectives are used to obtain carbon-neutral systems. Although, studies have shown that achieving the decarbonization of the current system requires a significant restructuration and a deeper electrification of the energy mix, with solar and wind power expected to become the primary source of low-carbon electrical power [3, 4, 5]. Moreover, the decentralization and intermittency of those technologies bring new challenges such as managing the electricity distribution, stabilizing the power grid, and storing energy. Additionally, the multiplication of energy vectors across systems, such as electricity, heat, oil, waste, biomass, and hydrogen, increases the complexity of reaching optimal configuration and operation of the system. Hence, it is common to consider a reduced number of energy vector to fulfill a specific demand. For example, a district heating network can be fed by waste valorization without considering a possible implementation of solar thermal, heat pump (HP)

or boiler unit. Likewise, most models focus on one specific energy sector (electricity, heat, mobility, etc.) and do not account for cross-sectoral synergy [6].

Finally, on a more general aspect, the system's scale varies substantially. As described by the IPCC [7], an energy system is defined as "all components related to the production, conversion, delivery, and use of energy". Following this definition, an energy system ranges from a simple HP to a national power grid. However, a large share of the literature focus on building energy system as the sector accounts for 31% of the final energy demand and has a high potential to reduce its GHG emissions [4].

2.1.1 Scale of energy systems

Building, district, and large-scale energy systems are the most prevalent scales for an energy system. Large-scale models range from a state or a province to an entire continent. Each scale has a specific role to play in the overall energy system, and requires different approaches and strategies to optimize its energy usage and minimize environmental impacts. While building-scale systems seek at supplying their service needs with high reliance and cheap energy cost. District-scale systems aim at improving energy efficiency between buildings, increase renewable penetration and decrease infrastructure usage, and large-scale systems prioritize efficient transmission and distribution of energy across vast geographic regions and accounts for centralized technologies.

Furthermore, systems can be aggregated to a larger scope (systems-in-systems approach), f.e. buildings can be assembled to construct a district representing an energy system by itself. Middelhauve [8] uses the term of energy renewable hub to describe such configuration, where the buildings represent micro energy hubs and are assigned to one macro hub, the district. This kind of configuration represents the future of energy system as the tendency is towards decentralized systems with the development of RE, such as wind and solar technologies [9, 10, 11]. The macro hub addition allows to consider centralized technologies that reduce cost and GHG emissions of the community due to an economy scale [8]. However, such system requires high interconnectivity between micro hubs and energy carriers (fuel cell, gasification process, etc.) to help tackle the stochasticity of RE alongside an optimal operation of the system.

The optimization of the system operation is computationally demanding, so typical periods are introduced to mitigate the complexity. Moreover, an optimal network exchanges schedules should consider the objective of the micro hub in himself and the objective of the entire community. The two objectives might be conflicting therefore a trade-off between the two systems has to be identified by the model. Hence the value to define embedded systems in order to consider a multi-scale optimal solution of the global system.

The limitation of aggregated system is the number of micro hubs as the consideration of multiple network nodes is computationally demanding [12]. In this regard, large-scale energy systems model the demand in a top-down perspective, using pre-defined aggregated energy demand profiles. In the case of building energy systems the approach is bottom-up, i.e. the energy demand is based either on statistical data or physical model. Whereas, the energy demand, in the case of large-scale energy system, is defined in a top-down perspective. The top-down method implies that the energy demand is often derived from socio-economic indicators. Therefore, the energy consumption can only be defined at the aggregated level, leaving out specifics about the energy system [13].

2.1.2 Resolution of energy systems

The spatial resolution of a large-scale energy system is a key feature of the system. It defines the different energy technologies available, energy vectors and the possibility to optimize the system operation, as the model characterization is related to its discretization [5, 6, 14]. Jalil-vega et al. [15] have analyzed the effect of spatial resolution on a large-scale energy system. They concluded that its impact was related to the homogeneity of the region, i.e., a heterogeneous region would be better captured by a thin resolution. Furthermore, their results highlighted that finer resolution improved the network design when using optimization models. In the same fashion, Aryanpur and al [16] have underlined in a review that using a single region to represent a national-scale system did not seize optimally the characterization of heterogeneous region. In this prospect, it is necessary to find a trade-off between spatial resolution and computational time. A suitable method is to define archetypes of meaningful scale to represent heterogeneous systems with a fine resolution. The following sub-sections discuss various sub-system resolution for large-scale energy system.

Building

Since the built environment accounted for a third of the final energy demand in 2019 [4], it is necessary to define large-scale energy system with a building-scale resolution [17]. Moreover, the built environment is quite heterogeneous as it can be spread across multiple meteorological zone, construction decades, density construction, etc. hence the importance to define representative archetypes. Kotzur et al. [18] have developed a bottom-up model based on residential building stock. They have deployed an aggregation algorithm to define archetype buildings. Their configuration and operation, which included the buildings interaction, were then optimized to acquire cost effective solution. The archetypes solutions were extrapolated to represent the German building stock. The model was validated with a case study and emphasized the importance of photovoltaic (PV) and HP deployment to reduce GHG emissions. However, those new technologies tended to increase the electricity availability gap between summer and winter.
De Jaeger et al. [19] explored building clustering using complete-linkage which is an agglomerative hierarchical method. The input data were geospatial data and provided acceptable result to model

the peak and annual energy demand. Stadler et al. [20] built 9 variations of three typical buildings, each with a different construction year, resulting in a different energy demand as the building envelope is correlated to the construction year. The buildings were then individually optimized and extrapolated to represent the built environment of Switzerland. None of the aforementioned studies considered the buildings interaction, which is representative of the current literature.

District

The use of isolated buildings to define energy demand does not consider the buildings interactions within the district. Several studies [8, 21, 22] presented such systems as macro energy hubs where each building is a micro energy hub, all controlled to optimally design and operate at the macro scale. They are also referred to as distributed energy systems in the literature and have been found to contribute significantly to support sustainable growth by promoting sustainable energy access, reducing GHG, creating jobs, and enhancing energy security [23, 24, 25, 26]. As discussed in the review on district-scale energy system by Allegrini et al. [27], there has been a significant improvement of models and tools used to analyze such system. In regard to the district scale, they found that multi-energy grids are critical technologies to improve RE penetration, as balancing demand and supply across the network improved load matching and resource utilisation.

In their paper, Morvaj et al. [28] performed an optimization of a district scale energy system composed of twelve buildings. For each building, an optimal design and operation have been identified and a district heating network associated was optimized to reduce the total expenditure (TOTEX) and the GHG emissions. Their results showed that the decentralization and the increase of energy conversion units availability helped to reduce objectives. Maroufmashat et al. [22] highlighted the importance of considering multiple energy hubs to observe significant cost and GHG emissions reduction. Their case study has showed that the implementation of distributed combined heat and power (CHP) units was limited while operating an electricity grid with low CO_2 emissions. And that the operation of interconnected energy hubs can significantly increase the robustness of the power grid, e.g., mitigation of congestion and ensuring reliability. Therefore the consideration of district systems within wider scale system can be beneficial to improve their optimal objectives.

On the other hand, increasing the size of the considered district has a direct impact on the computational time of the model resolution [29, 30]. Therefore, the use of district scale system to model large-scale energy system become infeasible in term of computational power. Hence, new solving methods are needed to improve optimization efficiency without reducing spatial and temporal resolution, which oversimplifies the problem [30, 31]. Instead, the use of archetype is an efficient way to reduce the computational cost. Eggiman et al. [32] attempted to define neighborhood archetypes, using various feature combinations, and up-scaling technique to define a national-scale energy system. However, the clustering steps of the building stock did not deliver acceptable results, which led to the definition of archetypes by an expert.

Geospatial cluster

Another resolution scale is the geo-cluster, which is interpreted as a region in which the selected indicators (features) are homogeneous. The GE2O project, supported by the European Commission, has defined a panel of energy efficient solutions for each identified geo-cluster in Europe. For their identification, both technological and non-technological aspects of the region, such as the wind generation potential and the age distribution, were considered [33]. Kuster et al. [34] defined 118 different geo-clusters within Europe, accounting for 16 parameters such as building types, climate, and socio-economic indicators. They have developed two tools to help users select appropriate technologies for their location by providing case studies with a similar environment. Slaymaker et al. [35] defined a national model based on interconnected region. The study provided a method for discretizing space using geographic data clustering. The geographical characteristics used were wind, solar and photovoltaic potential, as they are location dependent. As well as demographic characteristics such as population density, urban, agricultural, ecological areas, and distance to the power grid.

2.2 Global sensitivity analysis

The previous section emphasized the importance to develop large-scale energy system models to mitigate climate change. In addition it should even consider sub-systems to ensure a multi-scale optimal system and improve its objective function. However, the resolution of embedded systems would be computationally demanding. In this prospect, the large-scale system should have access to a panel of representative solutions of the sub-system to efficiently integrate them in its scope. To provide representative solutions it is necessary to explore the solution space of the sub-system. There exists several method to do so, such as multi-objective optimization, global sensitivity analysis, etc.. The further section discusses the application of GSA in the context of energy systems.
Sensitivity analysis (SA) is a powerful tool to assess the effect of input parameters on the model output. There exist two main types of SA: local and global. The most common method is the local sensitivity analysis (LSA), which evaluates the sensitivity by varying one parameter at the time around a specific value. Even though, its ease of implementation makes it popular, the sampling scheme does not scan the entire space of input parameters. This gap is filled by the GSA, which explore the solution space by varying several parameters at once. It can thus record the interactions between parameters.

Most of current models assume a perfect knowledge of the input parameters. This absence of uncertainty characterizes them as deterministic models [24]. However, various aspects such as current policy, renewable energy production and economic trends show that there are always some uncertainties to consider. Therefore, specific models are developed to consider such randomness: stochastic model. However, the resolution of the latter is computationally expensive. An alternative

to stochastic model is the application of SA to deterministic models as proposed in [36, 37]. Indeed, given the sensitivity of the model for uncertain parameters, it is possible to assess the general behavior of the model without extensive modifications.

A review on SA methods for building energy systems has been performed by Tian et al. [38] and emphasizes the importance of choosing the right method. They concluded that the choice should be based on the following criteria: research purpose, computational cost of energy models, number of input variables and familiarity with the methods. Another review has been performed by Westermann and Evins [39], 57 studies on building design were analyzed, focusing on: objective, sampling strategy and surrogate model type. Among all studies only 16 included a SA of the model. Their sampling strategies were primarily based on latin hypercube sampling, and only 3 used an optimization model, highlighting the low implementation of SA in optimization models.

Liu et al. [37] developed a framework to assess the uncertainty and sensitivity of district energy systems. The framework includes a two-stage SA. First, a screening method filters out non-influential parameters to reduce computational cost. Secondly, the sensitivity of the influential parameters was evaluated with a GSA. The methodology has been further validated with a case study of a distributed energy system (DES) in China, which indicated that the overall cost is more sensitive to some variables than others, including energy carriers price and the efficiency of thermoelectric generators.

Mavromatidis et al. [36] have applied the same methodology to a DES. The screening method used to identify the most influential parameters was the Morris method. It has been coupled with a Sobol sequence to obtain a proper sampling of the solution space for the final sensitivity analysis. The results showed that the variation from the system optimal cost came from energy carriers' prices and energy demand patterns. Concerning the installed units, a combination of heat pumps, cogeneration units and boiler achieved the lowest cost, and a phase-out of CHP was required to reduce GHG emissions.

Finally, Østergård et al. [40] have used a SA to explore multidimensional design space to help decision making regarding sustainable building design. The identical framework presented as presented in [36] has been used and provided a portfolio of possible building conception. The author developed a tool to select optimal designs based on specific parameter ranges.

2.3 Clustering

Once the solution space is explored it is necessary to identify typical solutions with the use of clustering algorithms, as the solution space can be composed of several thousands of optimizations. Clustering is used to explore data sets and determine their structure. It reassembles data points that have similar attributes into a cluster. Aggregation techniques are classified into two types: supervised and unsupervised (machine) learning. The supervised methods are given an object that has previously been tagged; their goal is to create an algorithm that will link a new object to a label. Unsupervised learning, also known as clustering, on the other hand, is not labeled. The algorithm must identify the groups (clusters) based on the features of each data point. The number of clusters is not necessarily specified by the user. As labels are not pre-identified with unsupervised learning, it is considered more difficult than supervised learning. Unsupervised learning methods can be further divided into two sub-categories: hierarchical and partitional clustering. The partitional algorithms separate the input data space in k clusters. Whereas, the hierarchical clustering method grows a tree-shaped structure of the data, also called dendrogram, to form nested sets of data.

A clustering algorithm can be expressed as a minimization problem in which the objective function represents the sum of the distances between each data point within the same cluster [41]. Lloyd et al. [42] have proposed the K-means clustering technique. It identifies k clusters among the data points using the algebraic mean as the cluster center and minimizes the intra-cluster distance. There exist many different methods to define the cluster center such as K-medoids using the most centered point as the cluster center [43] or K-medians using the median as the cluster center [44].

However, to perform well, these methods require a convex data space, a property which is inherent from the energy model formulation. Consequently, more robust algorithms based on density clustering are essential to aggregate arbitrary shaped data. The data distribution can come from several density functions and with different parameters [45]. Thus, Cheeseman and Sutz [46] have developed AUTOCLASS, an algorithm to detect clusters based on the distribution of the data. Ester et al. [47] have developed a density-based method named density-based spatial clustering of applications with noise (DBSCAN), which is able to detect non-convex shapes among large data sets. Campello et al. [48] have proposed a hierarchical version of DBSCAN named hierarchical density-based spatial clustering of applications with noise, which proved to be more robust to outliers.

The application of the aforementioned algorithms in the context of district energy systems is rare. One can mention the study published by Felsmann et al. [49], which examined multiple clustering techniques such as Single-Linkage, DBSCAN and OPTICS on a district heating network. The results showed that the GHG robustness outperformed the other algorithms. In order to identify typical buildings Stadler [31] has used the K-medoid technique to determine archetypes building among the entire Swiss building stock. The K-medoid was preferred to K-mean as it produced more robust results and identify existing buildings, not a fictive central cluster [50].

Since the choice of the clustering method is a problem in itself, it is a good practice to introduce a cluster validation index to compare several algorithms. Arbelaitz et al. produced an exhaustive comparison of cluster validation indices [51]. 30 indices were compared over several sets of configurations. The researchers created a tool to find the best suited index for a specific application, although the Silhouette index often performed the best. Moulavi et al. [52] introduced a novel index to evaluate the quality of density-based clustering as most of the validation index are developed for globular clusters.

Table 2.1: Literature overview on the resolution of large-scale energy system. The resolution indicates at which scale the investment or the demand profiles were considered. The labels *elec* refers to electricity. The pre-selection of conversion units, the use of energy demand profiles or the generation of scenarios are three approaches to assume the value of a variable and to fix it as a parameter. The column on interdependent systems refers to whether the study considered inter subsystems decisions in the methodology or in the analysis.

Case study		Method			Integration		Analysis		
Scope	Resolution	Model	Sampling	Solution identification	Multi-energy	Extrapolation	Interdependent systems	Tariff Impact	Reference
Region	Building	LP	✗	Scenario	Elec/gas	Profiles	✓	✗	[53]
City	Building	Simulation	✗	Single solution	Elec	Typical buildings (expert knowledge)	✓	✗	[54]
Country	Building	MILP	✗	Scenarios	Heat/elec/gas	Typical buildings	✗	✗	[18]
Country	City	MILP	✗	Scenarios	Elec/H_2	Profiles	✗	✗	[55]
City	City	MILP	✗	Single solution	Elec	✗	✗	✓	[56]
Country	Building	MILP	✗	Single solution	Gas/elec	Typical weather and buildings (expert knowledge)	✗	✗	[31]
Country	Building	MILP	Morris & Sobol	Clustering	Elec/gas/heat	Typical districts	✓	✓	This study

2.4 Gaps and contributions

Most of researches focusing on large-scale energy system use pre-defined or even constant demand profiles; assuming a certain technological integration and missing on the opportunity to model new consumption habits or life style variation. Some attempt to build bottom-up national model based on building archetypes without considering possible synergies within the built environment nor between the national scope and the building scope. Additionally, the exploration of solution space regarding district energy system is poorly studied. The majority of the studies consider scenarios to represent diverse solutions of the district but none consider an holistic strategy to define typical configurations based on its solution space.

This project aims to build a soft-link between a district and a national model in order to incorporate an optimal built environment. To do so the definition of optimal configurations for districts rep-

resentative of the country are necessary. This implementation will provide the tools necessary to address the following research questions:

- What are the typical investment and operation configurations in district energy systems?
  - Identification of the most influential parameters in the decision making
  - Exploration of the solution with a reduced number of optimizations and extraction of typical configurations
- How local energy systems decisions can be integrated in the national infrastructure?
  - Definition of a national energy system model based on a building-scale resolution
  - Validation of the national model based on the built environment
  - Impact assessment of a decentralized investment in the built environment
- What is the impact of the feed-in tariff of PV electricity on the national system configuration?
  - Determination of a threshold feed-in price of electricity for the deployment of decentralized PV production
  - Impact of the deployment of decentralized PV production on the national infrastructure
  - Effect of the tariff condition on the selection of district configurations

Chapter 3

Methodology

This chapter presents the distinct steps used to generate the results. First, a novel framework able to define typical technological solutions of a district energy system is presented. The framework is composed of a global sensitivity analysis to explore the solution space, which is then clustered to identify typical solutions. Subsequently, the methodology describes the modification performed on a national scale energy system optimizer to consider the district solutions. And finally, the typical districts used for the case study are presented.

3.1 Identification of typical district configurations

3.1.1 Solution space exploration

In order to identify robust typical solutions, it is primordial to explore the entire solution space. However, this can become an untrackable task as the solution space's size is directly related to the model complexity. The following methodology is based on the publication of Saltelli et al. [57] which assesses the state-of-the-art of GSA. A sensitivity analysis can be decomposed in four steps:

1. Identification: k input parameters of the model are selected
2. Sampling: the input parameters space is discretized by N samples
3. Evaluation: the model outputs are computed for each sample
4. Comparison: some metrics are derived from the N outputs of the model for the k parameters

As a sensitivity analysis can be quite expensive in term of optimizations, Saltelli recommend using a screening method to identify the most influential parameters of the model within a minor number of evaluations. Hence, the used methodology is a two-fold sensitivity analysis. The primary SA used as a screening is the method of Morris and the secondary SA is the variance-based Sobol method.

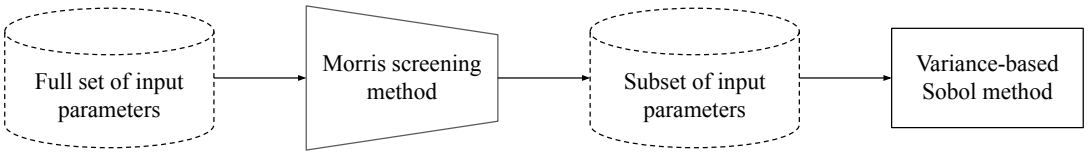


Figure 3.1: Scheme of the two-step GSA composed of first the Morris method as a screening and then the Sobol method is applied on the most influential parameters

Morris

The screening method allows to qualitatively compare the influence on the model output of a large number of parameters with a few evaluations [58]. The method discretizes the input parameters space, which is a k-dimensional hypercube, into a p-level grid, where k is the number of independent input parameters and p is a sampling parameter. Then, it performs a one-step-at-the-time method, i.e. it randomly modifies an input parameter by $\pm\Delta$ to generate r trajectories.
Then, it evaluates the elementary effect (EE) of the i_{th} input factors (EE_i) as a function of the model output $Y = f(\mathrm{X_1},...,\mathrm{X_n})$, see Equation 3.1. The EE can be interpreted as a local partial derivative, thereby it represents the sensitivity of the model at a specific point w.r.t the input parameter.

$$EE_i = \frac{[Y(\mathrm{X_1},\mathrm{X_2},\ldots,\mathrm{X_{i-1}},\mathrm{X_i}\pm\Delta,\ldots\mathrm{X_k}) - Y(\mathrm{X_1},\mathrm{X_2},\ldots,\mathrm{X_k})]}{\Delta} \tag{3.1}$$

Where Δ is defined as a function of p: $\Delta = \frac{\mathrm{p}}{2(\mathrm{p}-1)}$ and can be considered as the size of the discretization mesh . The total number of model evaluations amounts to k(k+1), where r is suggested between 4-10 [57]. The choice of p and r has to be made jointly to ensure that the k dimensions and their interactions are correctly sampled, Saltelli proposed p = 4 and r = 10 [59] whereas Morris used r = 4 in [58], which seems to be the minimum usable value.

In its original work, Morris proposed the computation of the mean μ_i (Equation 3.2) and standard deviations σ_i (Equation 3.4) of the elementary effect distribution for each parameter *i*. However, by doing so, the positive and negative effects cancel each other out, which would falsely influence the results of the mean value. Thus, the method has been improved by Campolongo et al. [60], by considering the absolute mean elementary effect μ_i^* (Equation 3.3).

$$\mu_i = \frac{1}{\mathrm{r}}\sum_{j=1}^{\mathrm{r}} EE_i^j \tag{3.2}$$

$$\mu_i^* = \frac{1}{\mathrm{r}}\sum_{j=1}^{\mathrm{r}} |EE_i^j| \tag{3.3}$$

$$\sigma_i^2 = \frac{1}{\mathrm{r}-1}\sum_{j=1}^{\mathrm{r}} \left(EE_i^j - \mu_i\right)^2 \tag{3.4}$$

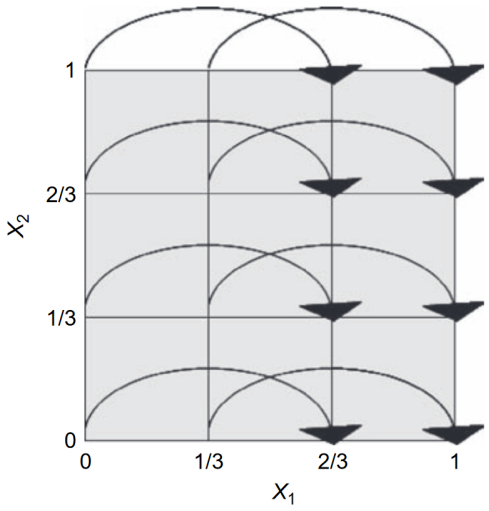


Figure 3.2: Illustration of a k-dimensional hypercube, with k = 2 and a four-level grid (p = 4) [57]. The arrows represent the 8 points required to estimate the EE of X_1.

Those indicators allow to compare the input parameters between each other. A small absolute mean value means a non-influential parameter. The standard deviation reflects the interaction between parameters: a high value means that the output is strongly dependent to the sampling point, i.e. to the other values of the input parameters. Conversely, a low value of σ indicates that the elementary effect is not subject to vary with different parameters values.
The representation of the mean absolute value of the EEs and their standard deviation makes it easy to identify to which group the parameter belongs. As represented on Figure 3.3, one can see the different zones and the line ($x = y$) separating quadrant 1 that defines whether parameters interact together or not. The different zones can be defined as follows:

1. Non-influential parameters
2. Influential, non-interacting parameters
3. Influential, interacting parameters
4. Influential parameters

The comparison of μ and μ^* gives an extra insight in the monotony of the model. If the model output increases with an augmentation of the parameter. The EE will stay positive thus μ^* and μ will have a similar value, whereas if the EE changes sign regularly its cumulative will be lower, i.e. μ will be lower than μ^*.

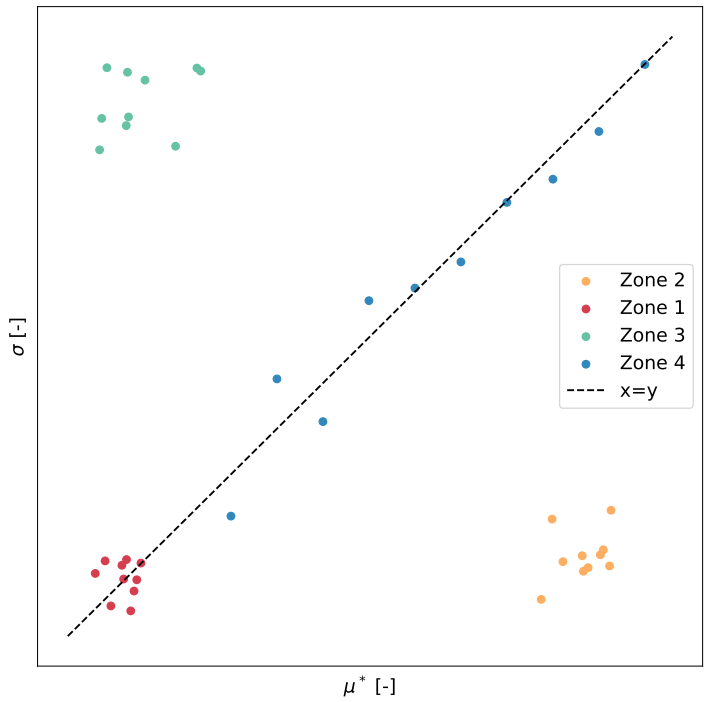


Figure 3.3: Identification of the typical zone on the $\mu^* - \sigma$ plane

Sobol

The Sobol method is a variance-based sensitivity analysis named after the mathematician Ilya M. Sobol. The method is developed in [57, 61] and uses Sobol's recommendation on the sequencing of quasi-random numbers. However, Saltelli extended her work in [62] to reduce the error rate when computing the sensitivity index.

Sensitivity indices

The method evaluates two different sensitivity indices, the first one is the first-order sensitivity coefficient (Equation 3.5). It results from the ratio of the variance of the output mean, considering all parameters except the ith, and the variance of the output. The secondary sensitivity coefficient is the total effect index, i.e. first and higher-order sensitivity coefficient (Equation 3.6).
Assuming: $Y = f(X_1, ..., X_n)$ is the model output, $\mathbf{X}_{\sim i}$ symbolizes all parameters but $\mathbf{X}_i$, $E_{\mathbf{X}_{\sim i}}(Y \mid X_i))$ is the mean of Y for every possible $\mathbf{X}_{\sim i}$ and finally the variance V_{X_i} is calculated for all values of X_i.

$$S_i = \frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)}{V(Y)} \tag{3.5}$$

$$S_{Ti} = \frac{E_{\mathbf{X}_{\sim i}}\left(V_{X_i}(Y \mid \mathbf{X}_{\sim i})\right)}{V(Y)} = 1 - \frac{V_{\mathbf{X}_{\sim i}}\left(E_{X_i}(Y \mid \mathbf{X}_{\sim i})\right)}{V(Y)} \tag{3.6}$$

The first-order coefficient consider only its own effect on the output value, but not the possible higher-order effect with other parameters. The total-order effect index evaluates the effect of a para-

meter considering all possible interactions with other parameters. Meaning that a parameter with a value of $S_T = 0$ can be considered as non-influential on the output Y. The total number of model evaluations is $N(k+2)$ with N being typically between 500 and 1000 [57], it is suggested to choose a power of two.

3.1.2 Clusters detection

This subsection describes the steps to obtain the typical system configurations of a district using the configurations obtained via the Sobol sampling. First, the data have to be standardized, as the results further used to perform the clustering can have a difference of several order of magnitude. Then, the chosen features are discussed and the aggregation method used for this data set are presented. Finally, the clusters are filtered to obtain representative configurations of the solution space.

Standardization of the data & features selection

The chosen standardization technique is the Z-score, see Equation 3.7. Each features X_i of the data set is standardized as follow:

$$Z_i = \frac{X_i - \mu_i}{\sigma_i} \tag{3.7}$$

where μ_i is the mean, σ_i the standard deviation of the feature X_i.
The features chosen for the clustering process consist in a combination of economic and technical attributes. Those features include the capital expenditure (CAPEX) and operational expenditure (OPEX) as key performance indicators. Additionally, the installed capacity of energy conversion and storage units is included in the clustering as it plays a significant role in determining the district configuration. When it comes to the exchange between the district and the network, both the yearly total and peak energy supply and demand for natural gas (NG) and electricity are taken into account.

Aggregation techniques

The selection of the aggregation technique, i.e. the clustering method, is a delicate task which highly influence the results outputs. The choice is based on the methodology described in [63]. As the solution space is non-convex usual clustering techniques such as k-mean, k-medoid, etc. would produce non-optimal results. This property is inherent from the energy system formulation which is a mixed integer linear programming (MILP). Therefore, such techniques are only used to get a sense of the optimal number of cluster. The chosen algorithm is the DBSCAN which aggregates data points based on the density of the data set. The method uses the notion of core points to define a cluster. Points are defined as core points of one specific cluster when they can reach a minimum of *minPts* neighbours within a ϵ distance. Hence the algorithm is used with a fine tuning of its hyperparameter ϵ and a selection of clusters representing at least 5% of the samples.

3.2 Integration of a built environment into a national model

The district scale energy system optimizer used within this project is a Renewable Energy Hub Optimizer (REHO), developed in [31, 64, 65]. Regarding the national energy model, EnergyScope (ES) is used to select an optimal collection of district energy system configuration [6]. This section presents the evolution of both tools and the adaptations to link the micro-energy management within REHO to a national energy system modelled by ES.

3.2.1 Renewable Energy Hubs Optimizer

REHO is a MILP model which evolved with several publications. The initial model was proposed by Luc Girardin in [65] which optimized a building energy system. This model had a unique feature in that the data for the buildings was represented in the form of a GIS. Making it easier to connect to multiple databases and access on site measurements. Paul Stadler further improved the model with modern conversion and storage technologies in its paper on the impact of model predictive control on building energy systems [20]. Finally, Middelhauve et al. [8] proposed a novel decomposition strategy to reduce computational time of the district optimization to allow a centralized optimization of the district in reasonable time among other contributions.

The tool optimizes a district energy system fulfilling electricity and heat demand (space heating and domestic hot water) with energy conversion and storage units (Table 3.1) fuelled by an electricity and a natural gas district grid. The space heating (SH) demand is defined by a 1R1C model of the building detailed by [65]. The demand-supply schedule is optimized on an hourly timescale using typical periods representing an annual time series. Energy and mass balance equations coupled with a heat cascade represents the majority of the constraints.

The model can be optimized in two distinct ways: centralized and decentralized. The decentralized approach optimizes each building individually and then aggregates all buildings within the district using postprocessing. On the contrary, the centralized strategy design and operate all buildings individually and the exchanges within the district networks in a single optimization, reducing total cost and GHG emissions of the system.

Regarding the modelling language, REHO core optimization code is a mathematical programming language (AMPL) which is run via a Python wrapper. The wrapper is critical since it enables: the connection to the GIS database, pre- and post-processing for the cluster identifications. Moreover, the data adaptation (subsection 3.2.3) is also performed using Python.

Table 3.1: District energy system technologies, table from [8].

technology	input stream	output stream	reference unit
energy conversion technologies			
natural gas boiler	natural gas	heat	[kW_{th}]
heat pump	ambient heat, electricity	heat	[kW_e]
electrical heater	electricity	heat	[kW_{th}]
PV panel	solar irradiation	electricity	[kW_p]
energy storage technologies			
thermal storage SH	heat	heat	[m^3]
thermal storage DHW	heat	heat	[m^3]

3.2.2 EnergyScope

ES is a model for strategic energy planning of regional and national energy systems. Firstly, presented by Moret and al [31] the model has known many iterations [66, 67]. The last iteration of the model accounts for several energy carrier grid layers, multi-energy, and multi-sector demands. It allows to point-out precisely which grid level is critical for various scenario.

The model is operated on twelve typical periods representing months. The energy balance is achieved between the end use demand (EUD) of the consumer and the output of energy conversion or storage units converting resources. The EUD are divided in four distinct sectors: households, services, industry, and transportation. The demand is further discretized within each sector in categories such as electricity at various voltage, heat demand at low or high temperature, etc.

In terms of coding language, ES is coded in AMPL and can be launch without any additional program. However, in this project as it is necessary to load the districts configurations an R wrapper is used to load the district configurations and post-process the results.

3.2.3 Data adaptation

In order to enable the connection between the two models some adaptation are required such as temporal transformation and upscaling. Indeed, the original demand profiles of ES are fulfilled by the districts configurations. Therefore, ES can selected which configurations are the most cost-effective in its regard, ie. which technologies should be deployed at the district level. Regarding the temporal adaptation, the generated results produced by REHO are composed of 10 typical day (TD). These TD can be relocated on an annual time series and aggregated to produce monthly results matching ES

inputs data. Finally, as the typical districts represent the whole country, one must adapt the demand in consequence. This is achieved by using an upscaling factor defined by the ratio of the total energy reference area (ERA) of the clustered districts to the ERA of the district itself.

3.2.4 Additional photovoltaic potential

Originally, ES has a decentralized PV potential deployment at the LV level to meet the electrical demand of industry, services and households. However, as the introduced districts meet only the households demands and that the original decentralized PV potential is set to zero, the other sectors cannot meet theirs demands with decentralized solar energy. Therefore, in order to provide decentralized solar energy to each sector and allow further modelling, the remaining PV potential of the district is proposed to ES at the low voltage level in the form of a bidding table between the districts, see Table E.1 for the identified feed-in tariffs by district.

Each district has a maximum PV production profile which is defined from its roofs orientation, meteorological situation, altitude, etc.. Therefore, by annualizing the investment cost of the PV installation over its lifetime, it is possible to calculate the price at which this additional electricity production should be sold. As mentioned, the PV production is different from one district to another, but the investment cost of PV is not, which will result in different electricity prices for each district, generating the bidding table.

3.2.5 Modelling

Please note the parameter, variable and set convention:

- Variables are written in *italic*
- Parameters are written in normal text
- Sets are written in $\mathcal{UPPERCASES}$

District configurations cost

In order to allow ES to select the optimal shares of district configurations, some modifications on the original code proposed by Schnidrig et al. [67] are necessary. The adjustments are based on an additional project from the same author [68]. The following parts describe in detail the sets, variables, constraints, and parameters added to the model.

Two new sets are created to receive the district ($\mathcal{DIS}$) and their possible configurations ($\mathcal{CONFIG}$).

A decision factor (f_c^{ID}) is defined to select the share of configurations for each district. As each district represents a different part of the country, the sum of the decision factor on the district configurations

must be unity.

$$\sum_c f_c^{ID}(d,c) = 1 \qquad \forall d \in \mathcal{DIS}, c \in \mathcal{CONFIG} \tag{3.8}$$

Subsequently, the total cost of the configurations (C_{config}) is defined by the investment in each configuration (c_{inv}^{conf}), calculated based on REHO, and the decision factor. Operational cost is composed of the resources which are already accounted in the resources cost balance.

$$C_{config} = \sum_d \sum_c f_c^{ID}(d,c) \cdot c_{inv}^{conf}(d,c) \qquad \forall d \in \mathcal{DIS}, c \in \mathcal{CONFIG} \tag{3.9}$$

The total GHG emissions of the configuration (G_{config}) is calculated similarly as the total cost.

$$G_{config} = \sum_d \sum_c f_c^{ID}(d,c) \cdot g_{inv}^{conf}(d,c) \qquad \forall d \in \mathcal{DIS}, c \in \mathcal{CONFIG} \tag{3.10}$$

$$\tag{3.11}$$

Additional photovoltaic capacity

Regarding the additional PV potential of household roofs, a new decision factor (f_c^{PV}) is introduced. It is used as a scaling factor multiplied by the normalized maximum PV production profile, see Equation 3.12. The sum of the additional PV production and the PV production within the district, which is defined by the multiplication of the decision factor (f_c^{ID}) and the PV production of the configuration (P_{PV}), is upper bounded by the maximum production for each district in Equation 3.13, as illustrated in Figure 3.4.

$$P_{PV}^{add}(d,t) = f_c^{PV}(d) \cdot \frac{P_{PV}^{max}(d,t)}{\max_t P_{PV}^{max}(d,t)} \qquad \forall d \in \mathcal{DIS}, t \in \mathcal{PERIODS} \tag{3.12}$$

$$P_{PV}^{max}(d,t) \geq \sum_c P_{PV}(d,c,t) \cdot f_c^{ID}(d) + P_{PV}^{add}(d,t) \quad \forall d \in \mathcal{DIS}, c \in \mathcal{CONFIG}, t \in \mathcal{PERIODS} \tag{3.13}$$

This additional production is proposed to ES as a bidding table. The price of the PV electricity ($Price_{PV}^{ID}$) represents the minimum tariff to annualize the cost of the maximal PV installation (C_{PV}^{max}) over its lifetime (l) at an interest rate (i). The same principle is applied to get its construction emissions (G_{PV}^{ID}).

$$Price_{PV}^{ID}(d) = \frac{C_{PV}^{max}(d)}{\sum_t P_{PV}^{max}(d,t)} \cdot \frac{i \cdot (1+i)^l}{(1+i)^l - 1} \qquad \forall d \in \mathcal{DIS}, t \in \mathcal{PERIODS} \tag{3.14}$$

$$G_{PV}^{ID}(d) = \frac{G_{PV}^{max}(d)}{\sum_t P_{PV}^{max}(d,t)} \cdot \frac{i \cdot (1+i)^l}{(1+i)^l - 1} \qquad \forall d \in \mathcal{DIS}, t \in \mathcal{PERIODS} \tag{3.15}$$

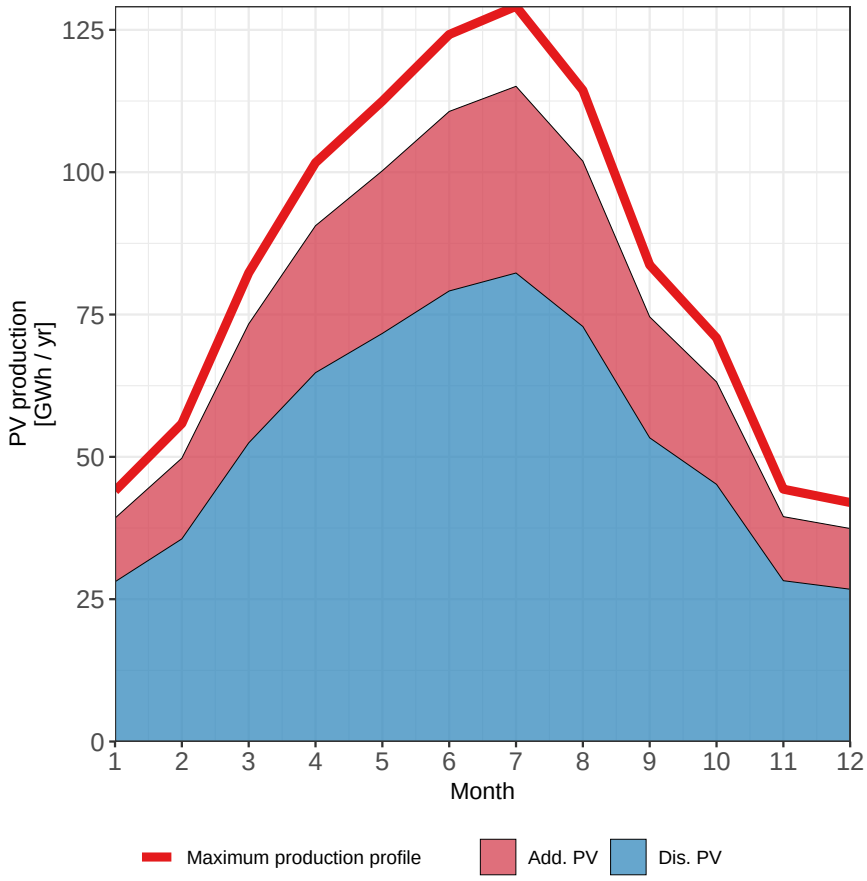


Figure 3.4: Representation of the different PV profiles modelled within ES. The district PV profile is a composition of the selected configurations PV profiles. The maximum production profile is the maximum possible production of the district. And the additional PV profile is the amount of PV ES chooses to install in addition, here it is 70% of the remaining potential of the district

The PV capacity within ES is set to zero as it is replaced by the combination of district configurations and the additional PV capacity. The original resource balance equation is completed by the import and export of the district at the low voltage and pressure for electricity and NG, respectively. The additional PV production of the district is also introduced at the low voltage level. The CO_2 emitted from the construction of the district technologies is introduced into the total emission balance. The newly defined costs and emissions are introduced into the total costs and emissions calculations in an identical manner.

$$C_{tot} = \sum_{tec} C_{inv}(tec) \cdot \frac{\mathrm{i} \cdot (1+\mathrm{i})^{\mathrm{l}}}{(1+\mathrm{i})^{\mathrm{l}} - 1} + C_{config} + \sum_{d} \sum_{t} \mathrm{Price}_{\mathrm{PV}}^{\mathrm{ID}}(d) \cdot P_{PV}^{add}(d,t) \tag{3.16}$$

$$\forall d \in \mathcal{DIS}, tec \in \mathcal{TEC}, t \in \mathcal{PERIODS} \tag{3.17}$$

3.3 Case study

The methodology is applied to 10 typical districts representing Switzerland. The districts were obtained following a methodology developed by Girardin et al. [65] and Loustau et al. [69]. A GIS database combined with a Python repository generates a dataset of building with their respective energy characteristics. Loustau additionally introduced the work of Gupta et al. [70] to spatially estimate the district localisation based on the low voltage grid topology. Then, based on environment, infrastructure and real estate characteristics, a clustering algorithm is applied to select representative districts.

The initial number of identified districts in the country is 175'507. The GaussianMixture aggregation methods were used on the following features: ERA, number of buildings, space heating demand, solar roof space, meteorological zone in the country, grid exchange, irradiance, etc. The optimal number of clusters ranges between 9 and 14. The selection of the specific cluster's results was based on the following criteria.

- The total number of buildings should range between 100 and 400 to apply the configurations identification framework in reasonable computational time.
- The identified district should be widespread across the country as Switzerland is a heterogeneous land.

The chosen clustering is made of ten districts, see Figure 3.5. Their size range from 1 to 104 buildings for a total of 362 buildings. This choice induced a slight bias as the clustering with large districts where not selected to prevent computational burdens. Each district is associated to a meteorological cluster defined by Moret in [71].

3.3.1 Districts analysis

The districts characteristics are presented in Table 3.3. Where one can see the energy reference area (ERA) by district along side the clustered ERA, i.e. the total ERA within the database that is represented by the district. The ratio between the two represents the upscaling factor used to extrapolate the districts optimizations from REHO to the national demand modelled in ES. The total clustered ERA (806.16 [Mm^2]) is only 2% lower than what is reported in [72, 73], this small variation validates that the majority of the built environment is considered. As the district buildings are classified following Schweizerische Ingenieur- und Architektenverein (SIA) regulations, it is possible to define share of the ES sector for each district [74], see Figure 3.6.

The distribution shows that the service and industry sectors are poorly represented in comparison to the households sector. This is a direct consequence of the database selection as the RegBL accounts mainly for the buildings and dwellings. This results is supported with the comparison of the heat and electricity demands obtained from the districts and the national demand declared by the Swiss Federal Office of Energy (SFOE), see Table 3.2. The districts do not represent well the service and

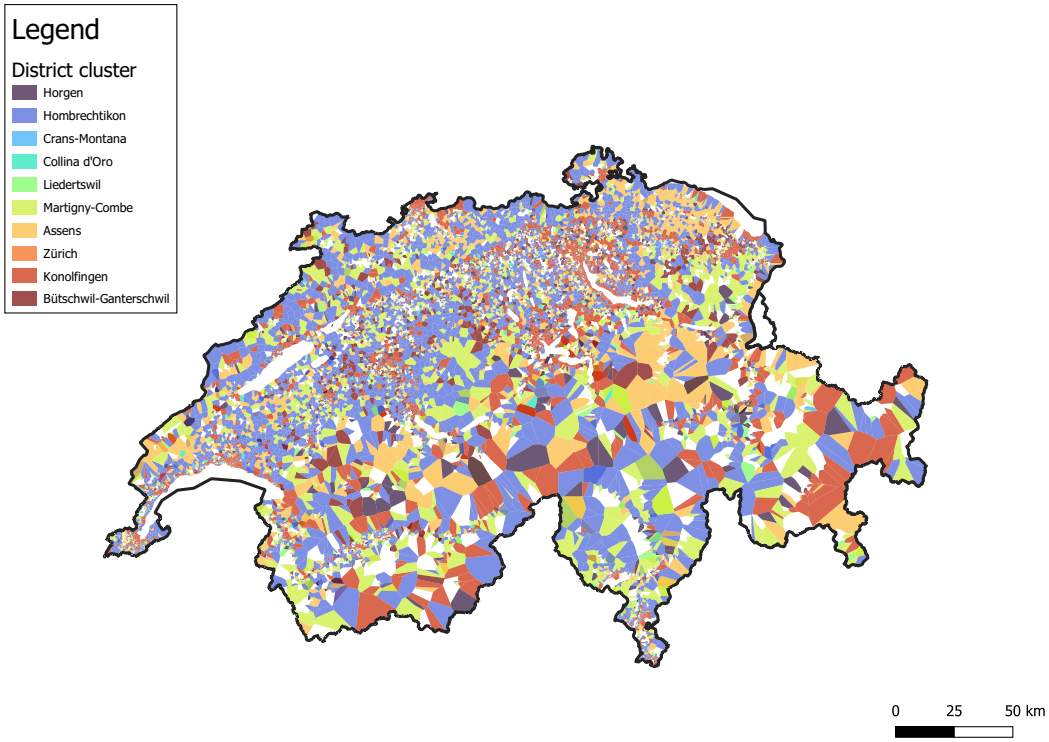


Figure 3.5: Repartition of the district in Switzerland.

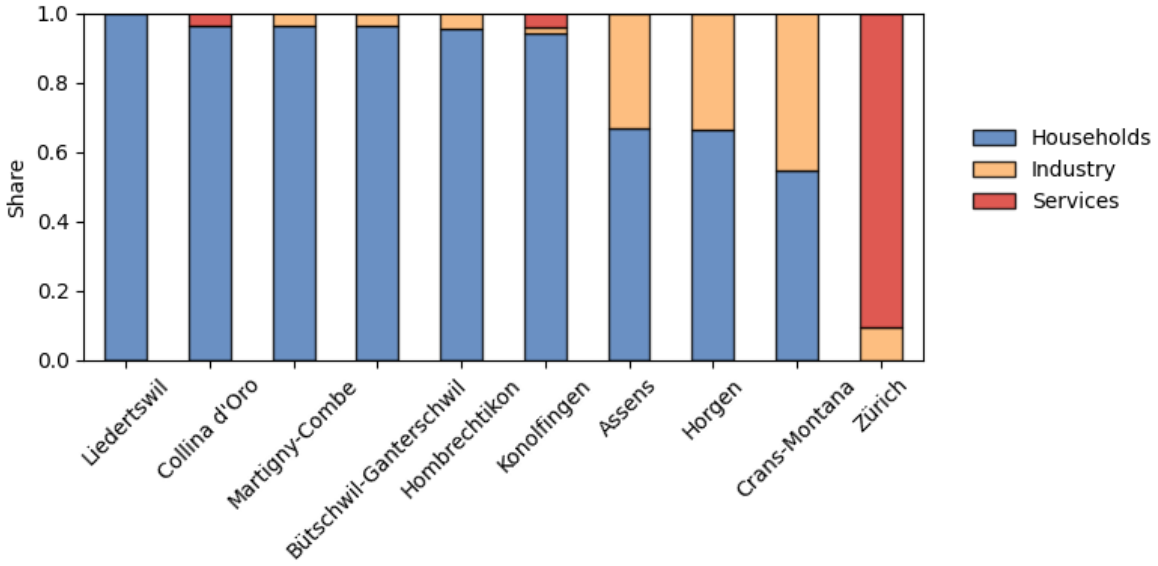


Figure 3.6: Distribution of the sectors within each district using the SIA class defined in the database and description of the class by SIA.

industry low voltage electric demand compared to ES and reference values, therefore they will not be considered being modelled by the introduced built environment. Only the SH and hot water (HW) low temperature demands are fulfilled for all sectors by the introduced districts.

Table 3.2: Modelled and reference energy demands in [GWh/yr]. The energy demands related to REHO represents the upscaled demands of the districts. The electricity demands presented in this table is low voltage and the SH and HW demands are for low temperature demands.

	REHO			EnergyScope			Reference		
	SH	HW	Electricity	SH	HW	Electricity	SH [73]	HW [73]	Electricity [72]
Household	76.13	6.92	16.18	45.6	7.1	18.8	42.11	8.97	20.10
Service	1.64	0.03	0.85	14.1	2.4	15.1	19.43	2.91	17.36
Industry	4.81	0.13	1.59	3.5	0.5	0	3.24	0.76	16.26
Total	82.58	7.08	18.62	63.3	10	33.9	64.78	12.64	53.72

Figure 3.7 illustrates that more than half of the demand comes from the district of *Konolfingen*, and represents 56% of the total clustered ERA. Therefore, its configurations will have a major impact on the final energy system. Additionally, *Hombrechtikon* is the second largest district with 26% of the total clustered ERA, the two combined represent more than 80% of the total demands: Electricity, SH and HW. Therefore, the identified configurations of those districts and the ones selected by ES will have a major impact on the national energy system.

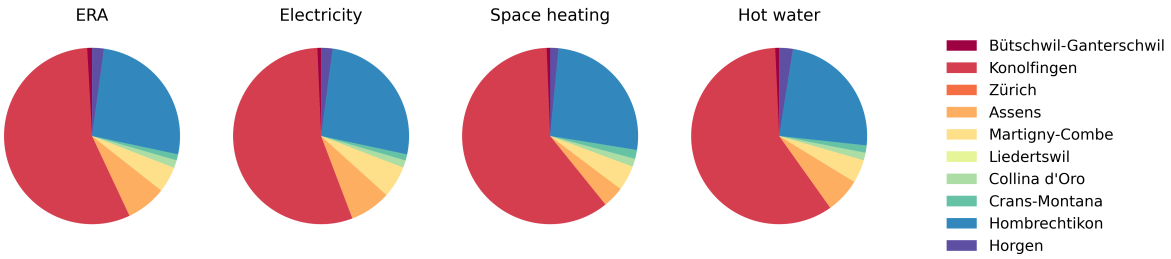


Figure 3.7: Repartition of the ERA and demands by district.

3.3.2 Scenario definition

The EnergyScope base national scenario envisions a Swiss energy system in 2020 that is not only economically optimal but also environmentally sustainable, with no net emissions and complete independence from energy imports. Despite the absence of nuclear power plants, which are being decommissioned in preparation for a complete phase-out by 2050, the scenario is intended to meet the country's energy needs in an efficient and effective manner.

Table 3.3: Characteristics of the typical districts for Switzerland

District name	Buildings [-]	ERA [m^2]	Upscaling factor [-]	Clustered ERA [Mm^2]
Liedertswil	1	94.86	316.0	0.03
Zürich	2	1'079.37	204.0	0.22
Horgen	10	6'318.42	2'752.0	17.39
Assens	16	6'894.09	8'616.0	59.40
Martigny-Combe	16	3'316.38	11'516.0	38.19
Crans-Montana	31	21'326.77	416.0	8.88
Bütschwil-Ganterschwil	43	16'933.44	428.0	7.24
Collina d'Oro	45	15'260.37	750.0	11.45
Hombrechtikon	94	33'231.69	6'353.0	211.11
Konolfingen	104	80'562.18	5'614.0	452.25
Total	362	185'017.57	-	806.16

Chapter 4

Results and discussions

4.1 District configurations

This section presents the most influential parameters obtained through the Morris and the Sobol method, and the typical configurations obtained for each district.

4.1.1 Screening method

This subsection presents the results obtained during the screening phase of the sensitivity analysis. To reduce the number of input parameters considered for the global sensitivity analysis, the Morris method was employed, as explained in subsection 3.1.1. The energy conversion and storage unit parameters, as well as the energy carrier prices, were the parameters considered in this step, see Appendix A for more details on these parameters. A total of 46 parameters were included in the analysis, and following the sampling methodology recommended by Morris, 460 optimizations were performed to compute the sensitivity metrics. To compare the optimizations, a specific output value was chosen, namely the total expenditure assuming a economically rational behaviour, as each run generated a complete system configuration and operation.

The results of the Morris method applied for each district can be found in Appendix A. However, in order to get an overall vision of the most influential parameters of the model, an average ranking of the parameters based on the absolute mean value of the EE distribution is presented in Table 4.1. The retail tariff of the electricity and natural gas grids are the most influential parameters and in a consistent manner as they place first and second 8 out of the 10 districts, respectively. Naturally, the variable cost and baremodule of the technology are influential parameters as they impact directly the objective function. The fact that the parameters related to photovoltaic and heat pump technologies are highly ranked highlights the cost competitiveness of an electrified system. Conversely, the presence of natural gas boiler parameters confirms the low cost of this technology.

Technology	Parameter	Average ranking
Electricity grid	Retail tariff	1.2
Natural gas grid	Retail tariff	2
PV	Baremodule	4.3
PV	Variable cost	5.3
Heat pump	Baremodule	6
Natural gas boiler	Baremodule	7.6
Natural gas boiler	Fixed cost	8.4
PV	Efficiency	9.1
Heat pump	Variable cost	9.8
Heat pump	Fixed cost	10.4

Table 4.1: The ten most influential parameters over all districts with respect to the Morris method

As discussed in subsection 3.1.1, it is recommended as a good practice to graph the absolute mean value (μ^*) and standard deviation (σ) of the EE distribution for every input parameter. Figure 4.1 shows such plane for the district of Crans-Montana. A persistent result is the non-interactivity of the electricity retail tariff always located in the lower right half, whereas, most of the influential parameters sit on the non-interacting zone.

4.1.2 Sobol sampling

This subsection focuses on the final sampling of the chosen input parameters using a Sobol sequence. The input parameter space is explored with 1024 samples following recommendations presented in subsection 3.1.1. The chosen parameters used to explore the solution space is the retail tariffs of NG and electricity. The number of trajectory used is 256. In order to reduce the computational time of the space exploration, by reducing the number of optimisations. The effect of the decentralized method within REHO and a reduced number of trajectory has been analyzed in Appendix B. The conclusion is that the number of trajectory can be reduced by a factor 2 without significant reduction of the sampling resolution. Additionally, the decentralized method can be used to reduce the computational time. However, this reduce the configurations diversity, ie. there is less configurations, but the nature of the configurations is preserved.

The Sobol's sensitivity indices cannot be computed as some of the sampling point causes a crash of the optimizer. Therefore, no objective values were obtained. As these indices are secondary results, no model correction was undertaken.

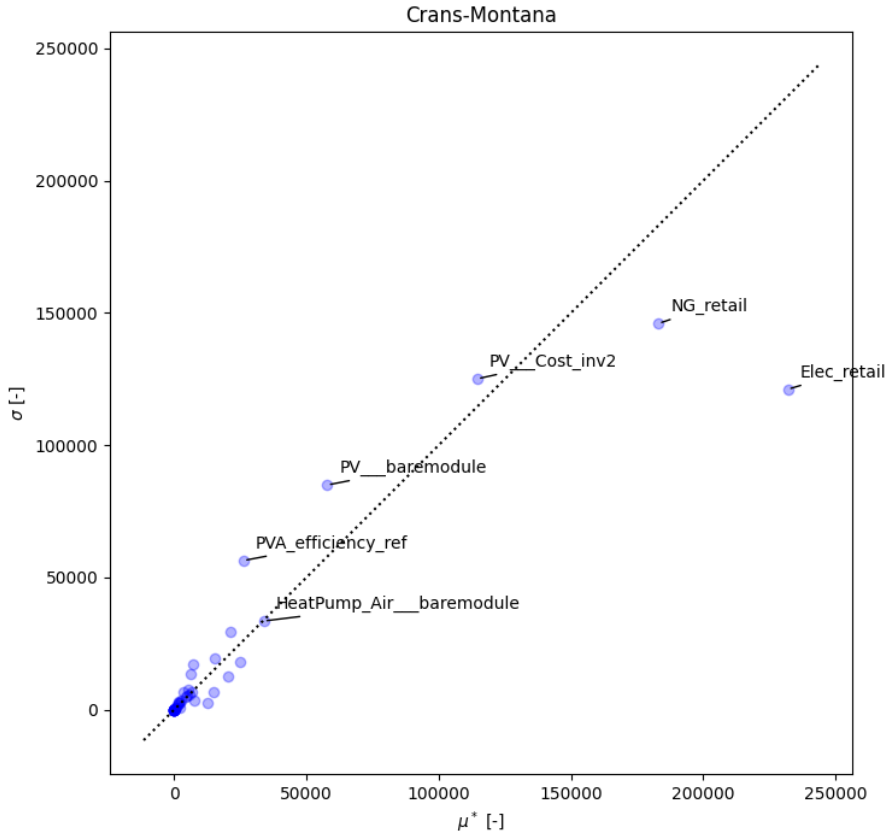


Figure 4.1: Morris analysis for the district of Crans-Montana on a $\mu^* - \sigma$ plane

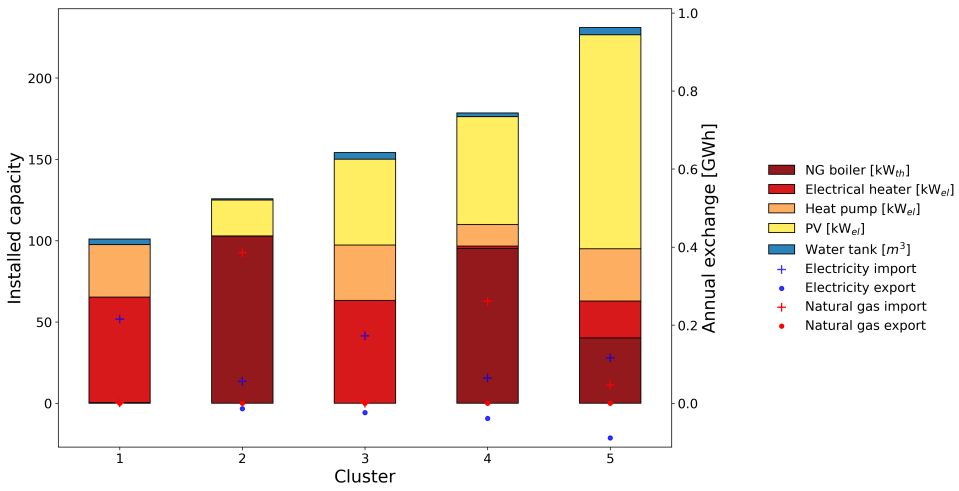


Figure 4.2: Identified configurations for the district of *Martigny-combe* and their grid exchanges

4.1.3 Clustering

The clustering algorithm identified between 3 and 7 typical configurations for the ten districts, accounting for a total of 46 configurations. All districts have at least one configuration based on NG and one on electricity. However, as demonstrated in Appendix C, the hyperparameter ϵ of the DBSCAN algorithm seems to need some adjustments in function of the district size. Large district, ie. dozens of buildings, solution space were better separated with low ϵ value whereas smaller district configurations were well identified with high value.

Figure illustrates the available configurations for the district *Martigny-Combe* associated with the meteorological cluster of *Piotta*. The configurations panel showcases both electric and fossil-fuel based solutions but also an hybrid solution, number 5. The latter probably operates the NG boiler only in high demand situation as its NG import is relatively low.

The main variation between configurations is the total installed capacity, ranging from single to double. This discrepancy is due to the extreme period (high demands and rash environment) included in the model. The model installs a minimum heating capacity to supply heat in any condition. This minimum heating capacity appears in all configurations and is either composed of a NG boiler or a combination of electrical heater and heat pump.

The extensive utilization of HPs and PV in various configurations highlights their high potential in district energy systems. Although, the electric grid is more strained with PV implementation as export are increased, but import are not necessarily reduced, requiring a sufficient absorption capacity of the grid. The installation of a HP triggers the deployment of a water tank to serve as heat buffer.

Figure 4.3 represents the distribution of the identified clusters over the sampling space. The heat map in the background illustrates the amount of natural gas imported for each optimization. As observed

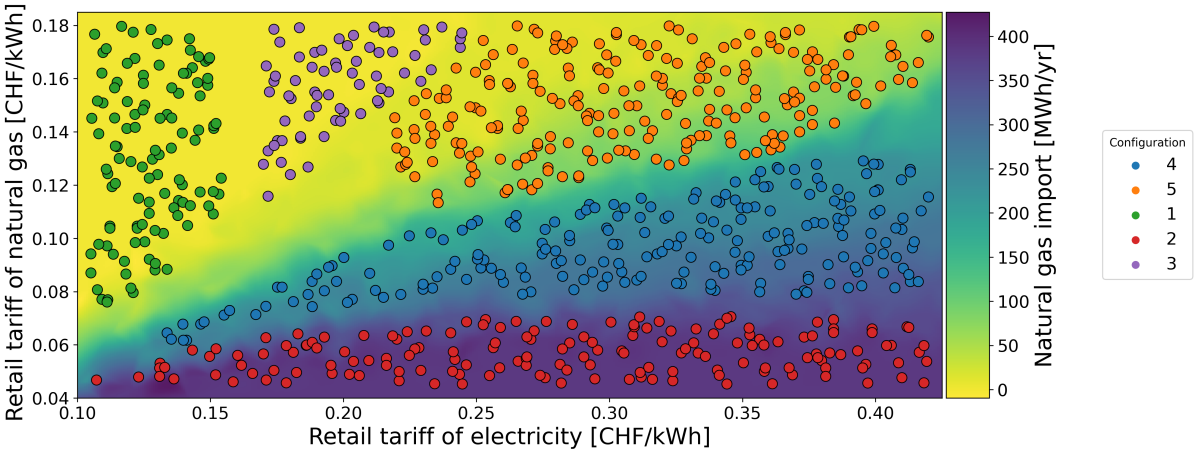


Figure 4.3: Typical configurations distribution in the retail tariffs space for the district of *Martigny-combe*

in subsection 4.1.2, the output of the model, i.e. the district configuration, is strongly correlated to the retail tariff of NG and electricity. This relation can be observed in the figure below as the different configurations can clearly be identified. Additionally, the sampling space naturally separates itself in half, the separation is highlighted by the amount of NG imports. The configurations 2 and 4 are based on NG and the configurations above, 5 to 10, on electricity. The NG configurations are located on the bottom right corner where the electricity tariff is high. Inversely, the electricity based ones are in the top left region where the NG price is high. One can note that the space has less samples in some region, this is due to the data selection done in subsection 3.1.2. The hybrid configuration is justified by high retail tariffs, top right corner, which prevent the configuration to be based solely on an energy vector.

4.2 Validation of the national system based on the built environment

The results from ES incorporating the optimizations of the built environment of 2020 (SREHO2ES) is compared to the Swiss ES (SES) of the same year. The main difference between the two models is the means used to satisfy the low temperature SH and HW demands, as well as the electric demand at low voltage (LV) for the households sector. In SREHO2ES, those demands are met within the district optimizations performed in REHO, whereas ES technologies are used in the case of SES.

There is a 15% difference of the annual cost in favor of SES, which is explained by the SH demand modelled by the district which is 23% higher and the oversimplification of the problem when using a centralized approach.
Figure 4.4 represents the decomposition of the annual cost per capita for both models, accounting for a Swiss population of 10 MCapita. The difference in operational costs is due to the amount of wood purchased which is more than double in the SREHO2ES model. This additional wood is further gasified to produce NG in order to feed the households boiler installed within the districts optimizations.

The maintenance costs have an important disparity inherent from the modelling approach of the maintenance cost in REHO. There is no maintenance cost in the district optimization as it is incorporated into the investment cost. Therefore, all investments in district technologies are overvalued in comparison to ES technologies cost. In this regard, the maintenance cost of ES PV and HP accounts for 43% of the total maintenance cost of SES. As a result, the maintenance cost of the SREHO2ES model represents the maintenance cost of the national infrastructure. The main systemic difference is the investment in gasification technologies to produce the NG from wood, for SREHO2ES, and in geothermal district heating network in SES. The heating sector is detailed in subsection 4.2.1, its investment gap is explained again by the maintenance cost modelling strategy. The wind and PV annual costs are similar; however the PV cost comes from REHO optimizations hence the maintenance cost is already represented, meaning that the SREHO2ES model installs less PV panels than SES. The *Others* category accounts for the rest of the investment that does not fit into previous categories. The principal contributors for SREHO2ES are the battery-electric vehicle. For SES, in addition to battery electric vehicles, there is a district heating network that replaces the districts configurations and an investment in deep saline carbon capture and storage technology. The overall difference comes from an important investment into medium range battery-electric vehicles in the SREHO2ES model that can absorb the electricity produces at LV by the district.

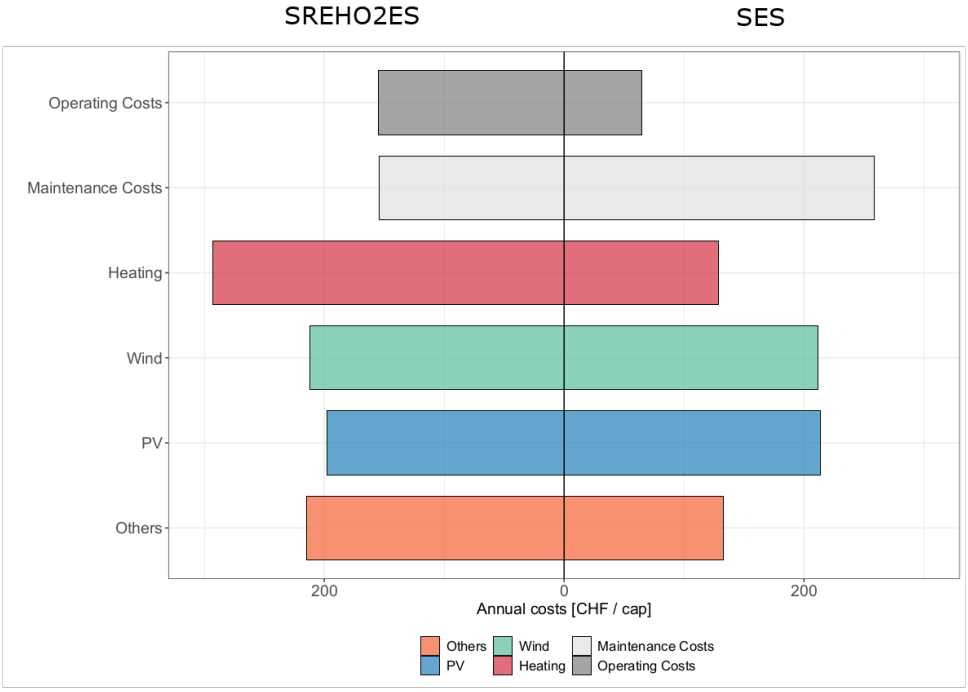


Figure 4.4: Comparison of the investment per capita for the SES and the SREHO2ES model

4.2.1 District technologies comparison

Figure 4.5 illustrates side by side the detailed annual costs of the heating category represented in Figure 4.4. The category is composed of district technologies used to respond to the heating demand, which is composed of low temperature demand for SH and HW and high temperature demand for industry. The difference of 164 [CHF/cap] is partly due to the maintenance cost, which in the case of the ES HP used in SES, accounts for the same annual cost as the investment. Moreover, the SREHO2ES is implemented with 2020 technologies costs at the district level, whereas the ES technologies have 2050 projected costs, increasing the cost discrepancy between the technologies of the models. The use of water tank in the SREHO2ES model derives from the finer temporal resolution used to optimize the district. As the operation is based on an hourly timescale the implementation of a thermal buffer is required to optimally redistribute the heat produced by coupled HP and PV. SES implements two district scale technologies for the heating network, electrical heat pumps and geothermal units. Note that such technologies are not modelled in REHO and only district large enough would justify their installation. The rest of the investments is distributed between electrical heaters and waste boilers. Although the total costs are different, the technologies used are the same and in the same proportions. Note that the PV is not represented in this graphic as more detailed results are presented in subsection 4.2.2.

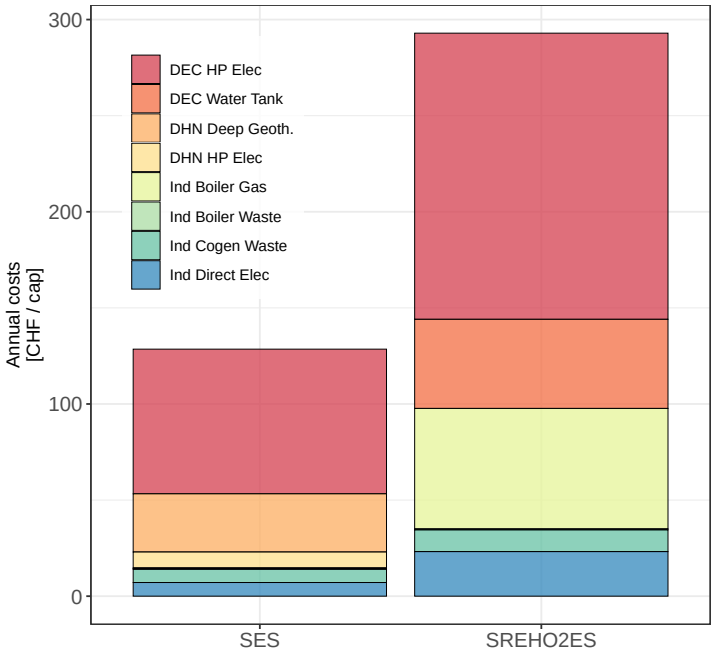


Figure 4.5: Comparison of the investment per capita for the heating category of the SES and the SREHO2ES model

4.2.2 Geolocalization of PV deployment

As described in subsection 3.2.4, there is two distinct types of PV: one installed during the districts optimizations and another installed by ES using the district PV characteristic to supply the demand. The district configurations use at best 64% of the maximum production capacity in the district of *Crans-Montana*. For other districts, the utilization is even lower with an average of 6.4% of the total capacity. The district of *Konolfingen*, which represents more than half of the country's built environment, installs a quarter of its total PV capacity. *Liedertswil*, second most representative district, does not install any PV, probably because only one of its configurations has PV, see Appendix D. The detail share of PV production potential is illustrated in Figure E.3 in Appendix C. ES invests in additional PV only in the district of *Crans-Montana* as its bidding offer was the cheapest with 0.107 [CHF/kWh], the second cheapest option was in *Konolfingen* with 0.121 [CHF/kWh]. This behaviour signifies that ES can supply electricity at LV at a lower tariff either by using other energy conversion units or choosing configurations exporting electricity at least cost.

Until then only the investments in the technologies have been compared but as mentioned their costs are different within REHO and ES. The annual PV production for SREHO2ES and SES are 12.66 and 34.85 [TWh/yr] respectively. This difference comes from the technology characterization in the two models. Moreover, the monthly time resolution of the SES model erases the daily stochasticity of PV, which, in reality, is a major limitation of its penetration in the system.

Figure 4.6 represents the PV production potential share, ie. the amount of electricity produces with respect to its maximum production capacity, on the Swiss map. The main observation is the globally low penetration of PV over the territory with the exception of the district of *Crans-Montana* represented in red. The penetration is higher in urban area, ie. on lake shore and along rivers, which are represented by the district of *Konolfingen*. More maps representing district technologies production are available in Appendix C.

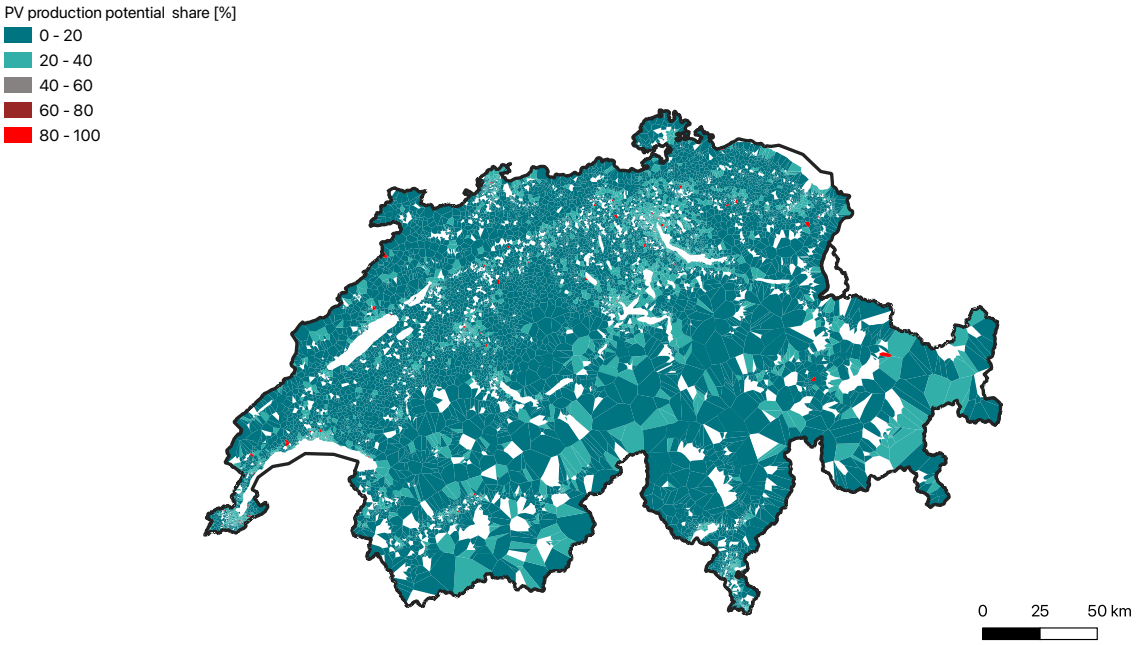


Figure 4.6: Distribution of the PV production potential share over the Swiss territory. Penetration of PV is higher in urban areas.

4.3 Impact of the feed-in tariff on the national system

This section analyses the role of the feed-in tariff of PV electricity on the behaviour of the model. The feed-in tariff is parameterized to vary from 0.01 to 0.12 [CHF/kWh]. Figure 4.7 illustrates the evolution of the annual cost composition and storage capacities of the system. The focus of the figure is the PV investment variation, which was previously referred to as additional PV. As for now on the PV within the district is incorporated into the district services category. The investment in PV increases until it hits a tipping point around 0.08 [CHF/kWh] and starts to decrease until reaching no investment for a tariff of 0.115 [CHF/kWh]. Another investment that fluctuates with the feed-in tariff is the district services which indicates a variation in the choice of the districts configurations. As the price of electricity increases and the model is not willing to purchase it anymore, it is necessary to either invest in new technologies or select district configurations which are more self-sufficient or even export electricity.

Regarding the storage hydro, it is at its maximum for the whole variation range. Whereas the NG storage is negatively correlated with the feed-in tariff as cheap electricity access pushes the system to use NG as a reserve resource. This storage capacity is used as a seasonal storage as there is no reduction of technologies' capacity fueled by NG. Additionally, the total annual cost is correlated to the feed-in tariff, as the access to cheap electricity becomes limited.

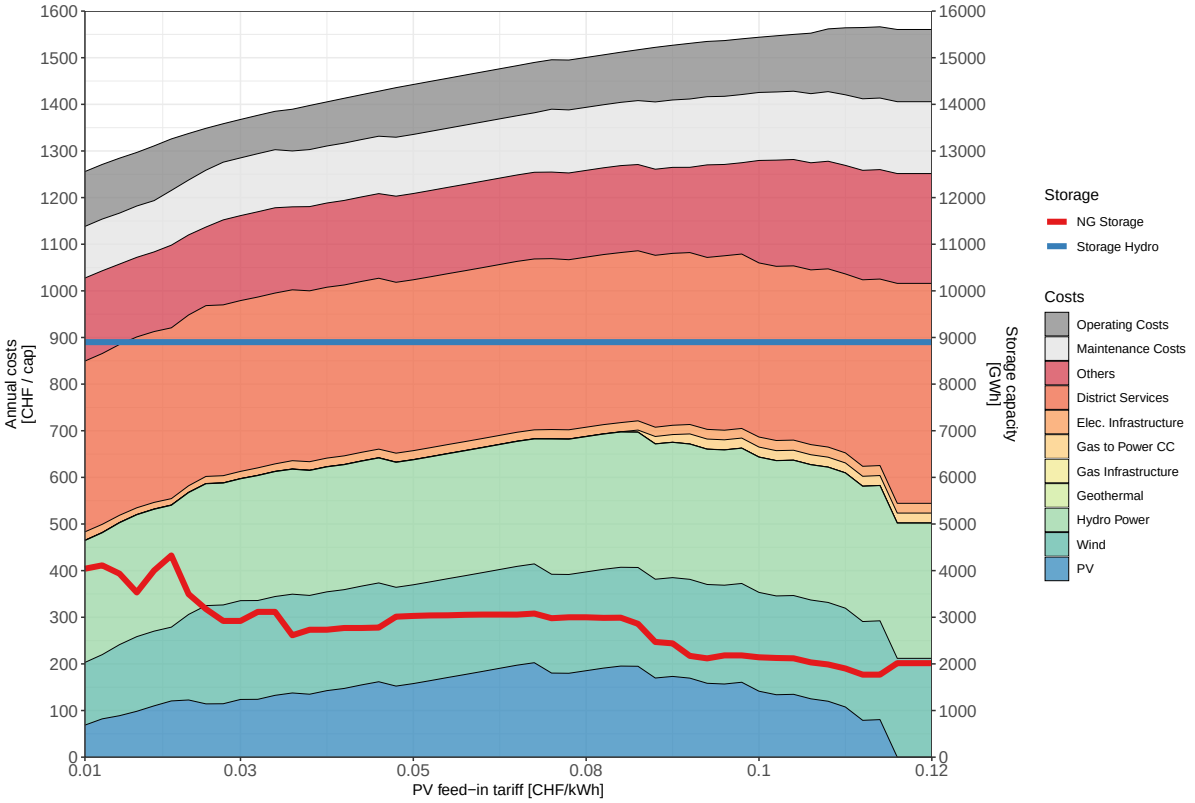


Figure 4.7: Evolution of the annual cost with the feed-in tariff of electricity.

4.3.1 Adaptation of the district configurations

From Figure 4.7, it is important to decorrelate the investment from the PV production as the feed-in tariff is not fixed. Therefore, even if the investment in the technology increases, this does not mean that the production increases.

Figure 4.8 represents the cost for both categories of PV: within the district and the additional capacity proposed to ES, and their respective output. The additional PV production is decreasing with the augmentation of the feed-in tariff even when the investment increases. ES invests in additional PV until it reaches 0.08 its tipping point where it starts to decrease. It corresponds to a shift towards NG based configurations as represented in Figure 4.9. The electrical heater and HP investments drop suddenly at the expense of NG boilers.

The additional PV capacity is latter replaced by district PV production at a tariff of 0.115 [CHF/kWh], which might be compared to the levelized cost of electricity (LCOE) within the selected district configurations. This value represents a threshold after which the system can supply cheaper electricity than the feed-in tariff, either via the district configurations or ES technologies.

The total PV production reduces with the increase of the feed-in tariff until the district takes over but

the overall production is drastically reduced. Therefore, one would expect a reduction of the electric infrastructure cost with the refraction of PV deployment, as a high production of electricity at LV would most likely be used to fulfill medium voltage (MV) demands.

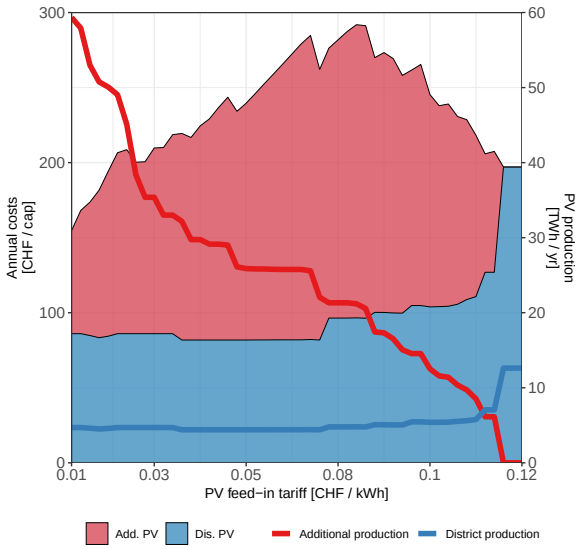


Figure 4.8: Evolution of the PV production and investment with the feed-in tariff.

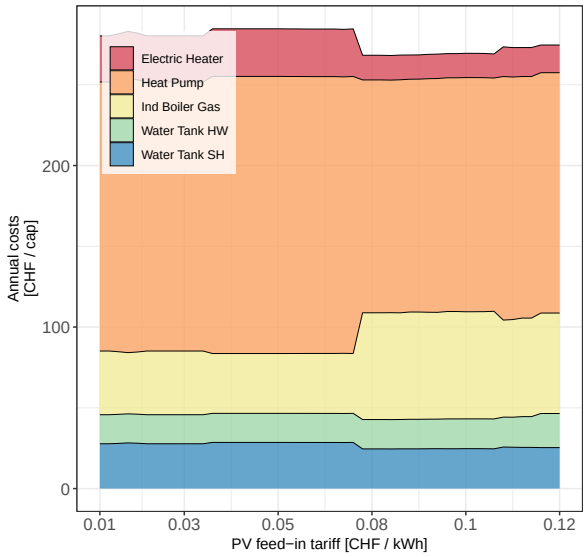


Figure 4.9: Evolution of the investment in district technologies with the feed-in tariff.

4.3.2 Evolution of the electric infrastructure

Figure 4.10 represents the evolution of the installed capacities of the electric infrastructure with the feed-in tariff variation. The total capacity decreases until a tariff of 0.05 [CHF/kWh], after what it stabilizes with the augmentation of the tariffs. As the PV output is fed into the LV level of the network, once the LV demands are fulfilled, it is transferred to the MV. Low feed-in tariffs generate such situations. The consequences on the network are an increase in capacity of superior levels to supply higher voltage demands. At the LV level, the demand and supply profiles cancel each other out as a monthly resolution is used, which reduces the required network capacity. As the tariff increases, the PV production is reduced and new technologies are introduced to provide electricity at the MV in order to fulfill the district imports at LV and high voltage (HV) demand. During this transition, the MV grid capacity is reduced and stabilizes at 0.05 [CHF/kWh], whereas the LV capacity is continuously increasing with the tariff. Regarding the HV and extra high voltage (EHV) grids, their capacity is decreasing because they were used to transfer the surplus electricity production to the dams connected to the EHV grid for seasonal storage. The spikes of the capacity are due to some instability of the model and can be disregarded as the resulting investment price correct, see Figure 4.7.

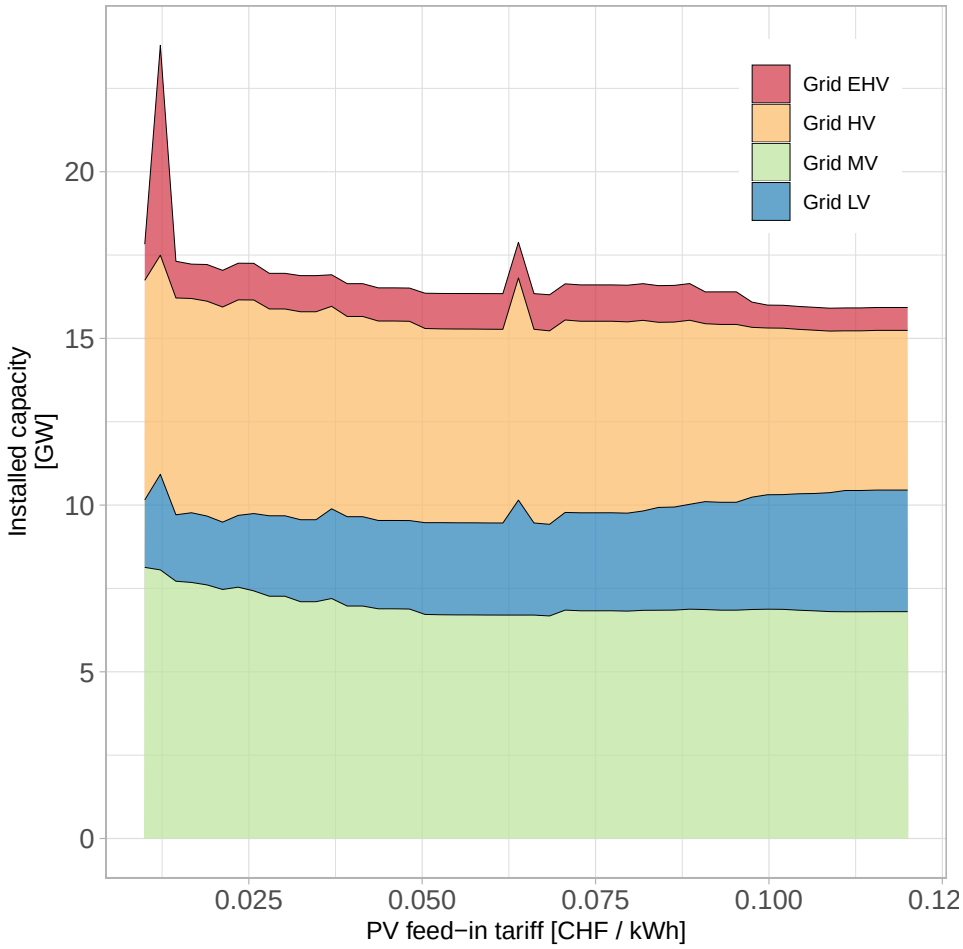


Figure 4.10: Evolution of the electric grids capacity with the feed-in tariff.

4.3.3 Identification of the district levelized cost of electricity

The variation of the feed-in tariff allows to analyze the activation of the additional PV capacity. A threshold tariff above which ES does not invest in this additional capacity can be defined for each district. Figure 4.11 illustrates this threshold value on the Swiss map. The district of *Konolfingen*, which represents the urban areas, has a lower threshold than the rest of the districts. However, a lower cut-off tariff does not mean a lower PV deployment. As illustrated in Figure 4.6, *Konolfingen* is the district with the second highest PV penetration for a feed-in tariff much higher than its threshold. As discussed in subsection 4.3.1, this threshold value can be compared to the LCOE of the different configurations. As the deployment of PV within a district produces electricity at least expensive [8]; the LCOE in urban areas, as a result of subsection 4.2.2, is therefore lower than in rural district. This justifies the presence of higher threshold rate in rural regions.

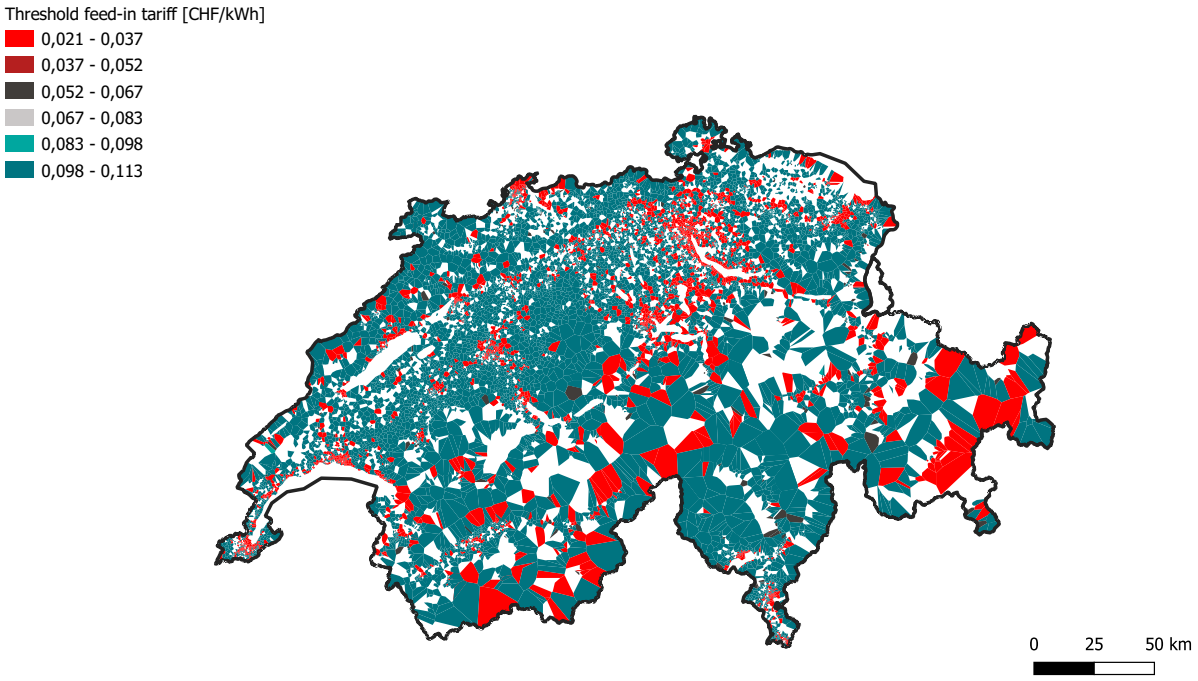


Figure 4.11: Representation of the activation threshold feed-in tariff of electricity in the district.

4.4 Outlook

The outlook is discussed for the main steps of the research. First the districts configurations identification can be improved to provide more accurate configurations to the national model. Then the quality of the initial districts representing Switzerland can be improvement as they have a direct impact on the final results. Finally, some harmonisations between the technological costs of the district and the national model might provide additional insights on the models interaction.

First, as the ideal approach would be to let ES define the districts configuration, it is important to propose the most representative and diverse panel of configurations. As discussed in the B, the decentralized optimization reduces the number of identified configurations. Therefore, additional computing power would allow to use the decomposed method and identified a larger amount of configurations. What is more, the configurations characteristics (key performance indicators (KPI), operation schedule, etc.) are defined by averaging the districts of the cluster. An alternative would be to do a final optimization of the district with the capacities of all buildings units fixed to the averaged values of the cluster. By doing so, the KPI used would correspond to a *truly* optimal district solution and not an average of optimums.

Secondly, the typical districts integrated to the framework represented poorly the built environment of Switzerland as only two districts amounted to 80% of the energy demand. The introduction of the municipality environment classification of the district as proposed in [75] as well as considering the network infrastructure density could help improve the results.

Finally, the difference in the technologies cost composition and definition presents an issue to represent the operational and maintenance cost. Most importantly, the choice between deploying centralized or decentralized energy sources is deeply biased if theirs cost are defined at different horizons. In our case, the technologies defined in REHO have actual market costs, while ES uses learning curves to project costs to 2050. The consideration of the units already deployed within the building stock and at large-scale would provides results more accurate to achieve an independent and carbon neutral system.

Chapter 5

Conclusion

This research presented the impact of renewable energy hubs configurations on a national energy system. First, a panel of optimal district configurations has been identified for 10 typical districts representing Switzerland. The built environment is then integrated in a large-scale energy system model and compared to a baseline model. Finally, a case study regarding the impact of feed-in tariff on the national infrastructure and the districts composition was performed. The main outcomes of the project are listed in the following.

The district technological solutions are predominantly defined by the energy carriers tariffs. Indeed, the gas and electricity retail tariffs are the most influential parameters of the model and were further used as sampling parameters for the global sensitivity analysis. As a result, the typical decisions taken to design the district energy systems were clearly distinguishable based on their energy exchanges. Moreover, the nature of the configurations were either based on gas or electricity depending on their location in the sampling space, ie. tariff condition.

The integration of district configurations in the national model produces a 15% increase of the annual cost to reach carbon neutrality and an energy independent system compared to the model of reference. This difference is due to the space heating demand being 23% higher within the built environment. Therefore, both approaches provide similar results, but the decentralization informs on favourable environments to develop RE. The PV capacity of the decentralized approach represents only 36% of the centralized approach. However, the results emphasized the importance to deploy PV panels within urban areas to ensure on-site production.

The impact of the feed-in tariff variation on the system shows the utilization of natural gas as seasonal storage to mitigate the PV electricity production gap between summer and winter. The investment in PV reaches a tipping points at a tariff of 0.08 [CHF/kWh]. Afterwards, the model selects more gas based configurations for the districts and reduces its investment in the additional PV production. Finally, the model identifies a threshold feed-in tariff for each district, above which no additional PV capacity is installed as the district is able to produce cheaper electricity, ie. the LCOE within the dis-

trict is lower than the feed-in tariff.

The implementation of the built environment in the national model provides relevant insights on the possibility to develop decentralized renewable energy capacity at the district level. Additionally, it allows to distinguish the required investment for the consumer and the national infrastructure to reach carbon neutrality.

Bibliography

[1] Hannah Ritchie, Max Roser, and Pablo Rosado. CO and greenhouse gas emissions.

[2] GlobalABC, IEA, and UNEP. GlobalABC roadmap for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector.

[3] Daniel M. Kammen and Deborah A. Sunter. City-integrated renewable energy for urban sustainability. 352(6288):922–928.

[4] P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. \{McCollum\}, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, and J. Malley. Climate change 2022: Mitigation of climate change. contribution of working group III to the sixth assessment report of the intergovernmental panel on climate change.

[5] Hans-Kristian Ringkjøb, Peter M. Haugan, and Ida Marie Solbrekke. A review of modelling tools for energy and electricity systems with large shares of variable renewables. 96:440–459.

[6] Gauthier Limpens, Stefano Moret, Hervé Jeanmart, and Francois Maréchal. EnergyScope TD: A novel open-source model for regional energy systems. 255:113729.

[7] IPCC. AR5 synthesis report: Climate change.

[8] Middelhauve, Luise. On the role of districts as renewable energy hubs.

[9] Towards energy-autonomous cities using CO2 networks and power to gas storage. Meeting Name: 29th International Conférence on Efficiency, Cost, Optimization, Simulation and Environmental Impact of energy systems.

[10] Samuel Henchoz, Céline Weber, François Maréchal, and Daniel Favrat. Performance and profitability perspectives of a CO2 based district energy network in geneva's city centre. 85:221–235.

[11] Tillmann Lang, Erik Gloerfeld, and Bastien Girod. Dont just follow the sun – a global assessment of economic performance for residential building photovoltaics. 42:932–951.

[12] Eugenia D. Mehleri, Haralambos Sarimveis, Nikolaos C. Markatos, and Lazaros G. Papageorgiou. A mathematical programming approach for optimal design of distributed energy systems at the neighbourhood level. 44(1):96–104.

[13] Wenliang Li, Yuyu Zhou, Kristen Cetin, Jiyong Eom, Yu Wang, Gang Chen, and Xuesong Zhang. Modeling urban building energy use: A review of modeling approaches and procedures. 141:2445–2457.

[14] Miguel Chang, Jakob Zink Thellufsen, Behnam Zakeri, Bryn Pickering, Stefan Pfenninger, Henrik Lund, and Poul Alberg Østergaard. Trends in tools and approaches for modelling the energy transition. 290:116731.

[15] Francisca Jalil-Vega and Adam D. Hawkes. The effect of spatial resolution on outcomes from energy systems modelling of heat decarbonisation. 155:339–350.

[16] Vahid Aryanpur, Brian O'Gallachoir, Hancheng Dai, Wenying Chen, and James Glynn. A review of spatial resolution and regionalisation in national-scale energy systems optimisation models. 37:100702.

[17] Kody M. Powell, Akshay Sriprasad, Wesley J. Cole, and Thomas F. Edgar. Heating, cooling, and electrical load forecasting for a large-scale district energy system. 74:877–885.

[18] Leander Kotzur, Peter Markewitz, Martin Robinius, Gonçalo Cardoso, Peter Stenzel, Miguel Heleno, and Detlef Stolten. Bottom-up energy supply optimization of a national building stock. 209:109667.

[19] Ina De Jaeger, Glenn Reynders, Chadija Callebaut, and Dirk Saelens. A building clustering approach for urban energy simulations. 208:109671.

[20] P. Stadler, Luc Girardin, and F. Maréchal. The swiss potential of model predictive control for building energy systems.

[21] Mohammad Mohammadi, Younes Noorollahi, Behnam Mohammadi-ivatloo, Mehdi Hosseinzadeh, Hossein Yousefi, and Sasan Torabzadeh Khorasani. Optimal management of energy hubs and smart energy hubs – a review. 89:33–50.

[22] Azadeh Maroufmashat, Ali Elkamel, Michael Fowler, Sourena Sattari, Ramin Roshandel, Amir Hajimiragha, Sean Walker, and Evgueniy Entchev. Modeling and optimization of a network of energy hubs to improve economic and emission considerations. 93:2546–2558.

[23] M. Di Somma, B. Yan, N. Bianco, G. Graditi, P. B. Luh, L. Mongibello, and V. Naso. Operation optimization of a distributed energy system considering energy costs and exergy efficiency. 103:739–751.

[24] M. Di Somma, B. Yan, N. Bianco, G. Graditi, P. B. Luh, L. Mongibello, and V. Naso. Multi-objective design optimization of distributed energy systems through cost and exergy assessments. 204:1299–1316.

[25] Hongbo Ren, Weisheng Zhou, and Weijun Gao. Optimal option of distributed energy systems for building complexes in different climate zones in china. 91(1):156–165.

[26] David A. Copp, Tu A. Nguyen, Raymond H. Byrne, and Babu R. Chalamala. Optimal sizing of distributed energy resources for planning 100% renewable electric power systems. 239:122436.

[27] Jonas Allegrini, Kristina Orehounig, Georgios Mavromatidis, Florian Ruesch, Viktor Dorer, and Ralph Evins. A review of modelling approaches and tools for the simulation of district-scale energy systems. 52:1391–1404.

[28] Boran Morvaj, Ralph Evins, and Jan Carmeliet. Optimising urban energy systems: Simultaneous system sizing, operation and district heating network layout. 116:619–636.

[29] Thomas Schütz, Xiaolin Hu, Marcus Fuchs, and Dirk Müller. Optimal design of decentralized energy conversion systems for smart microgrids using decomposition methods. 156:250–263.

[30] Luise Middelhauve, Cedric Terrier, and Francois Marechal. Decomposition strategy for districts as renewable energy hubs. page 10.

[31] Paul Michael Stadler. Model-based sizing of building energy systems with renewable sources.

[32] Sven Eggimann, Michael Wagner, Yoo Na Ho, Mirjam Züger, Ute Schneider, and Kristina Orehounig. Geospatial simulation of urban neighbourhood densification potentials. 72:103068.

[33] CENTRE SCIENTIFIQUE ET TECHNIQUE DU BATIMENT. Geo-clustering to deploy the potential of energy efficient buildings across EU.

[34] Corentin Kuster, Jean-Laurent Hippolyte, Yacine Rezgui, and Monjur Mourshed. A simplified geo-cluster definition for energy system planning in europe. 158:3222–3227.

[35] Demographic and geographic region definition in energy system modelling. a case study of canada's path to net zero greenhouse gas emissions by 2050 and the role of hydrogen.

[36] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems. 214:219–238.

[37] Tianjie Liu, Wenling Jiao, and Xinghao Tian. A framework for uncertainty and sensitivity analysis of district energy systems considering different parameter types. 7:6908–6920.

[38] Wei Tian. A review of sensitivity analysis methods in building energy analysis. 20:411–419.

[39] Paul Westermann and Ralph Evins. Surrogate modelling for sustainable building design – a review. 198:170–186.

[40] Torben Østergård, Rasmus L. Jensen, and Steffen E. Maagaard. Early building design: Informed decision-making by exploring multidimensional design space using sensitivity analysis. 142:8–22.

[41] Hrishikesh D. Vinod. Integer programming and the theory of grouping. 64(326):506–519. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1969.10500990.

[42] S. Lloyd. Least squares quantization in PCM. 28(2):129–137. Conference Name: IEEE Transactions on Information Theory.

[43] Partitioning around medoids (program PAM). In *Finding Groups in Data*, pages 68–125. John Wiley & Sons, Ltd. Section: 2 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470316801.ch2.

[44] David Dohan, Stefani Karp, and Brian Matejek. K-median algorithms: Theory in practice. page 21.

[45] Amit Saxena, Mukesh Prasad, Akshansh Gupta, Neha Bharill, Om Prakash Patel, Aruna Tiwari, Meng Joo Er, Weiping Ding, and Chin-Teng Lin. A review of clustering techniques and developments. 267:664–681.

[46] P. Cheeseman and J. Stutz. Bayesian classification (AutoClass): Theory and results.

[47] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, pages 226–231. AAAI Press.

[48] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. Density-based clustering based on hierarchical density estimates. In Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda, and Guandong Xu, editors, *Advances in Knowledge Discovery and Data Mining*, Lecture Notes in Computer Science, pages 160–172. Springer.

[49] Clemens Felsmann, Luise Mann, and Vera Boß. Identification of urban cellular structures for flexible heat and temperature distribution in district heating networks. 7:9–17.

[50] Thomas Schütz, Markus Hans Schraven, Marcus Fuchs, Peter Remmen, and Dirk Müller. Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis. 129:570–582.

[51] Olatz Arbelaitz, Ibai Gurrutxaga, Javier Muguerza, Jesús M. Pérez, and Iñigo Perona. An extensive comparative study of cluster validity indices. 46(1):243–256.

[52] Davoud Moulavi, Pablo A. Jaskowiak, Ricardo J. G. B. Campello, Arthur Zimek, and Jörg Sander. Density-based clustering validation. In *Proceedings of the 2014 SIAM International Conference on Data Mining*, pages 839–847. Society for Industrial and Applied Mathematics.

[53] Bo Wang, Payman Dehghanian, and Dongbo Zhao. Chance-constrained energy management system for power grids with high proliferation of renewables and electric vehicles. 11(3):2324–2336.

[54] Alaa Alhamwi, Wided Medjroubi, Thomas Vogt, and Carsten Agert. Modelling urban energy requirements using open source data and models. 231:1100–1108.

[55] C. Doga Demirhan, William W. Tso, Joseph B. Powell, Clara F. Heuberger, and Efstratios N. Pistikopoulos. A multiscale energy systems engineering approach for renewable power generation and storage optimization. 59(16):7706–7721.

[56] C.A.G. MacRae, A.T. Ernst, and M. Ozlen. A benders decomposition approach to transmission expansion planning considering energy storage. 112:795–803.

[57] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis. The Primer*. Wiley. Edition: 1.

[58] Max D. Morris. Factorial sampling plans for preliminary computational experiments. 33(2):161–174. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/00401706.1991.10484804.

[59] Andrea Saltelli. Making best use of model evaluations to compute sensitivity indices. 145(2):280–297.

[60] J {and} Saltelli A Campolongo, F {and} Cariboni. Sensitivity analysis: the morris method versus the variance based measures.

[61] Will Usher, Jon Herman, Calvin Whealton, David Hadka, Xantares, Fernando Rios, Bernardoct, Chris Mutel, and Joeri Van Engelen. Salib/salib: Launch!

[62] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. 181(2):259–270.

[63] *Application of levelized infrastructure-connected regionalisation in energy systems modelling*.

[64] Luise Middelhauve, Francesco Baldi, Paul Stadler, and François Maréchal. Grid-aware layout of photovoltaic panels in sustainable building energy systems. 8:573290.

[65] Luc Girardin. A GIS-based methodology for the evaluation of integrated energy systems in urban area.

[66] Jonas Schnidrig, Tuong-Van Nguyen, and Xiang Li. A modelling framework for assessing the impact of green mobility technologies on energy systems. In Francois Marechal, editor, *Proceedings of ECOS 2021*, page 13. ECOS 2021 Local Organizing Committee. ZSCC: 0000000.

[67] Jonas Schnidrig, Rachid Cherkaoui, Yasmine Calisesi, and Manuele Margni. On the role of energy infrastructure in the energy transition case study of an independent and CO2 neutral energy system for switzerland.

[68] Jonas Schnidrig. *Energyscope - Building energy systems integration*.

[69] Joseph Loustau. *Clustering and typification of urban districts for energy systems modelling*.

[70] Rahul Gupta, Fabrizio Sossan, and Mario Paolone. Countrywide PV hosting capacity and energy storage requirements for distribution networks: The case of switzerland. 281:116010.

[71] Stefano Moret. Strategic energy planning under uncertainty.

[72] Office fédéral de l'énergie OFEN. Statistique globale de l'énergie.

[73] DETEC. Analyse des schweizerischen nergieverbrauchs 2000–2021 nach verwendungszwecken.

[74] SIA. vernehmlassung sia 380/1 heizwärmebedarf.

[75] Mashael Yazdanie, Martin Densing, and Alexander Wokaun. The nationwide characterization and modeling of local energy systems: Quantifying the role of decentralized generation and energy resources in future communities. 118:516–533.

[76] L. V. D. Maaten and Geoffrey E. Hinton. Visualizing data using t-SNE.

[77] Luise Middelhauve, Cédric Terrier, and François Maréchal. Decomposition strategy for districts as renewable energy hubs. 9:287–297. Conference Name: IEEE Open Access Journal of Power and Energy.

[78] scikit-learn-contrib/hdbscan. original-date: 2015-04-22T13:32:37Z.

Appendix A

Global sensitivity analysis

A.1 Morris screening

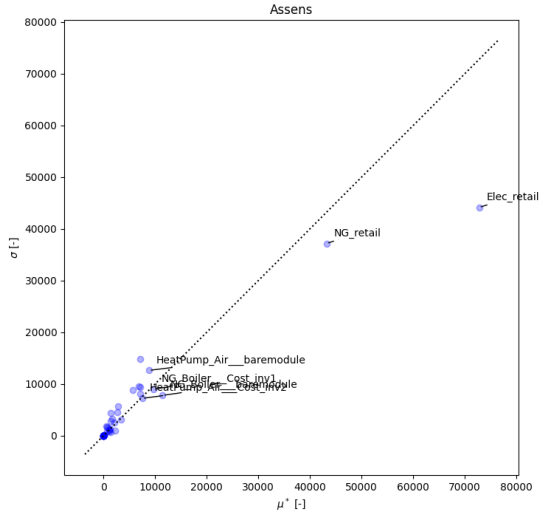
A.1.1 Input parameters

Parameter	Default	Min	Max
Elec_retail	0.2100	0.10500	0.4200
Elec_feedin	0.0800	0.06000	0.1000
NG_retail	0.0900	0.04500	0.1800
NG_feedin	0.0001	0.00005	0.0002
PVA_efficiency_ref	0.2000	0.10000	0.4000
PV___Cost_inv2	2000.0000	1000.00000	4000.0000
NG_Boiler___Units_Fmin	0.1000	0.05000	0.2000
NG_Boiler___Units_Fmax	100000.0000	50000.00000	200000.0000
NG_Boiler___Cost_inv1	3800.0000	1900.00000	7600.0000
NG_Boiler___Cost_inv2	100.0000	50.00000	200.0000
NG_Boiler___baremodule	1.8000	0.90000	3.6000
NG_Boiler___lifetime	20.0000	10.00000	40.0000
HeatPump_Air___Units_Fmin	0.1000	0.05000	0.2000
HeatPump_Air___Units_Fmax	100000.0000	50000.00000	200000.0000
HeatPump_Air___Cost_inv1	4000.0000	2000.00000	8000.0000
HeatPump_Air___Cost_inv2	1000.0000	500.00000	2000.0000
HeatPump_Air___baremodule	1.8000	0.90000	3.6000
HeatPump_Air___lifetime	20.0000	10.00000	40.0000
PV___Units_Fmax	100000.0000	50000.00000	200000.0000
PV___Cost_inv1	5000.0000	2500.00000	10000.0000
PV___baremodule	1.0000	0.50000	2.0000
PV___lifetime	20.0000	10.00000	40.0000
WaterTankSH___Units_Fmin	0.1000	0.05000	0.2000
WaterTankSH___Units_Fmax	1000.0000	500.00000	2000.0000
WaterTankSH___Cost_inv1	700.0000	350.00000	1400.0000
WaterTankSH___Cost_inv2	2000.0000	1000.00000	4000.0000
WaterTankSH___baremodule	1.8000	0.90000	3.6000
WaterTankSH___lifetime	20.0000	10.00000	40.0000

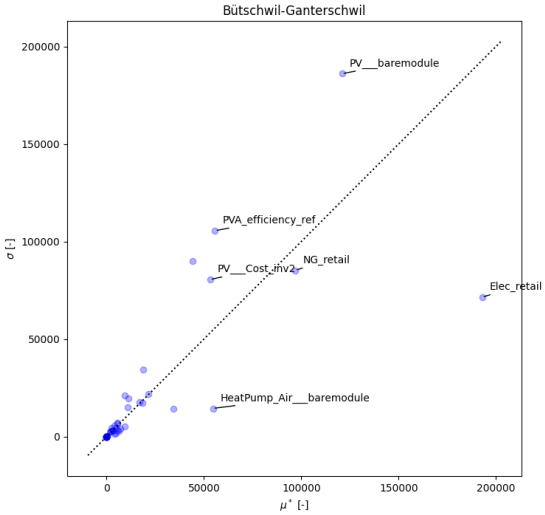
WaterTankDHW___Units_Fmin	0.0001	0.00005	0.0002
WaterTankDHW___Units_Fmax	1000.0000	500.00000	2000.0000
WaterTankDHW___Cost_inv1	300.0000	150.00000	600.0000
WaterTankDHW___Cost_inv2	4000.0000	2000.00000	8000.0000
WaterTankDHW___baremodule	1.8000	0.90000	3.6000
WaterTankDHW___lifetime	20.0000	10.00000	40.0000
ElectricalHeater_SH___Units_Fmin	0.1000	0.05000	0.2000
ElectricalHeater_SH___Units_Fmax	100000.0000	50000.00000	200000.0000
ElectricalHeater_SH___Cost_inv1	1000.0000	500.00000	2000.0000
ElectricalHeater_SH___Cost_inv2	50.0000	25.00000	100.0000
ElectricalHeater_SH___baremodule	1.0000	0.50000	2.0000
ElectricalHeater_SH___lifetime	20.0000	10.00000	40.0000
ElectricalHeater_DHW___Units_Fmin	0.1000	0.05000	0.2000
ElectricalHeater_DHW___Units_Fmax	100000.0000	50000.00000	200000.0000
ElectricalHeater_DHW___Cost_inv1	1000.0000	500.00000	2000.0000
ElectricalHeater_DHW___Cost_inv2	50.0000	25.00000	100.0000
ElectricalHeater_DHW___baremodule	1.8000	0.90000	3.6000
ElectricalHeater_DHW___lifetime	20.0000	10.00000	40.0000

Table A.1: Morris input parameters and variation range

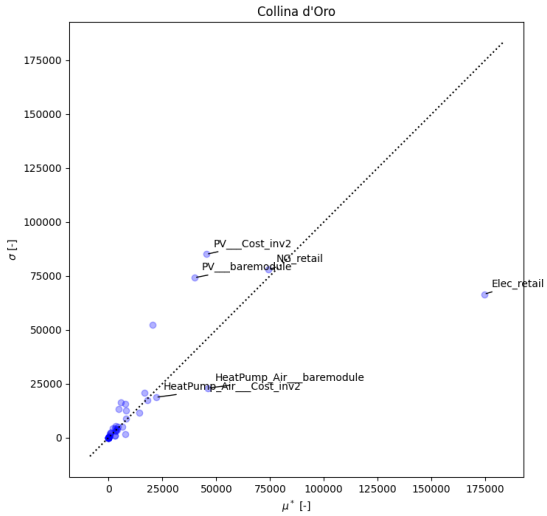
A.1.2 Elementary effects results



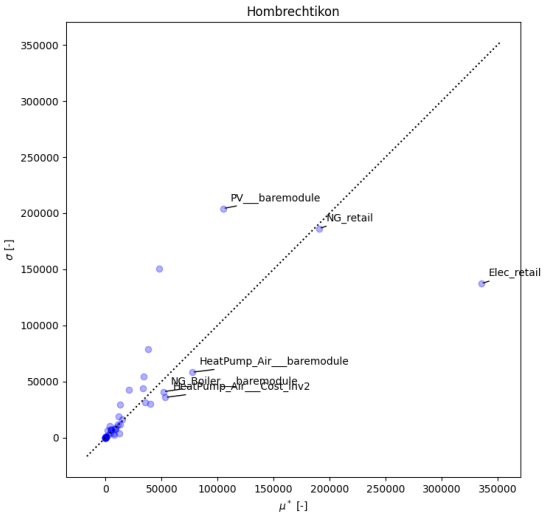
Assens



Bütschwil-Ganterschwil

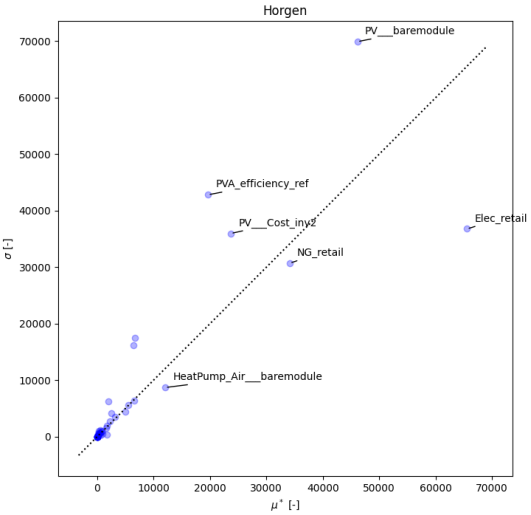


Collina d'Oro

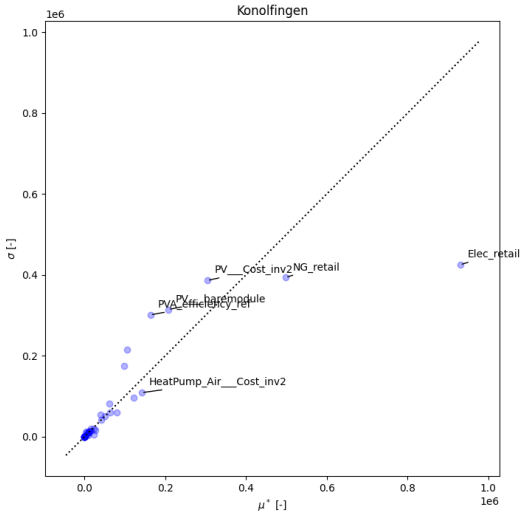


Hombrechtikon

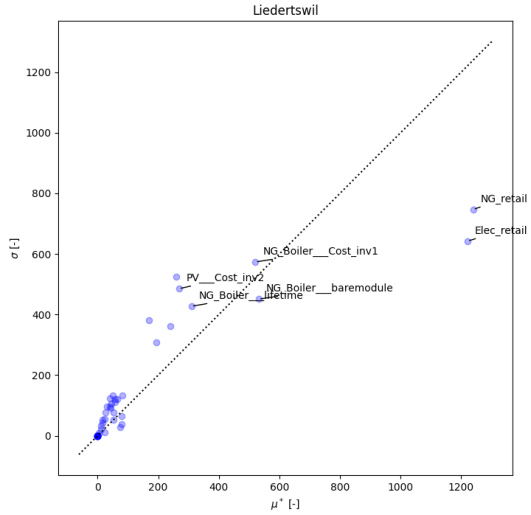
A.2 Sobol sampling



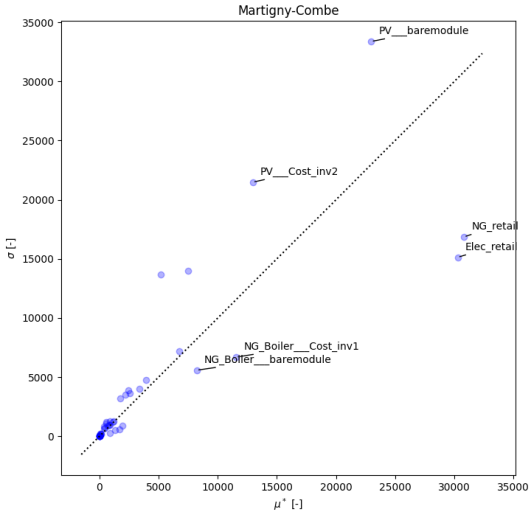
Horgen



Konolfingen



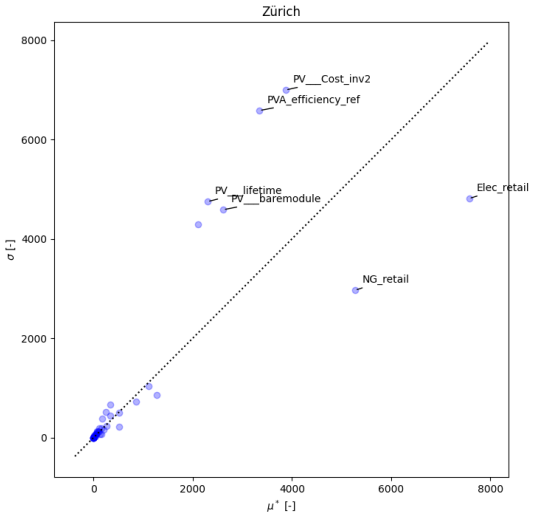
Liedertswil



Martigny-Combe

Parameters	Upper bound	Lower bound	Unit
Electricity retail tariff	0.105	0.42	[CHF/kWh]
NG retail tariff	0.045	0.18	[CHF/kWh]

Table A.2: Sampling parameters for the Sobol sampling



Zürich

Figure A.1: Morris results represented on the μ^* - σ plane for each district

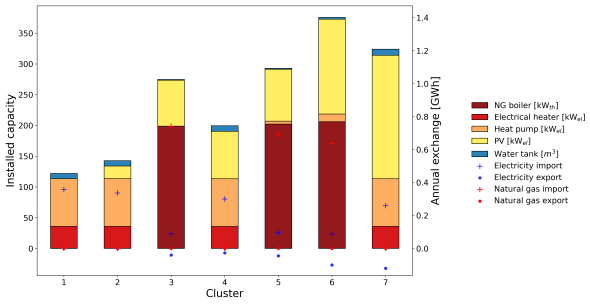
Table A.3: Clustering hyperparameter ϵ for the different districts

District name	ϵ	District size
Konolfingen	0.4	104
Hombrechtikon	0.3	94
Collina d'Oro	0.3	45
Bütschwil-Ganterschwil	0.3	43
Crans-Montana	0.3	31
Assens	0.3	16
Martigny-Combe	0.5	16
Horgen	0.5	10
Zürich	0.5	2
Liedertswil	0.5	1

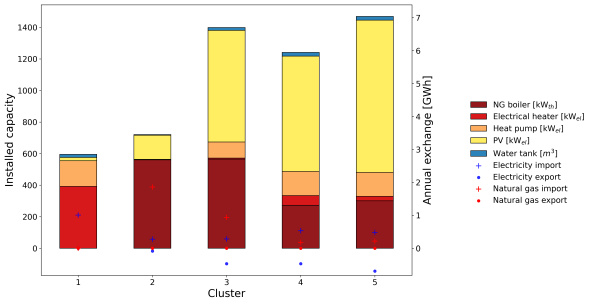
A.3 Typical configurations identification

A.3.1 Clustering hyperparameter ϵ

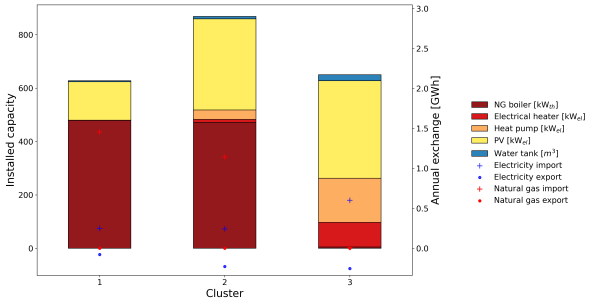
A.3.2 Districts configurations



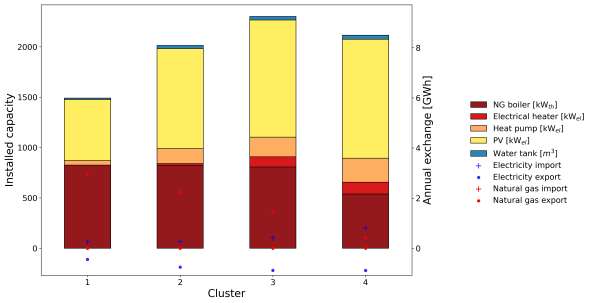
Assens



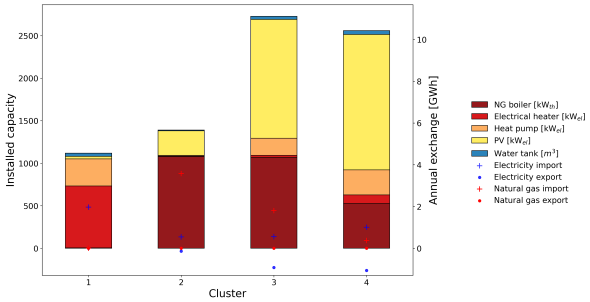
Bütschwil-Ganterschwil



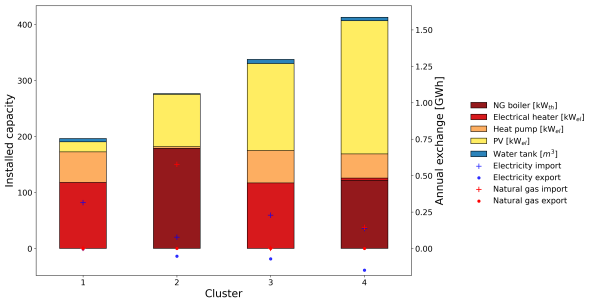
Collina d'Oro



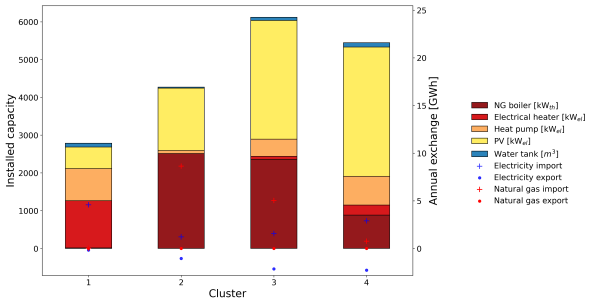
Crans-Montana



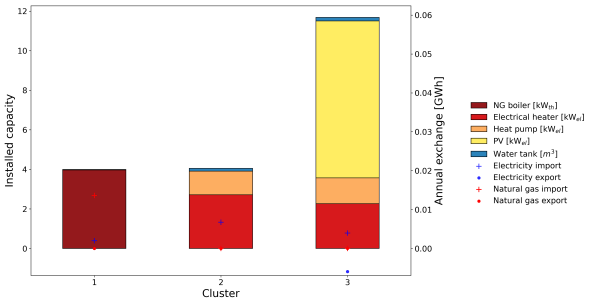
Hombrechtikon



Horgen

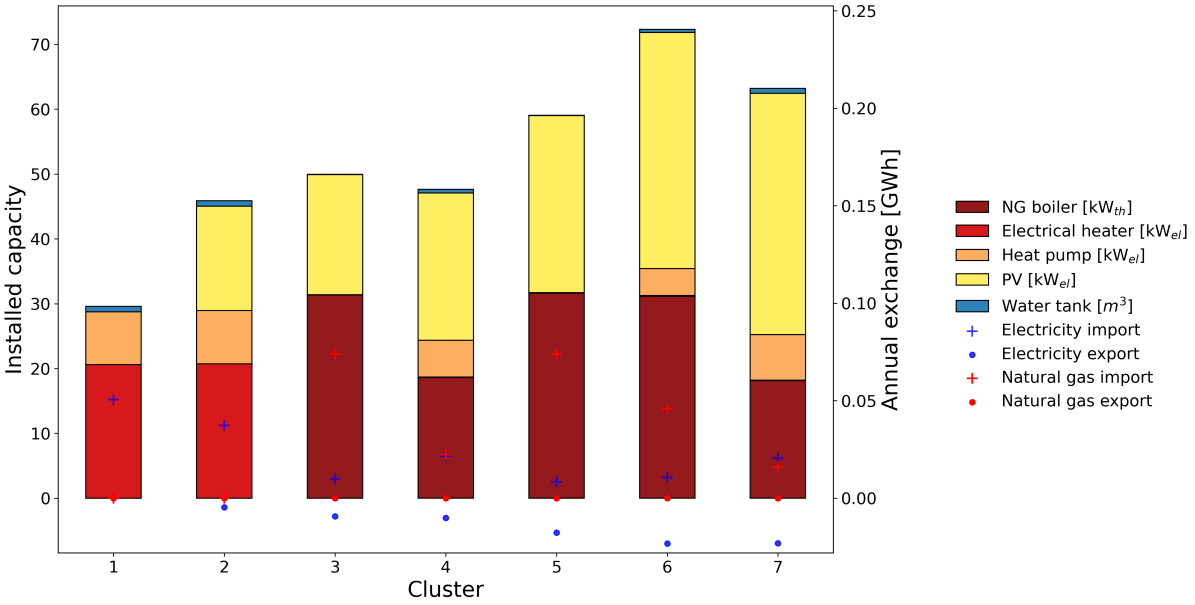


Konolfingen



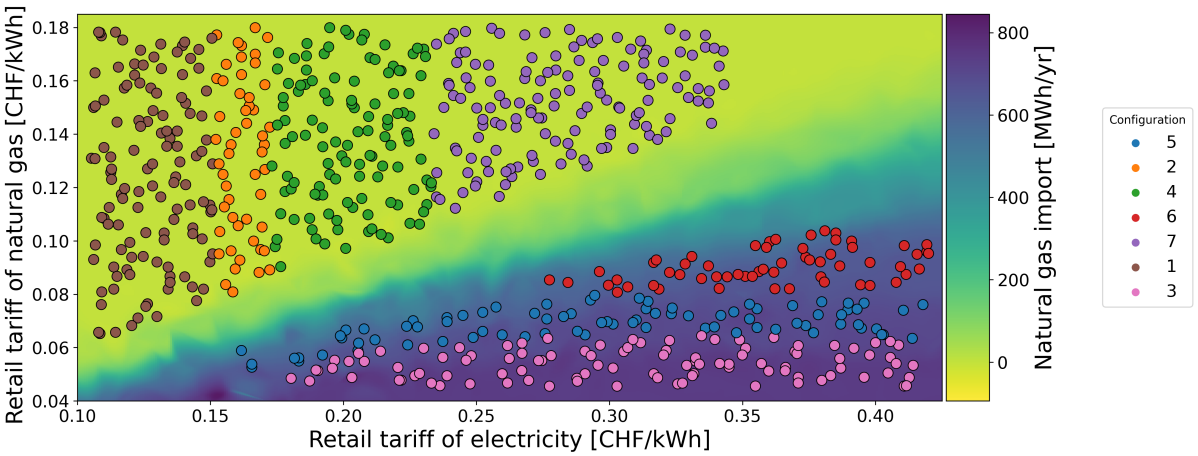
Liedertswil

A.3.3 Distribution of the configurations over the sampling space

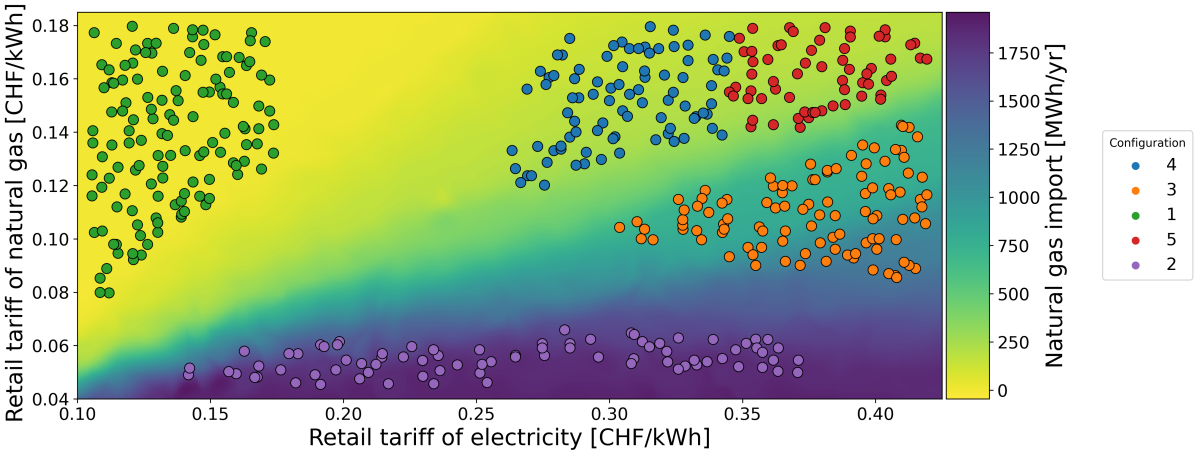


Zürich

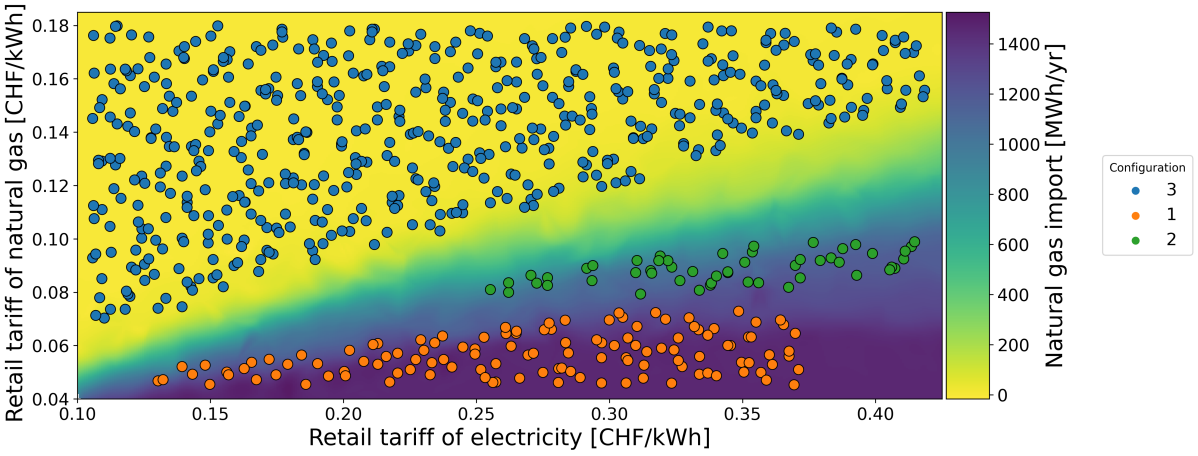
Figure A.2: Typical configurations.



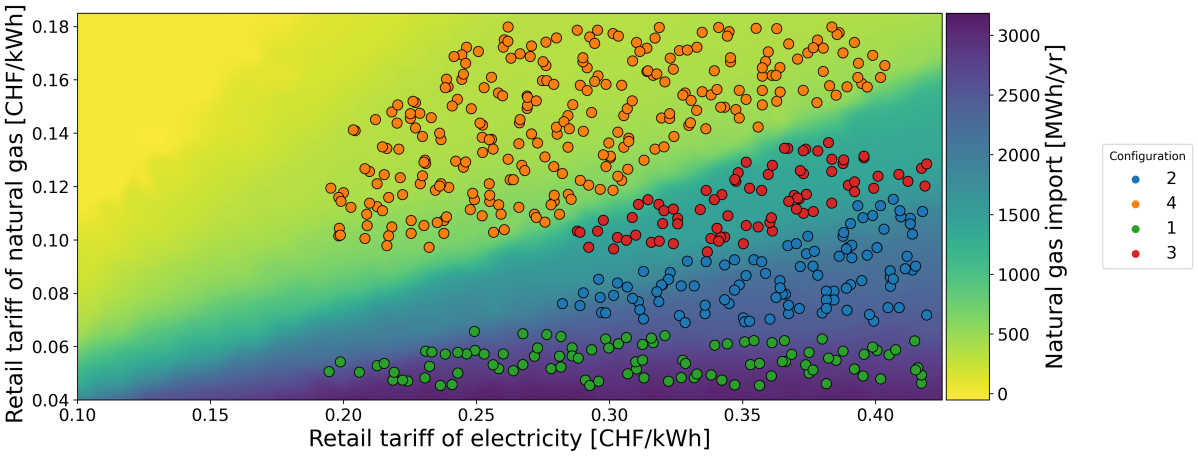
Assens



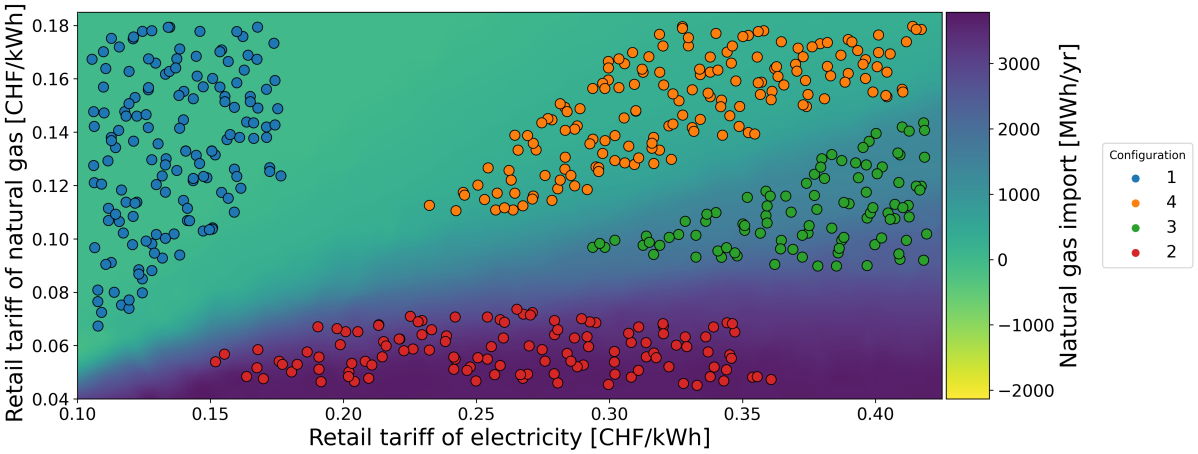
Bütschwil-Ganterschwil



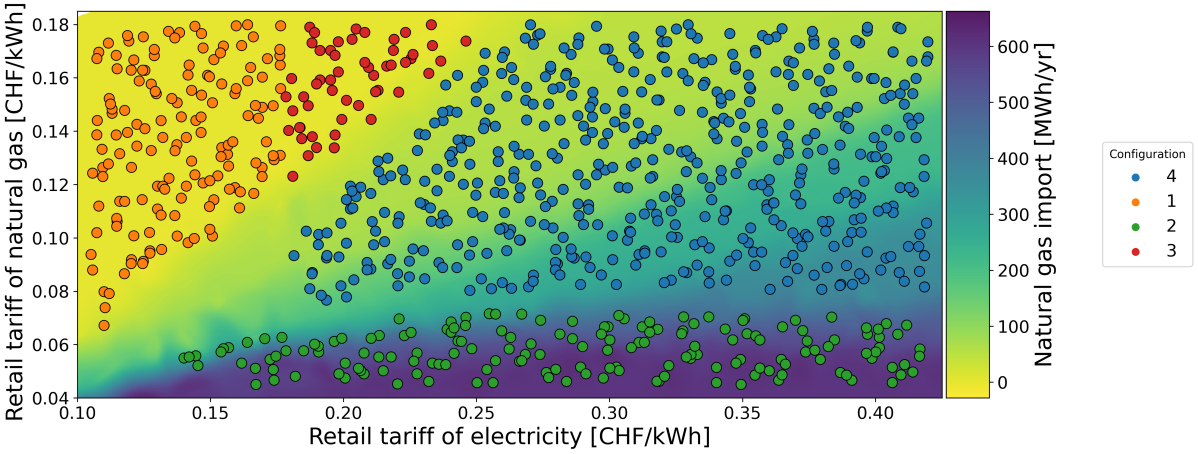
Collina d'Oro



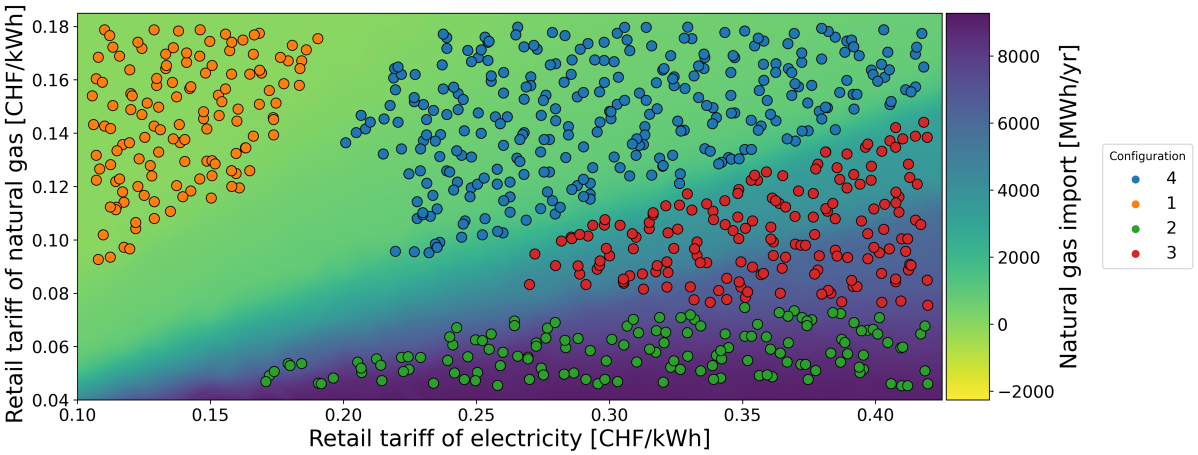
Crans-Montana



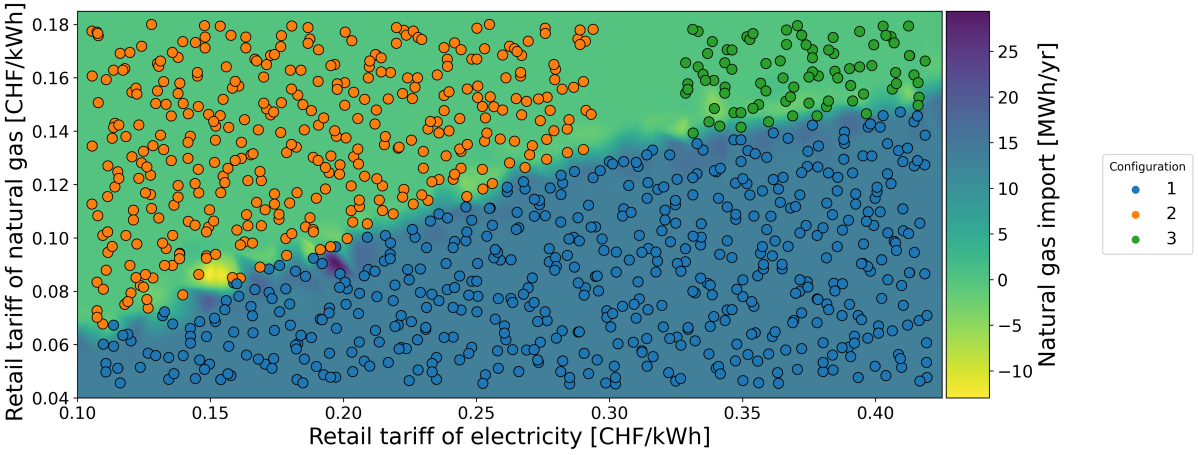
Hombrechtikon



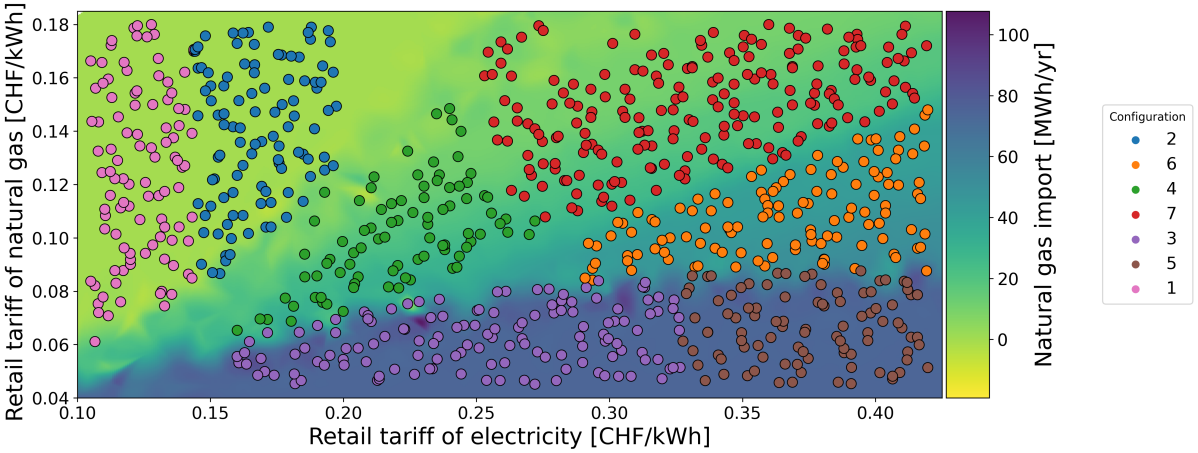
Horgen



Konolfingen



Liedertswil



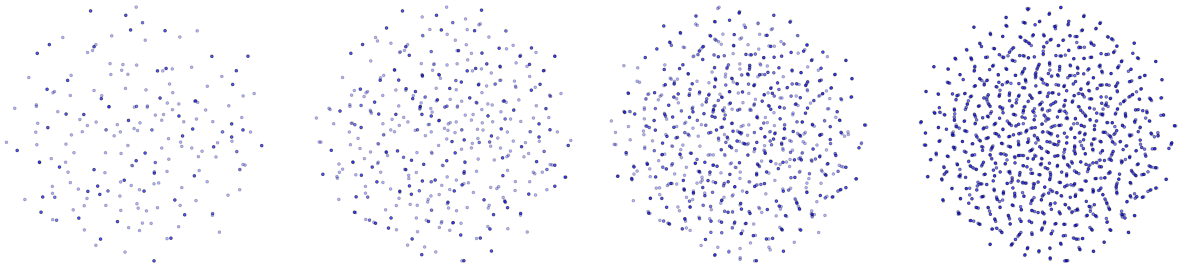
Zürich

Figure A.3: Typical configurations distribution in the retail tariffs space.

Appendix B

Effect of the decentralized method and a reduced number of trajectory on the clustering

In order to validate or not the use of the decentralized method within REHO a Sobol sampling has been applied on the retail tariffs of NG, electricity and on the feed-in tariff of electricity. Using 512 trajectories for the sampling, the space exploration required 2560 optimizations of the model, for both method, decentralized and decomposed. The sampling scheme of the Sobol sequence allows to select only a the i-ith first results to simulate a reduced number of trajectories. This behaviour is validated by Figure B.1 which illustrates the sampling space reduced to two dimensions, using a t-distributed stochastic neighbor embedding (T-SNE) algorithm [76], for a smaller number of trajectories. The sampling density is slightly decreased by reducing the number of sampling points by a factor of 2. This is a direct consequence of the robustness of Sobol sampling sequence.

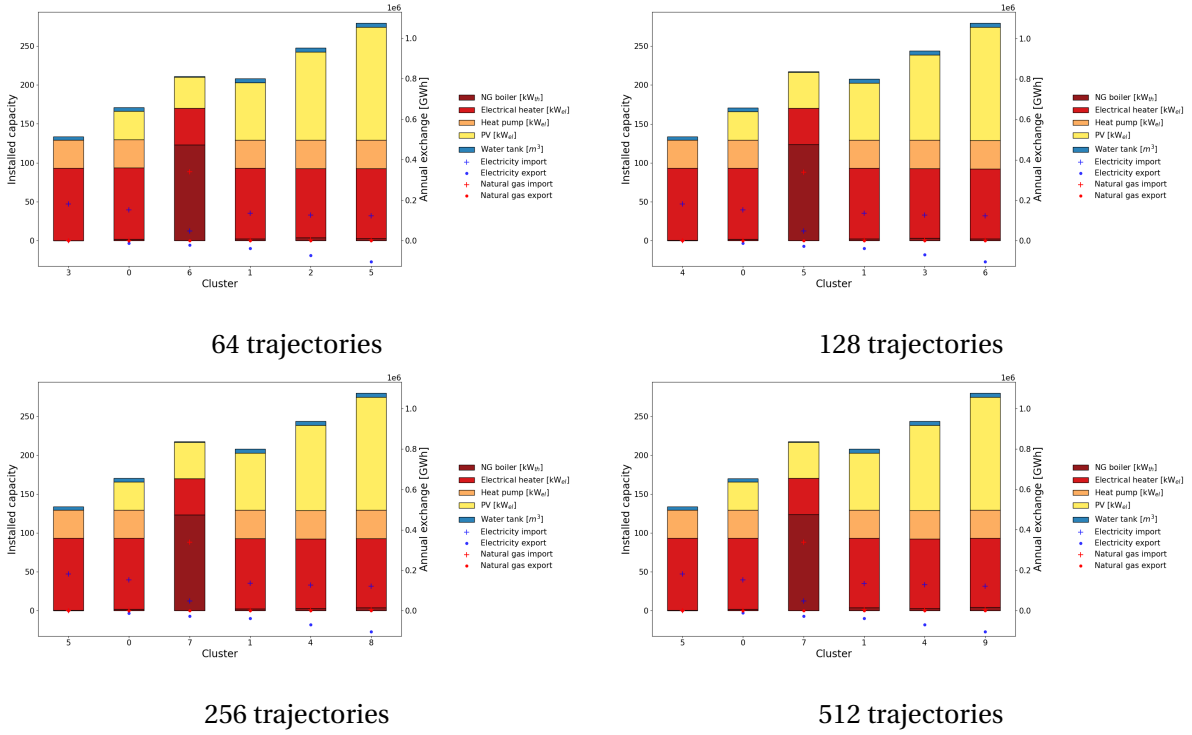


64 trajectories 128 trajectories 256 trajectories 512 trajectories

Figure B.1: Sampling space reduced to 2D for various number of trajectory

Even though, a reduced number of trajectories offers a good coverage of the sampling space, it is necessary to assess its impact on the clustering process. The following cluster configurations originates from a sensitivity analysis performed on a 10 buildings district optimized using the Dantzig-Wolfe decomposition (DWD). Figure B.2 represents four different clustering results with different number

of sampling points. The effect of a reduced sampling is almost non visible on the district configurations validating its usage for further computations.



64 trajectories

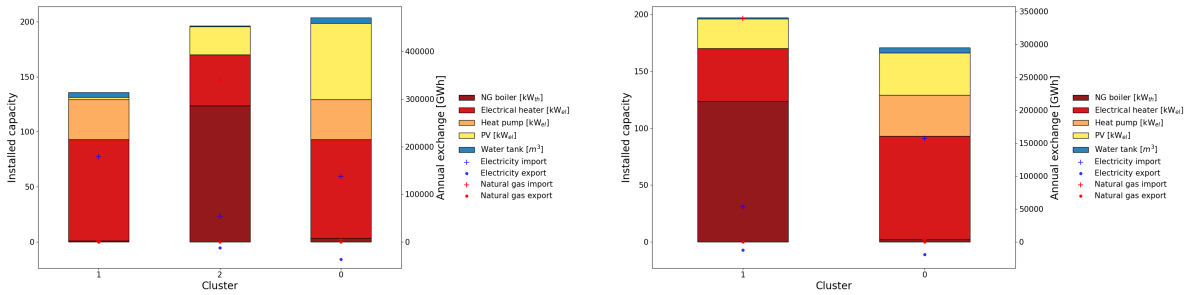
128 trajectories

256 trajectories

512 trajectories

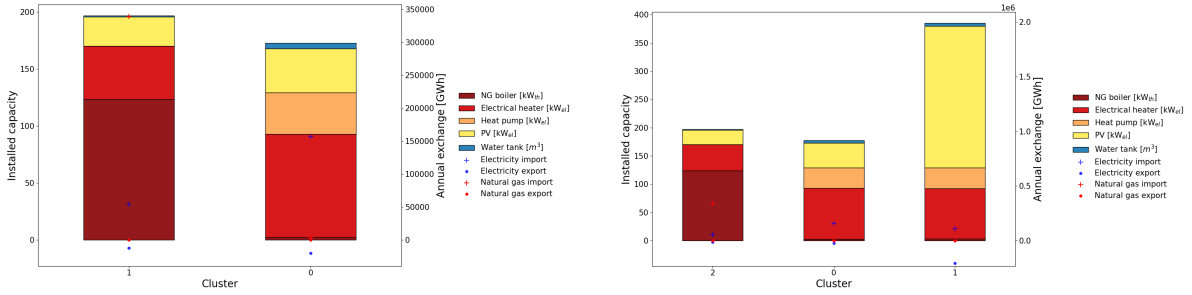
Figure B.2: Cluster configurations for reduced number of sampling points, ie. number of trajectory, obtained with the DWD method.

The next point to analyze is the impact of the decentralized method on the clustering. The decentralized method is an abuse of language as the decomposed method is also decentralized. The difference resides in the number of master optimization which is performed, the DWD will perform around 5 to 9 iterations [77] where the decentralized will only do a single iteration. This method allows to approach rapidly the optimal solution about 0.005% error of the objective function for district of a dozen buildings. Figure B.3 illustrates again the results of clustering for various number of sampling points but the optimizations has been performed using the decentralized method. Again, there is no spectacular difference between the number of chosen trajectory. However, there is a reduction in the configurations diversity, ie. there is less typical configuration identified. This loss might come from the final iterations of the DWD which allow to fine tune the system and reduce the TOTEX by adjusting the units size and operation. This behaviour is not a major problem for the future steps as the aggregation method still identifies typical configurations based on different energy carrier. However, the diversity of the configurations is reduced which might force the large-scale model to combine several configurations to obtain ideal profiles.



64 trajectories

128 trajectories



256 trajectories

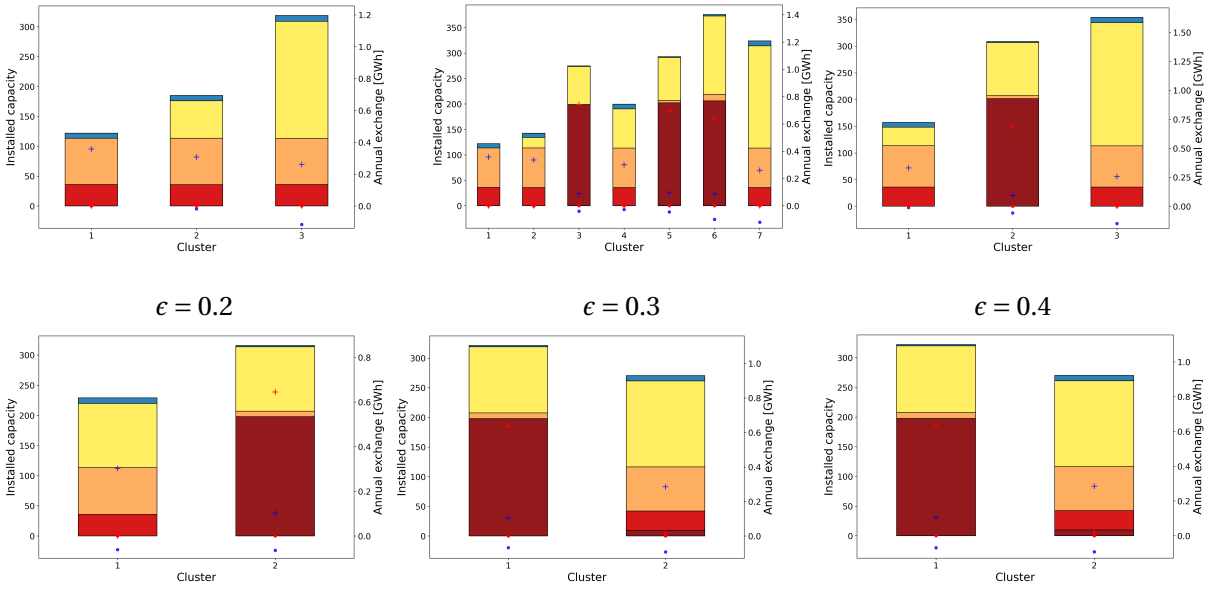
512 trajectories

Figure B.3: Cluster configurations for reduced number of sampling points, ie. number of trajectory, obtained with the decentralized method.

Appendix C

Clustering analysis

Clustering algorithms are powerful tools that can help to identify patterns and groupings within large datasets. While these algorithms are often easy to use, it is important to have a good understanding of their underlying principles and to fine-tune their hyperparameters in order to obtain accurate and useful results. The DBSCAN algorithm has two main parameters: $minPts$ and ϵ. The first one represents the number of points that should be within a ϵ distance to be considered a core points but it is not considered as the most impactful parameter, for more details see [78]. The second one, as mentioned, is a distanced that define, in some sort, the density of the cluster. The smaller the value is, the denser the cluster needs to be or, new clusters are identified. Therefore, the impact of the ϵ value on the configurations identification is analysed. Even if the size of the district should not have an impact on the clustering behaviour of the algorithm as the data is standardized, the analysis is performed on the district *Konolfingen* and *Assens*, which are 104 and 16 buildings respectively.

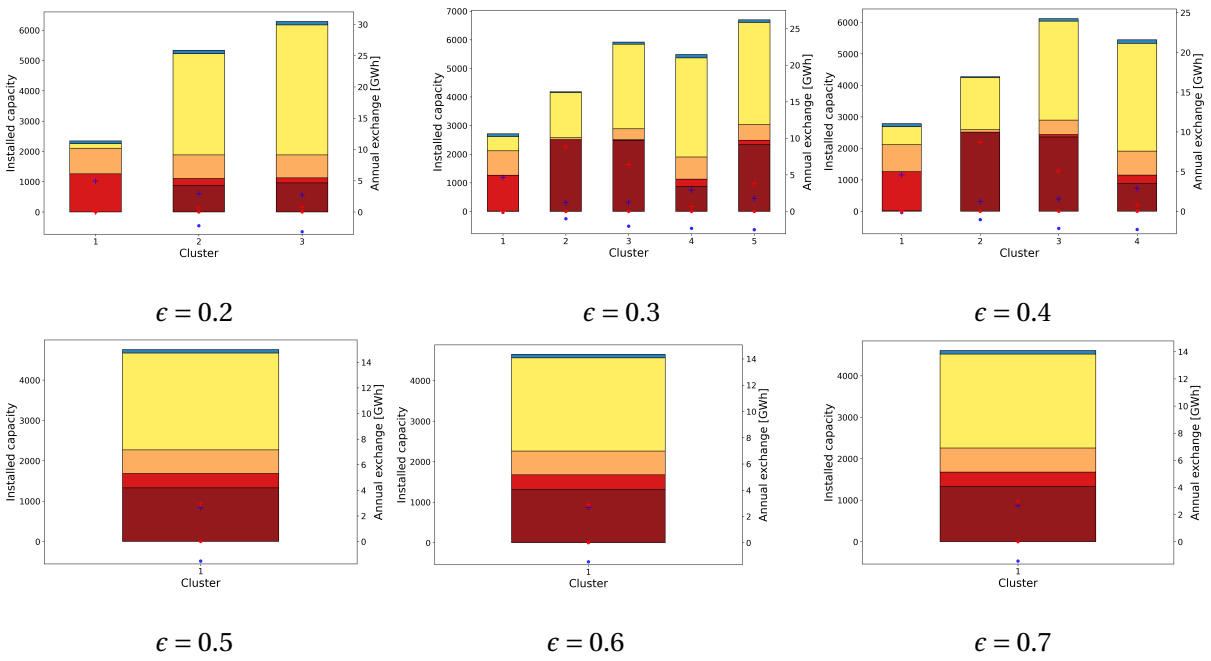


$\epsilon = 0.2$ $\epsilon = 0.3$ $\epsilon = 0.4$

$\epsilon = 0.5$ $\epsilon = 0.6$ $\epsilon = 0.7$

Figure C.1: Identified configurations with various ϵ for the district *Assens*.

Figure C.1 and Figure C.2 clearly show an optimal value of $\epsilon = 0.3$. Therefore, the districts that identified a limited number of configurations with the default value of $\epsilon = 0.5$ were clustered again with the optimal value. However, some districts didn't produced better results with various ϵ tested.



$\epsilon = 0.2$ $\epsilon = 0.3$ $\epsilon = 0.4$

$\epsilon = 0.5$ $\epsilon = 0.6$ $\epsilon = 0.7$

Figure C.2: Identified configurations with various ϵ for the district *Konolfingen*.

Appendix D

EnergyScope 2020 sectors demand

Table D.1: Swiss EnergyScope 2020 sector values

Sector	EndUses	Value
HOUSEHOLDS	ELECTRICITY_LV	0
HOUSEHOLDS	ELECTRICITY_MV	0
HOUSEHOLDS	ELECTRICITY_HV	0
HOUSEHOLDS	ELECTRICITY_EHV	0
SERVICES	ELECTRICITY_LV	15143.33
SERVICES	ELECTRICITY_MV	2327.528
SERVICES	ELECTRICITY_HV	0
SERVICES	ELECTRICITY_EHV	0
INDUSTRY	ELECTRICITY_LV	0
INDUSTRY	ELECTRICITY_MV	4807.66
INDUSTRY	ELECTRICITY_HV	8106.22
INDUSTRY	ELECTRICITY_EHV	1599.14
HOUSEHOLDS	HEAT_LOW_T_SH	0
HOUSEHOLDS	HEAT_LOW_T_HW	0
SERVICES	HEAT_LOW_T_SH	14095
SERVICES	HEAT_LOW_T_HW	2443
SERVICES	HEAT_HIGH_T	425
INDUSTRY	HEAT_LOW_T_SH	3482
INDUSTRY	HEAT_LOW_T_HW	546
INDUSTRY	HEAT_HIGH_T	24057
TRANSPORTATION	MOBILITY_PASSENGER	127555
TRANSPORTATION	MOBILITY_FREIGHT	27930

Appendix E

EnergyScope results

E.1 Additional results regarding the model validation

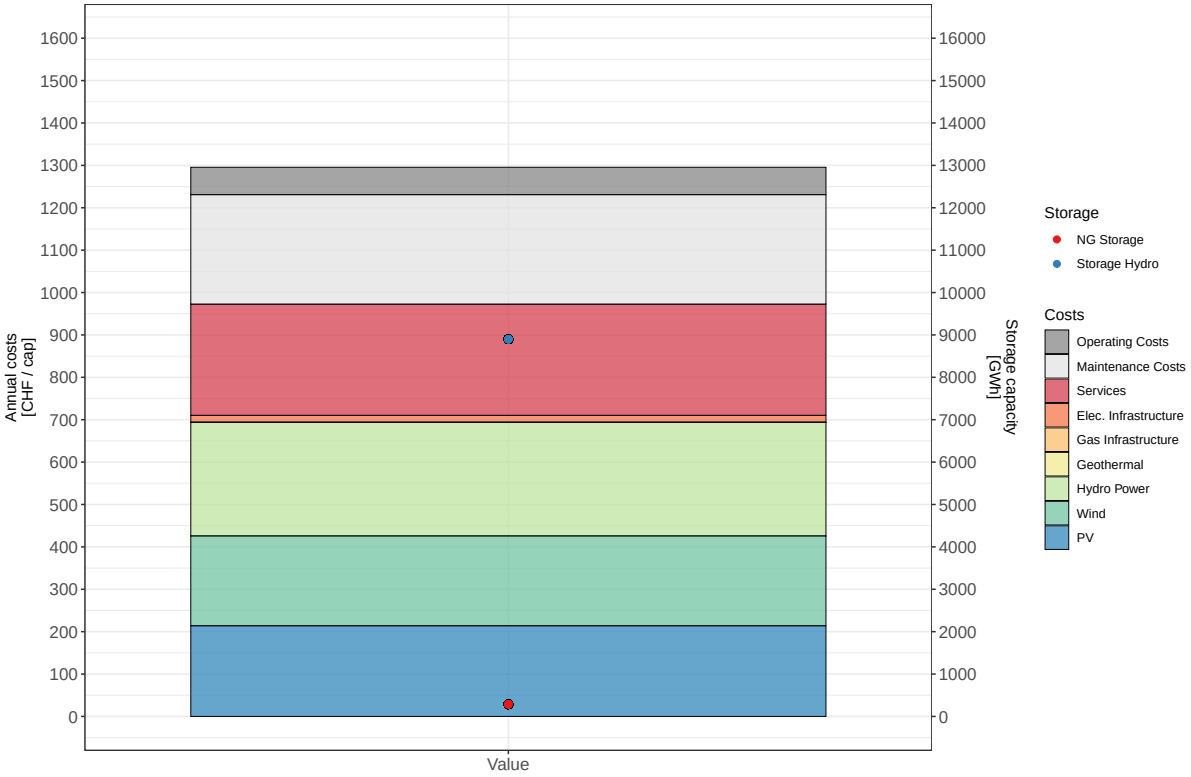


Figure E.1: SES 2020

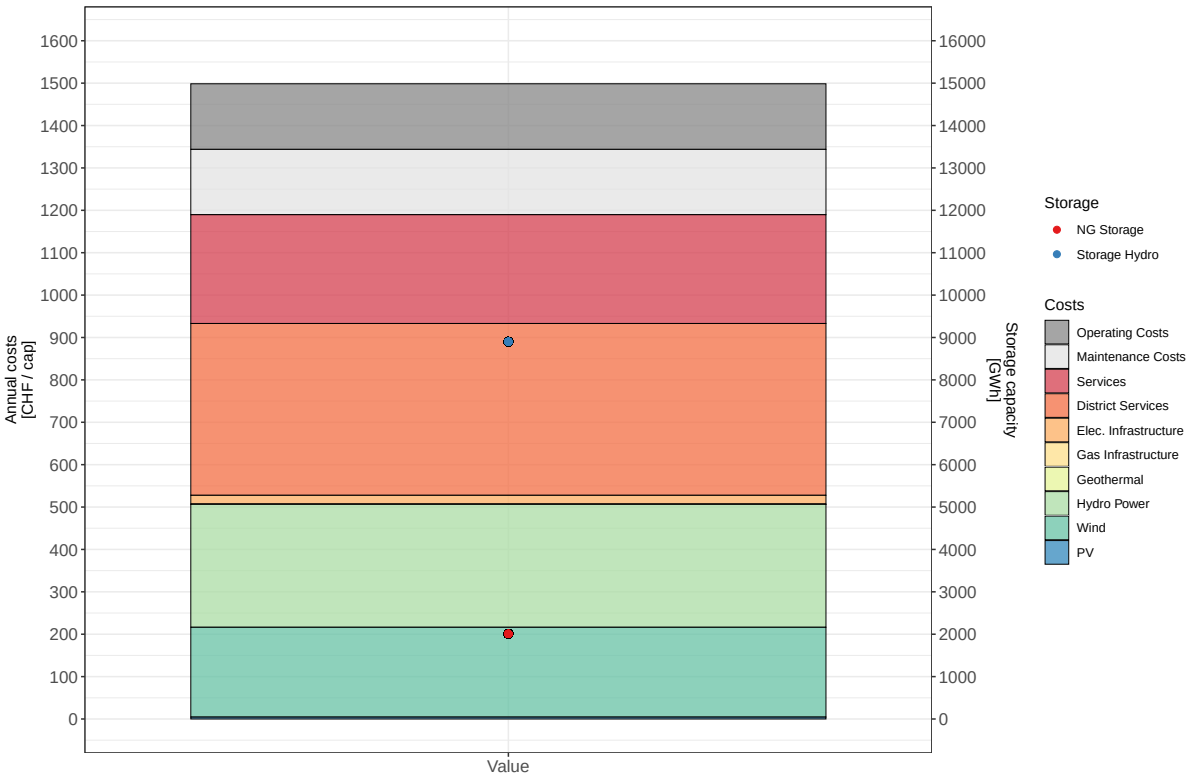


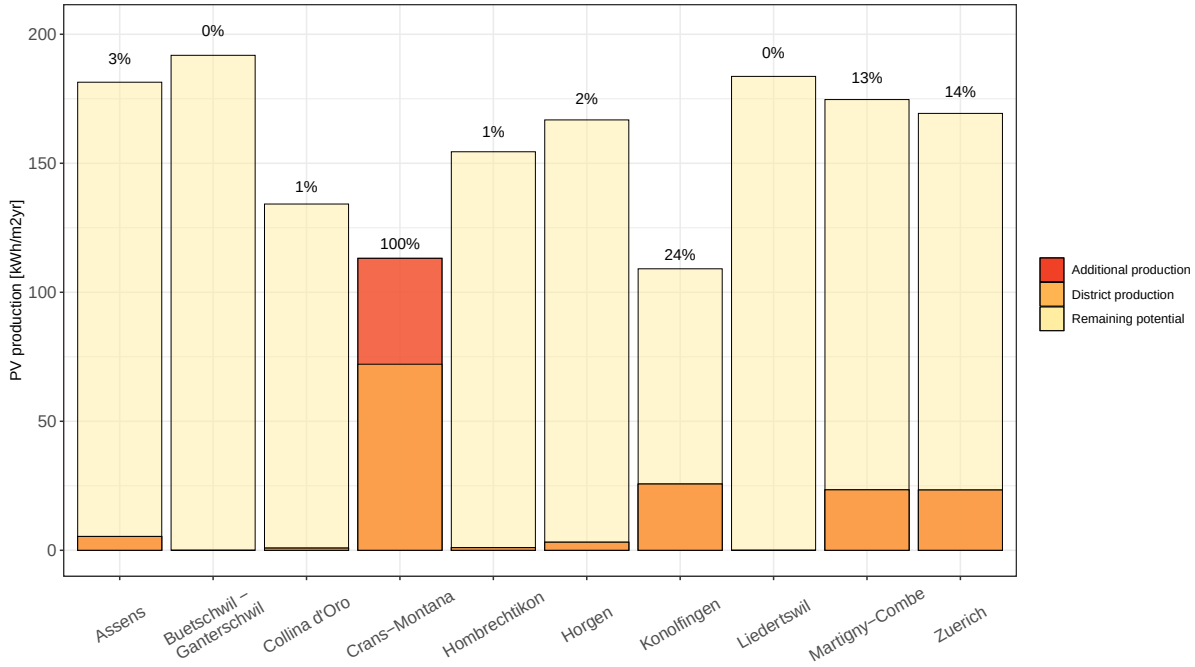
Figure E.2: SREHO2ES 2020

E.1.1 PV penetration in the built environment

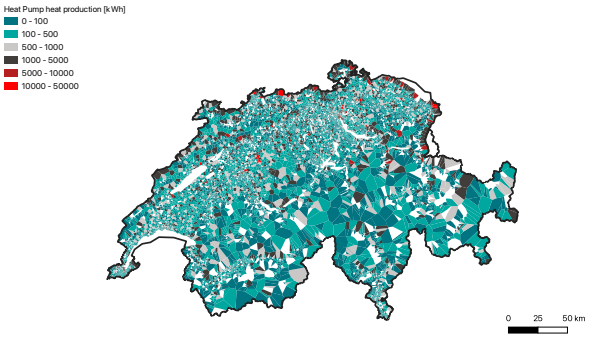
Table E.1: Identified bidding tariffs within ES.

District	PV feed-in tariff [CHF/kWh]
Assens	0.1274339
Bütschwil-Ganterschwil	0.1254299
Collina d'Oro	0.1228274
Crans-Montana	0.1074611
Hombrechtikon	0.1271358
Horgen	0.1241702
Konolfingen	0.1211480
Liedertswil	0.1402705
Martigny-Combe	0.1239286
Zürich	0.1246590

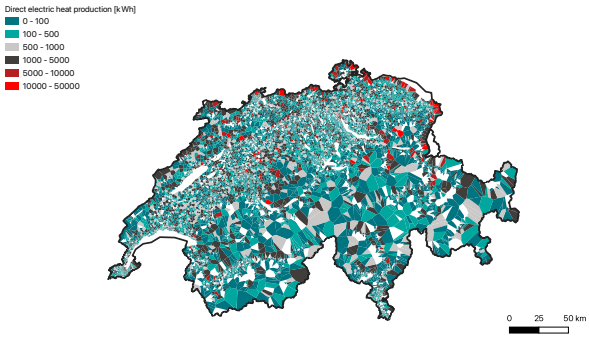
Figure E.3: Illustration of the installed PV capacity within the districts and their remaining potentials



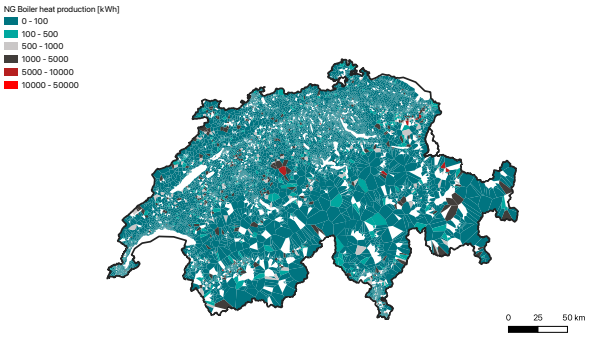
E.1.2 Distribution of the district technologies in Switzerland



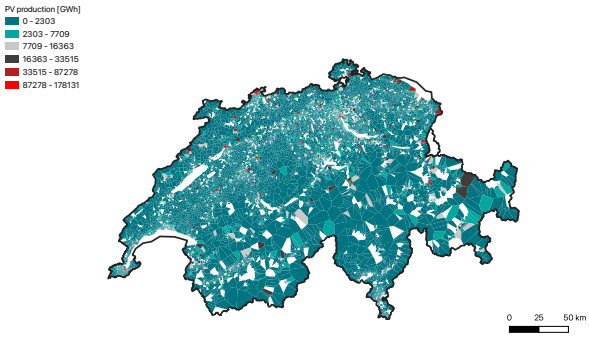
Heat pump



Electrical heater



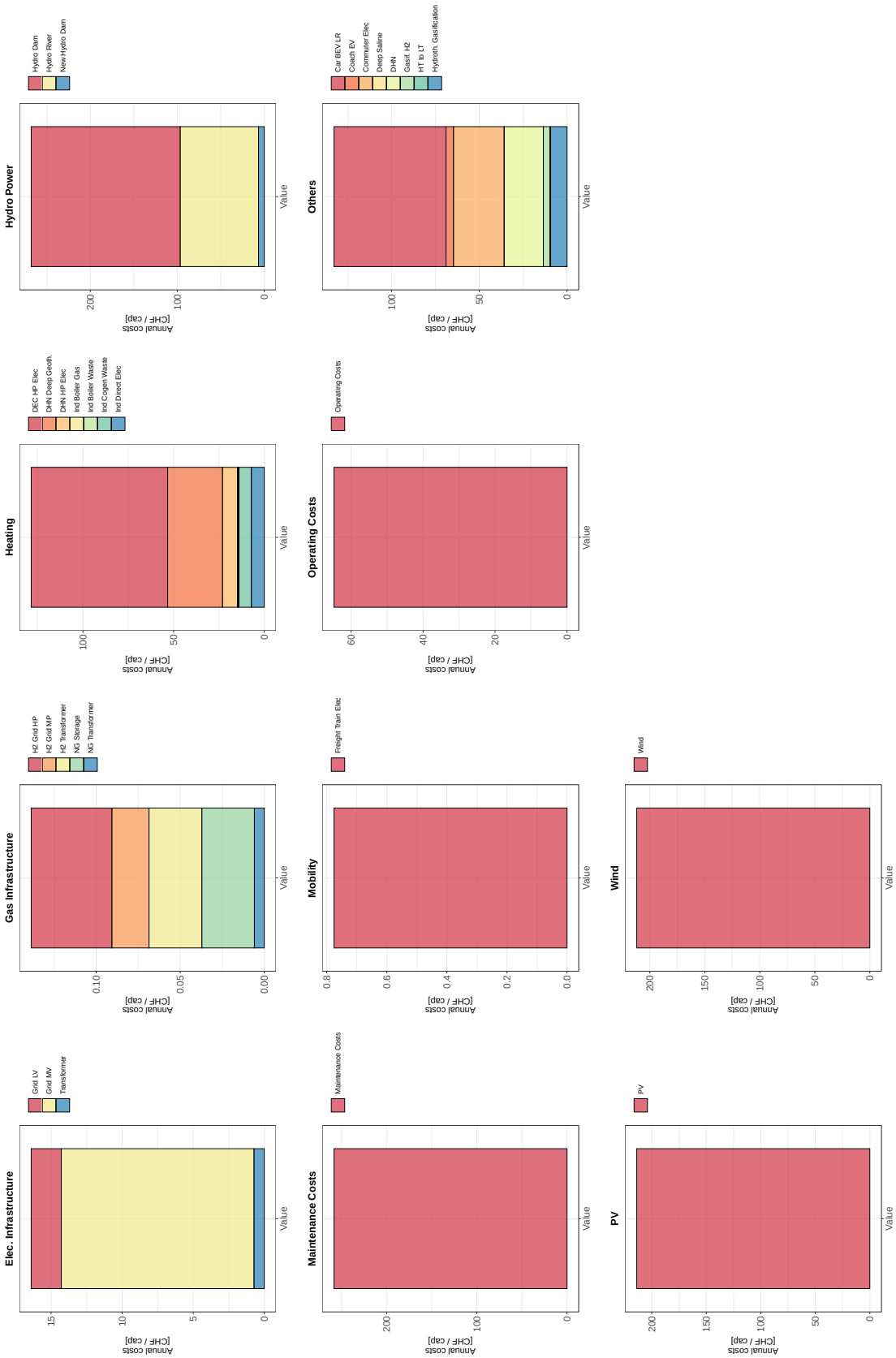
Natural gas boiler



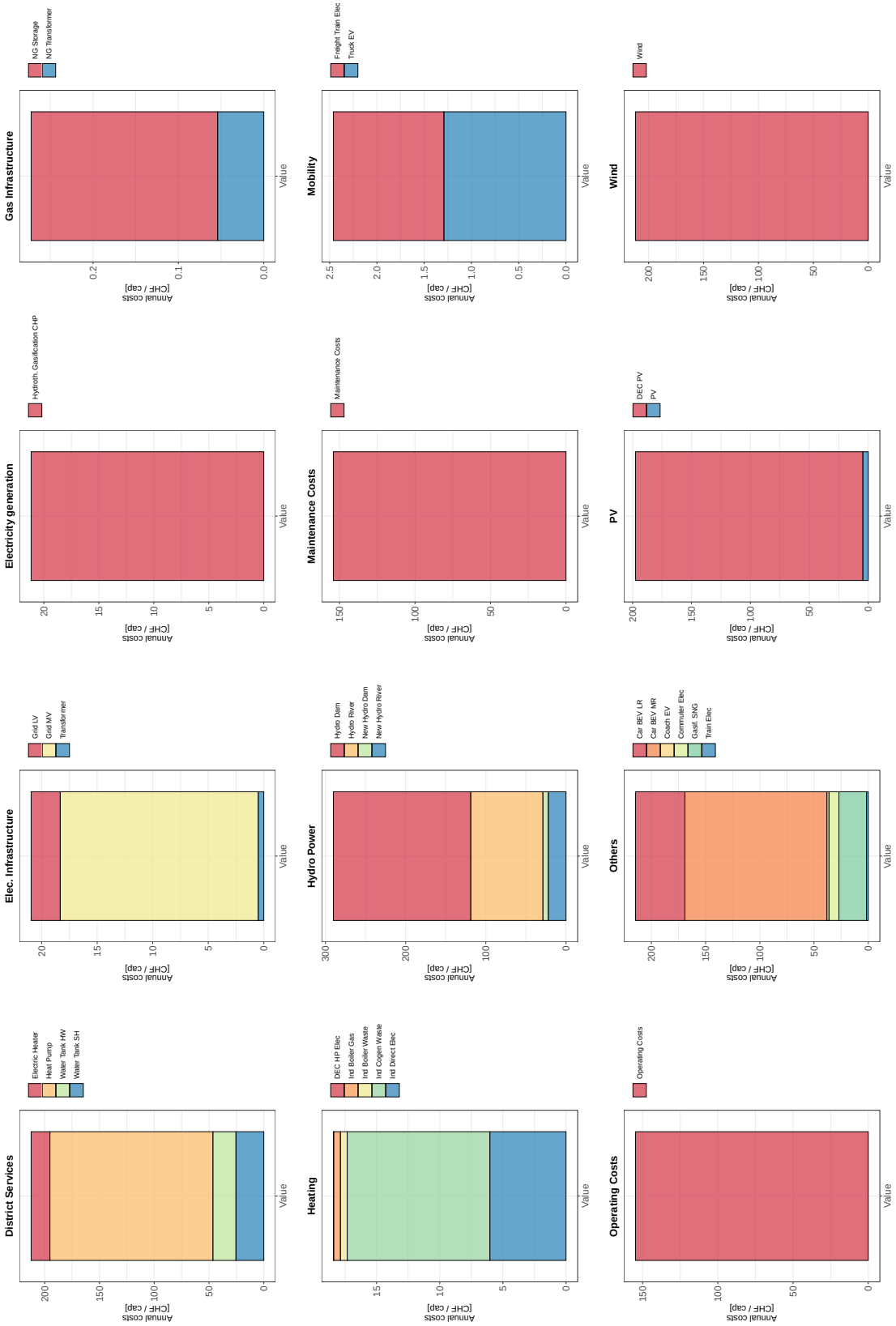
PV

Figure E.4: Distribution of technologies production across Switzerland.

E.1.3 Annual cost of SREHO2ES



E.1.4 Annual cost of SES



E.2 Additional results of the impact of feed-in tariff