Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of the Environment, Transport, Energy and Communications DETEC

**Swiss Federal Office of Energy SFOE**
Energy Research and Cleantech

# SWEET Call 1-2020: DeCarbCH

# Deliverable report

| **Deliverable n°** | 3.1.1 |
|---|---|
| **Deliverable name** | Selection of representative districts |
| **Authors**<br>The authors bear the entire responsibility for the content of this report and for the conclusions drawn therefrom. | Markus Auer, HSLU, **markus.auer@hslu.ch**<br>Luca Brauchli, HSLU, **luca.brauchli@hslu.ch**<br>Núria Duran Adroher, HSLU, **nuria.duranadroher@hslu.ch**<br>Philipp Schütz, HSLU, **philipp.schuetz@hslu.ch**<br>Willy Villasmil, HSLU, **willy.villasmil@hslu.ch**<br>Jörg Worlitschek, HSLU, **joerg.worliteschek@hslu.ch** |
| **Delivery date** | March 2023 |

## Table of contents



sweet swiss energy research for the energy transition

## Summary

This deliverable reports on the bottom-up methodology and results to select representative thermal network districts. Data from the Swiss Federal Office of Energy and the Federal Register of Buildings and Dwellings were used to locate the existing thermal networks and the connected buildings to them in a map. Spatial relations between networks and clusters of buildings were classified in 7 types. Measures on refining the clustering were taken where necessary. The final number of clusters was 1381. They were used to select the representative districts by means of a K-medoids algorithm. Thirteen input properties including the district size, the building density distribution, the share of building types, and building age were considered. Eight districts representative of the 1381 existing ones were obtained.

# 1 Introduction

Thermal networks are essential for the energy transition of the Swiss building stock. The building stock of districts connected to thermal networks is highly heterogeneous and so are the demand structures that set the requirements on the network. This makes it difficult to decide on what decarbonisation path should be applied to a specific district. If instead of the high number of current districts connected to a thermal network one had a selection of them, being the selected ones representative of the current districts, the task of assigning district decarbonisation paths would be considerably easier. This deliverable presents a bottom-up methodology to define such representative Swiss districts, which are called district archetypes. In contrast to network archetypes, that represent the variety of mainly technological properties of thermal networks to supply thermal services to the district, district archetypes are purely focussing on properties that define the demand structures and the various district customer groups. The installation of a thermal network in a district therefore combines the demand side that is represented by a district archetype and the supply side that is represented by a network archetype. Figure 1 presents an overview of this interconnection.

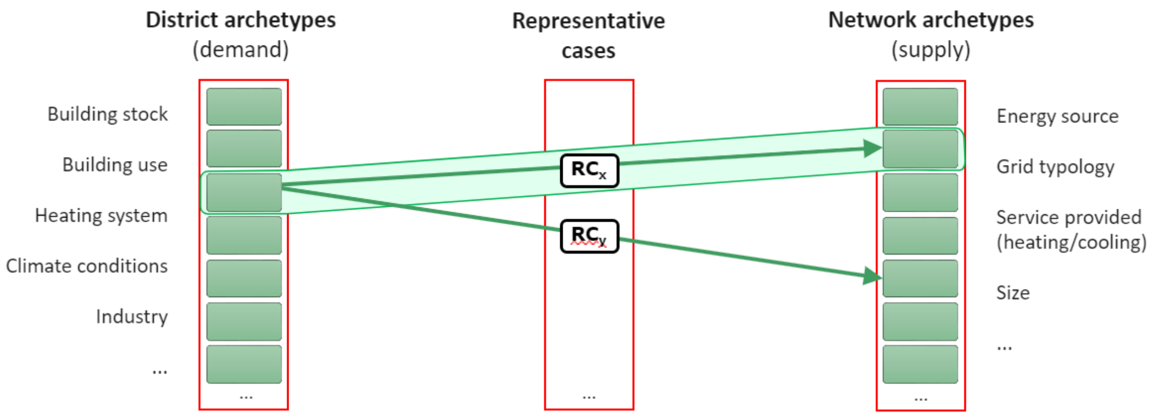



Figure 1: Interrelation of district and network archetypes. A specific demand archetype can possibly be served by multiple network archetypes. Case studies that are representative for one of these connections serve as reference cases (RC).

# 2 Deliverable content

## 2.1 Bottom-up identification of building clusters

Data from the Federal Register of Buildings and Dwellings (RBD) and from the list of thermal networks of the Swiss Federal Office of Energy (SFOE) were used [1]. From the RBD, the buildings with a thermal network connection were identified and from the SFOE list the registered thermal networks were considered. A total number of 80’955 connected buildings and of 1’150 thermal networks are registered in these data sets. The list of thermal networks has been elaborated throughout the past four years and is continuously updated. Therefore, not every single network in place is registered in this data set. The reliability of the information in the RBD depends on the local authorities of the individual municipalities, since it is their responsibility to keep the RBD up to date. Hence, it is possible that a network is registered in the SFOE list, but the RBD information does not give a clear indication about the territory of the network and vice versa.

Firstly, buildings with a thermal network connection (as indicated in the RBD) were clustered using the Density-Based Spatial Clustering of Applications with Noise algorithm (DBSCAN). This algorithm identifies points that are densely packed together and marks them as a cluster. Points that do not fulfil the density criterion are marked as outliers. A cluster is a conglomeration of buildings, and outliers are individual buildings not belonging to a cluster. The idea behind this is that buildings with a network connection that are close to each other most probably are connected to the same thermal network. In this study the density criterion is fulfilled if there are more than 5 buildings within a radius of 200 m. The identified locations of clusters of buildings are represented on a map together with the locations of the thermal networks, as shown in Figure 2. The blue flags indicate the location of thermal networks as given by the SFOE list and the coloured point clouds indicate the different clusters. Black points indicate outliers.

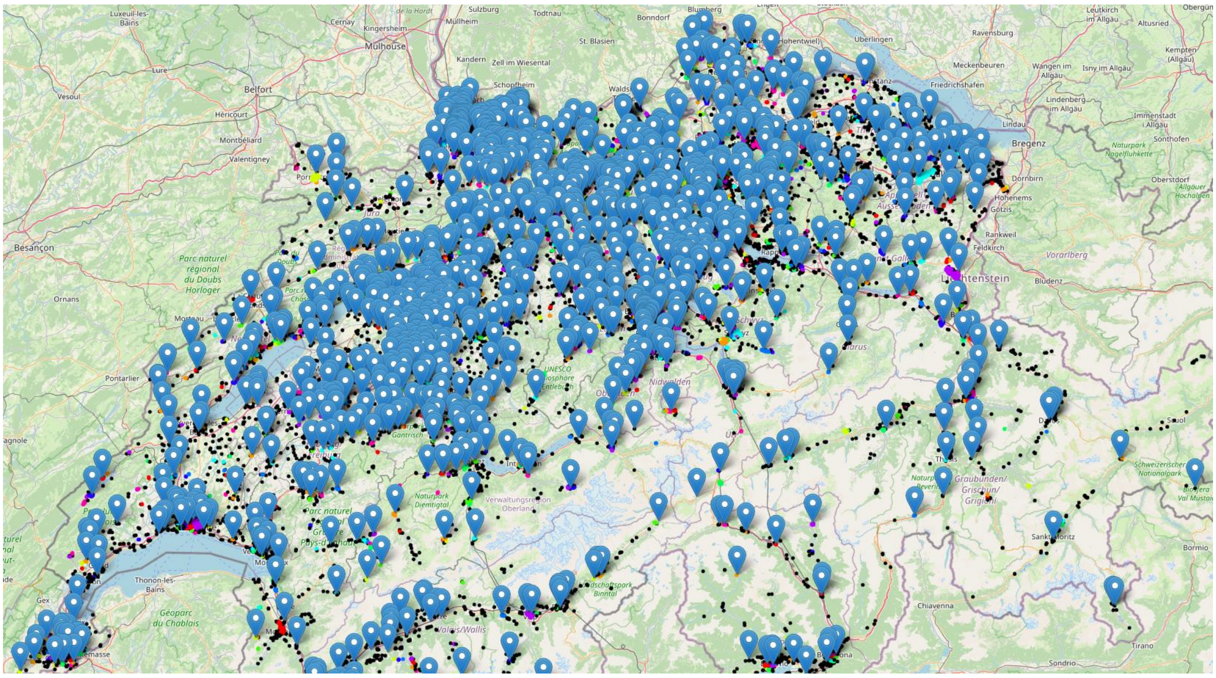


Figure 2: Spatial representation of buildings connected to thermal networks (coloured points) and existing thermal networks (blue flags). Black points represent buildings not belonging to any cluster and are called outliers.

Secondly, the spatial relations of the identified clusters and the networks were classified into seven types A-G as seen in Table **1**. This information helped to assess the probability whether the clustering algorithm correctly identified the buildings of a network or not and for which situations automated or manual measures were needed to correctly identify the districts. Type A contains multiple clusters that are close to the same thermal network. These clusters are merged in a 'supercluster', so that the buildings from the clusters close to the same network belong to the same conglomeration. Type B consists of isolated networks which are not close to any cluster and therefore are not considered in the analysis. Type C contains all the outliers and is not subject to the analysis. Type D consists of clusters not being close to a network and possibly correspond to networks not registered in the SFOE list. No measure is needed for them. Type E corresponds to the ideal situation where the algorithm correctly assigned the buildings close to a network into a single cluster. No measure is needed for them. Type F corresponds to a situation where multiple thermal networks are close to each other and closest to the same cluster. For most clusters of type F, no manual measure was taken due to the mostly small size of the corresponding networks. The effort of researching the information to manually select the buildings correctly outweighs the benefit due to their generally minor importance in terms total number of buildings affected (< 5%) . In type G multiple clusters and multiple networks are close to each other. A manual selection of those clusters was taken. The classification in types A, D, E, and F was done automatically considering the number of buildings of the cluster and the minimal distances between cluster and networks (and viceversa).

Thirdly, the measure for type G was taken: clusters of 92 districts of type G were manually selected with the program QGIS based on publicly available information about the network territories. Type-G clusters contained a total of 39'184 buildings and covered the major cities and the biggest networks in terms of installed power as reported by the SFOE list of thermal networks.

Fourthly, a subsequent DBSCAN clustering analysis with the remaining 41'771 buildings that were not manually assigned was performed. 1'348 clusters were identified, and 9'440 buildings were outliers (the black points in Figure 2 or type C). 109 clusters were of type A and were merged in 50 superclusters. 426 clusters were of type E, and 76 of type F. The total number of clusters after having taken the manual and automatic measures ended up to 1'381.

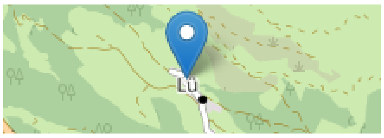
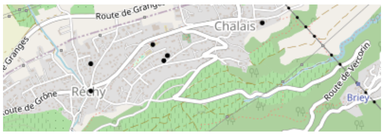
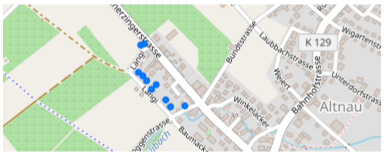
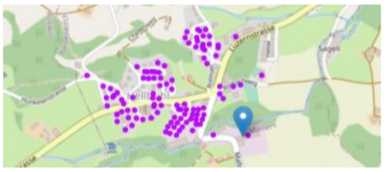
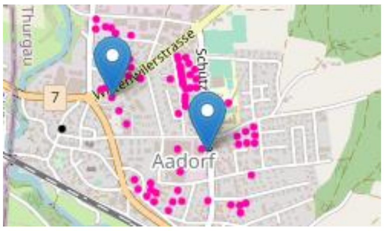
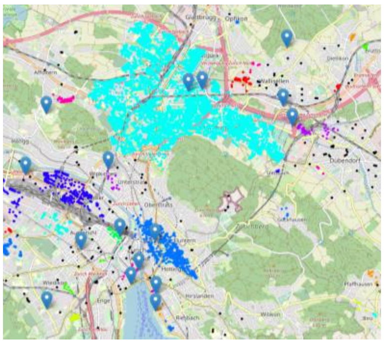
| Type | Example | Clusters | Thermal grids | Interpretation | Measure | Nr. of clusters affected | Nr. of buildings affected | Median / Mean nr. of buildings |
|---|---|---|---|---|---|---|---|---|
| A |  | > 1 | 1 | Algorithm wrongly divided the buildings of a network into multiple clusters | Merge clusters automatically into a supercluster | 109 (resulting in 50 superclusters) | 3'111 (3.84%) | 41 / 62.22 |
| B |  | 0 | 1 | RBD information insufficient OR wrong thermal network location | Ignore these buildings | - | - | - |
| C |  | 0 | 0 | RBD information insufficient | Ignore these buildings | - | 9'440 (11.66%) | 1 / 1 |
| D |  | 1 | 0 | Cluster shows a network that is possibly not yet registered in the SFOE list | No measure needed | 737 | 10'302 (12.73%) | 9 / 13.98 |
| E |  | 1 | 1 | Algorithm correctly assigned the buildings close to a network into a single cluster | No measure needed | 426 | 14'942 (18.46%) | 19 / 35.08 |
| F |  | 1 | > 1 | Algorithm wrongly merged the buildings of multiple networks OR wrong location of any of the thermal networks close to each other | Manual selection needed | 76 | 3'976 (4.91%) | 33.5 / 52.32 |
| G |  | > 1 | >1 | Algorithm is not able to identify the territories correctly as multiple networks are adjacent to each other and/or spatially separated areas are interconnected | Manual selection needed | 92 | 39'184 (48.4%) | 92 / 425.91 |

Table 1: Spatial classifications of cluster-network relations.
A supercluster is a conglomeration of 2 or more clusters

## 2.2 Identification of district archetypes

The methodology to derive district archetypes is partially based on the approach presented in [2]. In contrast to the assessment in [2], our analysis was not conducted within the boundaries of the political communities but within the boundaries defined by the 1381 clusters identified in the previous step. For each cluster, a raster field with cells of 100m x 100m was set up and only the cells containing buildings of the cluster were considered.


Figure 3: Example of a raster field that is considered as the territory of the district. The red points indicate the buildings connected to the network, the yellow points the remaining buildings within the raster field and the grey points the remaining buildings that were not considered as they are not enclosed by the field.

For each cell, the total floor area of all buildings in the cell (including the buildings that are not connected, i.e., the red and yellow points in *Figure 3*) was divided by the cell's area of 10'000 $m^2$. The total floor area is estimated by the ground area and the number of floors, both of which are available in the RBD. The resulting floor area ratio (FAR) indicates, for each cell, how densely the corresponding area is built. This quantity gives an indication on both the energy demand density and the availability of space for infrastructure. Based on the minimum requirements on the FAR ("Geschossflächenziffer") in the canton of Bern [3], a value of less than 0.4 was considered as low density (LD), a value above 1 as high density (HD) and between 0.4 and 1 as medium density (MD). The share of high, medium, and low density cells in percent were then amongst other parameters used to characterize the districts.

The building stock is characterized by the building type and age. The percentage of single family houses (SFH), multi family houses (MFH), buildings for offices and services (SER) and industrial buildings (IND) was assessed for each district according to the classification in the RBD (see **Table 2: Allocation of the Building class (GKLAS code in RBD) to the classification**Table 2). This classification does not only differentiate different demand structures but also different building ownership structures. As found in the ITC QUBE project, owners of single-family houses, for example, have different interests when it comes to investment decisions than owners of multifamily houses who might want to generate profit. Such differences between customer groups are highly relevant for thermal grid providers towards the initiation or expansion of networks.

| Building class (GKLAS code in RBD) | Classification |
|---|---|
| 1110, 1121 | SFH |
| 1122, 1130 | MFH |
| 1211, 1212, 1220, 1230, 1231, 1261, 1262, 1263, 1264, 1265 | SER |
| 1251, 1252 | IND |

Table 2: Allocation of the Building class (GKLAS code in RBD) to the classification

The age of the buildings was characterized by five intervals of construction years, chosen according to Table 3. Similar to the building type, the percentage of the number of buildings that was erected during each interval was calculated.

| Building period GBAUP code in RBD) | Period | Interval-Nr |
|---|---|---|
| 8011 | < 1919 | 0 |
| 8012, 8013, 8014 | 1919 – 1970 | 1 |
| 8015, 8016, 8017 | 1971 – 1990 | 2 |
| 8018, 8019, 8020 | 1991 – 2005 | 3 |
| 8021, 8022, 8023 | > 2005 | 4 |

Table 3: Allocation of the Building period (GBAUP code in RBD) to the interval number.

The density and building stock properties were applied to a K-Medoids algorithm to group the districts by similar characteristics. This algorithm divides the data into K groups so that the centre of each group (called 'medoid') has a minimal sum of distances to all objects in the group. The data lie in a N-dimensional space defined by N properties. In addition to the mentioned size-independent properties, the absolute size of each district by the number of house connections was also included in a normalized scale (i.e., values between 0 and 1). In total, N=13 properties were considered, creating a point cloud in a 13-dimensional space where the grouping takes place.

As around half of the buildings belong to networks with less than 200 house connections, the grouping was executed twice: once for districts with more than 200 connections and once for districts with more than 10 and less than 200. The reason for this is that outliers have a small weight in the K-Medoids algorithm and so the big thermal networks –of which there is a small number and therefore tend to be treated as outliers-- are underrepresented by the set of medoids. This contrasts with the fact that those networks are of high relevance towards decarbonisation. On the other hand, the 538 networks with less than 10 buildings were not considered since they only represent less than 5% of all buildings connected to thermal networks. The 804 districts with more than 10 and less than 200 connections account for around 42% of all buildings and the remaining 39 districts with more than 200 connections account for 54%. For both groupings, a target number of 4 groups was set. In the normalized space, the algorithm tries to minimize the sum of distances between each point and the medoid of its group. Therefore, the medoid is the district that is most representative for its group towards the above-mentioned properties. Table 4 shows the resulting medoid districts, their characteristics and the number of districts and buildings they represent (including themselves).

| Community / Network | Nr. of buildings of the network | SFH [%] | MFH [%] | SER [%] | IND [%] | LD cells [%] | MD cells [%] | HD cells [%] | < 1919 [%] | 1919 - 1970 [%] | 1971 -1990 [%] | 1991 - 2005 [%] | > 2005 [%] | Nr. of networks in the group | Nr. of buildings in the group |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Zürich (ERZ) | 6'580 | 30 | 54 | 12 | 2 | 11 | 33 | 56 | 8 | 57 | 8 | 10 | 16 | 2 | 15' 094 |
| Agro Energie Schwyz | 912 | 47 | 41 | 6 | 3 | 41 | 46 | 13 | 9 | 20 | 16 | 17 | 39 | 18 | 9'452 |
| Martigny | 449 | 16 | 72 | 10 | 0 | 21 | 39 | 40 | 12 | 44 | 24 | 8 | 12 | 12 | 7'600 |
| Münchenbuchsee | 229 | 66 | 27 | 6 | 0 | 59 | 39 | 2 | 4 | 11 | 18 | 40 | 27 | 7 | 5'925 |
| Schwarzenburg | 59 | 57 | 34 | 7 | 2 | 41 | 48 | 11 | 21 | 26 | 24 | 12 | 16 | 205 | 11'222 |
| Escholzmatt-Marbach | 48 | 71 | 24 | 3 | 3 | 71 | 29 | 0 | 22 | 21 | 21 | 17 | 19 | 222 | 9'897 |
| Lyss | 17 | 22 | 64 | 14 | 0 | 30 | 60 | 10 | 15 | 40 | 18 | 4 | 22 | 193 | 4'484 |
| Vevey | 10 | 1 | 64 | 16 | 10 | 25 | 12 | 62 | 22 | 40 | 14 | 12 | 12 | 119 | 2'102 |

Table 4: Table of the medoid districts for the identified 8 groups of districts. The last two columns show how many networks / buildings are represented by the shown medoid districts.

These preliminary results are plausible because different ranges of the various parameters are represented. However, they need to be subject to the following further investigations:

- On the one hand, the resulting set of medoid districts is highly sensitive to the chosen input parameters. The current set of parameters characterizes the built environment that reflects the heat demand densities and the customer groups with their behaviour. Process heat demand by industry is currently underrepresented by this set of parameters but it will be crucial when it comes to assess the exergetic performance of the networks. Acquiring data about the needs of different industrial sectors will therefore be tackled in the next steps. Generally, the energy and exergy demand has not been part of the analysis yet and will be included for all sectors. These will be based on the results of the deliverable D1.1.1 of WP01. To assess the exergetic demand patterns it will be essential to properly quantify the potential of industrial heat demand in terms of power and temperature requirements for different sectors. The quantification of these requirements will be tackled by WP04 and the geospatial information about the location and sector of industrial sites is currently developed by WP01.
- On the other hand, the medoid districts should ideally underly good RBD data quality and the possibility to access real consumption data from the operator, in order to build detailed cases for the other tasks and WP to work with. If this is not the case, it might be meaningful to choose a network that is close to the current medoid and fulfils these criteria as new representative of the group.

# 3 Conclusion

More than 1300 districts of currently installed thermal networks have been identified by a novel bottom-up approach based on information of the RBD. Approximately 600 of these districts consist of less than 10 buildings and together account for less than 5 % of all buildings connected to a thermal network and were neglected. The 39 biggest districts account for more than 50% of all buildings served by district heating. By a multidimensional K-Medoids grouping, groups of districts that have similarities in terms of size, building stock characteristics and building density were identified. Each group is represented by its medoid district. The groups and medoids that are chosen by the algorithm are highly sensitive to the properties chosen to characterize the districts. Energetic and exergetic demand properties can either be included characteristics for the grouping or as description of the groups presented here. The goal in both cases is to describe the chosen representative medoid districts in more detail by exergy demand, power demand profiles, etc. Assessing the exergy demand will hereby be of high relevance and therefore, a good knowledge of industrial energy demand patterns is crucial.

# 4 References


[1] Bundesamt für Energie BFE, “Thermische Netzte.” https://opendata.swiss/de/dataset/thermische-netze-nahwarme-fernwarme-fernkalte/resource/a8bc6284-4153-4744-b149-e642b7077663 (accessed Feb. 02, 2023).

[2] P. Murray, J. Marquant, M. Niffeler, G. Mavromatidis, and K. Orehounig, “Optimal transformation strategies for buildings, neighbourhoods and districts to reach CO2 emission reduction targets,” *Energy Build.*, vol. 207, p. 109569, Jan. 2020, doi: 10.1016/J.ENBUILD.2019.109569.

[3] L. Bühlmann, “Bericht Regelungen zur Förderung der Verdichtung und zur Beseitigung von Verdichtungshemmnissen,” 2019.