# IDENTIFICATION OF TYPICAL ENERGY GRID STRUCTURES BY CROSS-SECTORAL CLUSTERING

**Patrick Wintzek[1], Tobias Riedlinger[2], Shawki Alsayed Ali[2], Bernd Bastian Wierzba[2], Markus Zdrallek[2], Heike Barth[3], Sigrid Plötz[3]**

[1] Institute of Power Systems Engineering, University of Wuppertal, Rainer-Gruenter-Str. 21, 42119 Wuppertal, Germany, +49 202 439 1907, patrick.wintzek@uni-wuppertal.de, https://www.evt.uni-wuppertal.de/

[2] Institute of Power Systems Engineering, University of Wuppertal, Germany

[3] Rheinische NETZGesellschaft mbH, Parkgürtel 26, 50823 Köln, Germany

**Abstract:** This contribution focuses on the identification and analysis of correlated energy grid structures. Firstly, a *k*-means algorithm is implemented to cluster the low-voltage and medium-voltage electricity grids as well as gas grids. Secondly, the correlations of typical grids across voltage-levels as well as sectors are investigated. As a result, similar grid structures in the electricity and gas clusters can be predominantly detected in certain regions. Finally, the regionalization of heat pumps shows a high concentration in certain energy clusters. Additionally, the presence of electric heat pumps is analyzed on gas grid clusters. In conclusion, the structure of existing energy grids is recommended to be analyzed together to exploit synergy effects for both sectors.

**Keywords:** clustering, heat pumps, low- and medium-pressure gas grids, low-voltage and medium-voltage electricity grids

## 1 Introduction

In order to further develop distribution grid infrastructure in a future-proof manner, urban energy grids in particular are faced with the challenge of increasing electrification of the mobility and heating sectors. Thus, an infrastructural change in the energy grids is expected. Due to the future coupling of the different energy sectors, it is important to ensure corresponding sustainability through cross-sectoral analyses and grid planning. Therefore, a combined consideration of electricity and gas distribution grids is essential in the upcoming years.

### 1.1 Literature Review and Novelty

Publications such as [1-5] already describe exemplary data sets, but only in the area of electricity grids. This contribution analyzes cross-sectoral distribution grid structures for electricity and gas grids of an urban distribution system operator (DSO). It extends the analyses and results from [6], in which urban electricity distribution grids are identified based on extensive data sets and contrasts them with the gas grid for one city area. A cross-sectoral clustering is conducted and superimposed on the cross-voltage level clustering carried out in [6]. The results then serve, among other things, to estimate the effect of progressive electrification on gas grids and cross-sectoral grid planning while considering commercial market and geodata according to [7].

### 1.2 Structure and Objective

Section 2 describes the data on which the evaluations are performed. Section 3 then briefly explains the clustering procedure and used parameters. Section 4 presents the respective results based on the performed out clustering. Firstly, the clustering for the low-voltage (LV) level and the medium-voltage (MV) level are presented for the electricity grids. Then the correlation of these two voltage levels is analyzed. Subsequently, the method for clustering the low- and medium-pressure grids is explained and the corresponding results are presented. Afterwards, the correlation between the electricity and gas grids is analyzed. As a supplementary evaluation, the influence of a regionalization of electric heat pumps (HPs) on the analysis and future grid planning of the electricity and gas grids is presented. Section 5 then ends with a conclusion and recommendations as a summary of the results.

## 2 Dataset

Firstly, the complete dataset from [6] is used for the electricity grids. The data derives from six large DSOs in Germany (see **Figure 1**). They cover a large range of urban distribution grids (feeding from 100,000 up to 3.5 million residents) and thus can be classified as representative for other urban DSOs. The clustered dataset consists of 4,207 LV grids and 126 MV grids. In addition, all low- and medium-pressure gas grids are available for one DSO (operating in the city of Cologne), which are used as basis for the following analysis.

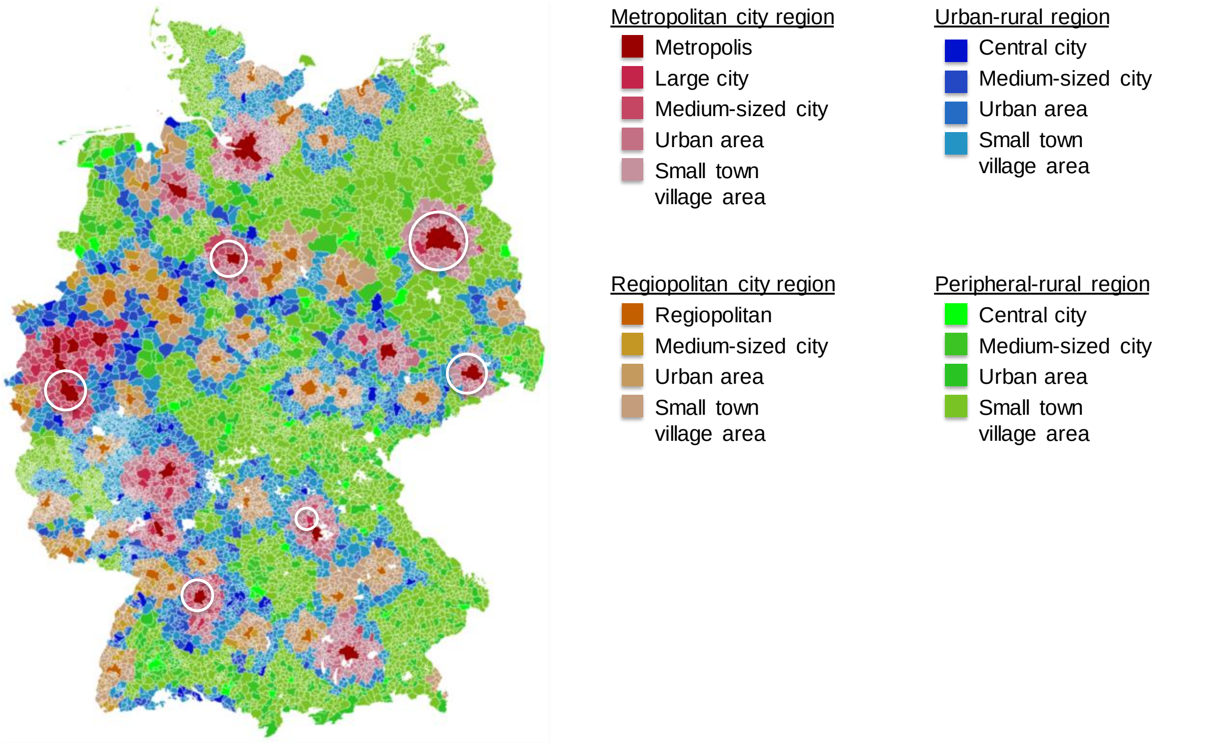


**Figure 1** German-wide community structure and six distribution system operators (white circles) whose electricity grids are available for the analysis [6]

Commercial market and geodata from [7] are used to supplement the analysis and evaluations with further data. An exemplary evaluation for Cologne can be found in **Figure 2**, which shows how the detached and semi-detached houses as well as multi-family houses are distributed across the city.

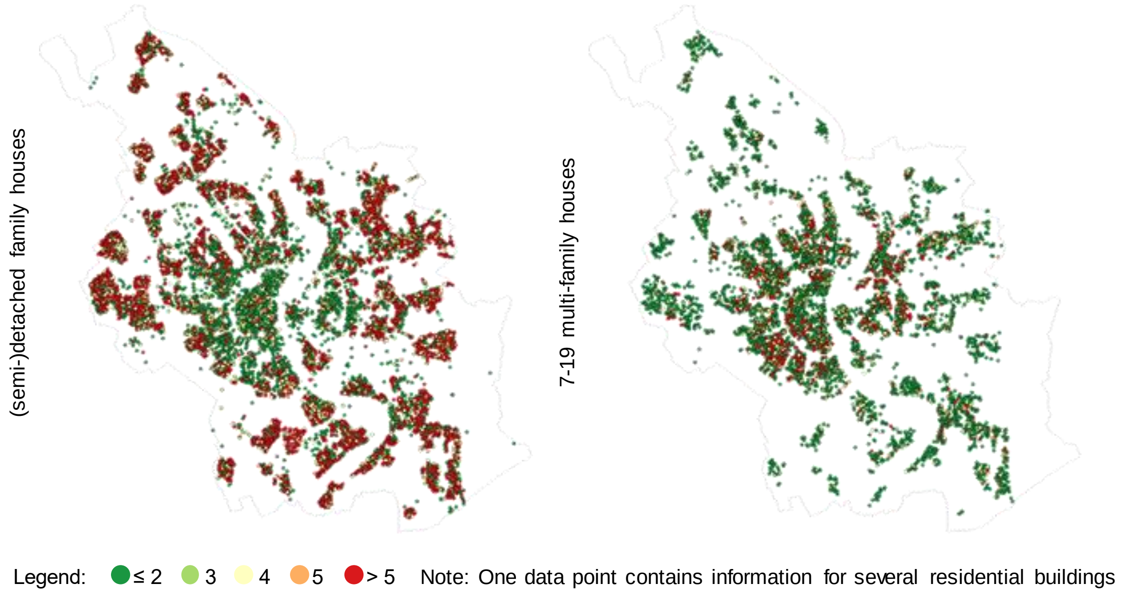


**Figure 2** Visualization of detached, semi-detached and 7-19 multi-family houses [7]

# 3 Clustering

Due to their early development and stability, partitioning algorithms such as *k*-means [8] are widely used for dataset analysis and are therefore used for the clustering in this contribution. Other algorithms such as hierarchical algorithms, e. g. Single Linkage [9] or density-based algorithms, e. g. DBSCAN [10] would have been equally applicable. [6]

The *k*-means algorithm follows a simple three-step procedure to perform data partitioning. Firstly, the algorithm randomly distributed *n* cluster centers (CCs) on the input data. By calculating the least distance between each data point and the CCs, each data point is assigned to one of the CCs. Secondly, a correction of the values of the CCs is done so that they now represent the center of their assigned data points. Thirdly, the data points are data validated to ensure they are correctly mapped to the closest data center. Otherwise, the data point(s) get reassigned to a different CC. The second and third step are iteratively repeated until no further CC need to be corrected and no data point needs to be reassigned. [6,11]

With the *k*-means algorithm, it is important to note that the number of clusters must be specified before the algorithm is applied to the dataset. Furthermore, the results are not fully re-traceable, as they depend on the initial random distribution. Repeating *k*-means clustering with the same parameters applied to the same dataset can lead to different clustering results which are very similar to each other, but still different, especially in the fringes. [6]

The two aforementioned restrictions of the *k*-means algorithm are overcome in this contribution. To overcome the first restriction, the *k*-means algorithm is repeated for different number of clusters and the resulting clusters are analyzed. The analysis determines the appropriate number of clusters so that the clusters are distinct from each other and the data points are consistent within the cluster. The second restriction is fulfilled by repeatedly applying the k-means algorithm for the determined number of clusters and recording the results. The clustering with the optimal results is then selected.

# 4 Results

In the following sections, the results for both the separate clustering of the energy grids as well as the correlations between them are presented.

## 4.1 Low-voltage clustering

Based on the existing data, the connection-density as number of building connections (BCs) per line length (LL), and the load-density as metering points (MPs) per BC, are set to be the clustering parameters for the LV electricity grids.

**Figure 3** shows the clustering results of all 4,207 LV grids on the left side and the resulting mean values of the dataset per cluster on the right side. This shows that the LV clusters are very different from each other in terms of LL, number of BCs and number of MPs per LV grid.

**Figure 4** then shows on the left side an excerpt of the clustering for the city of Cologne, which represents the focus of the present analysis. On the right side of **Figure 4**, the geographical visualization of the LV grid clusters is presented. Here it can be seen that certain clusters are located in certain areas. Especially in combination with **Figure 2**, it can be seen that the clusters LV-C05, LV-C06 and LV-C10 predominate in the suburban areas.

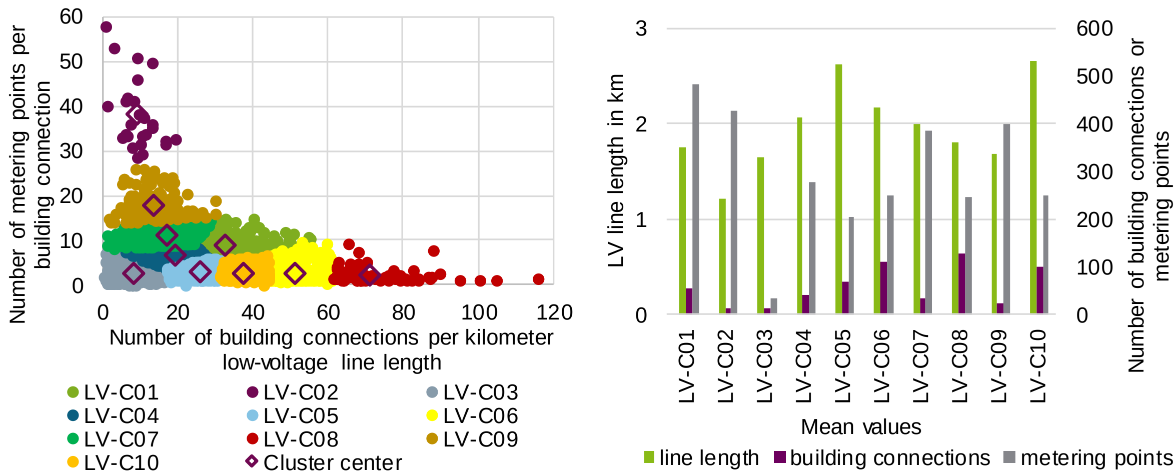


**Figure 3** Clustering of low-voltage (LV) grids (left) and mean values per cluster (right) [6]

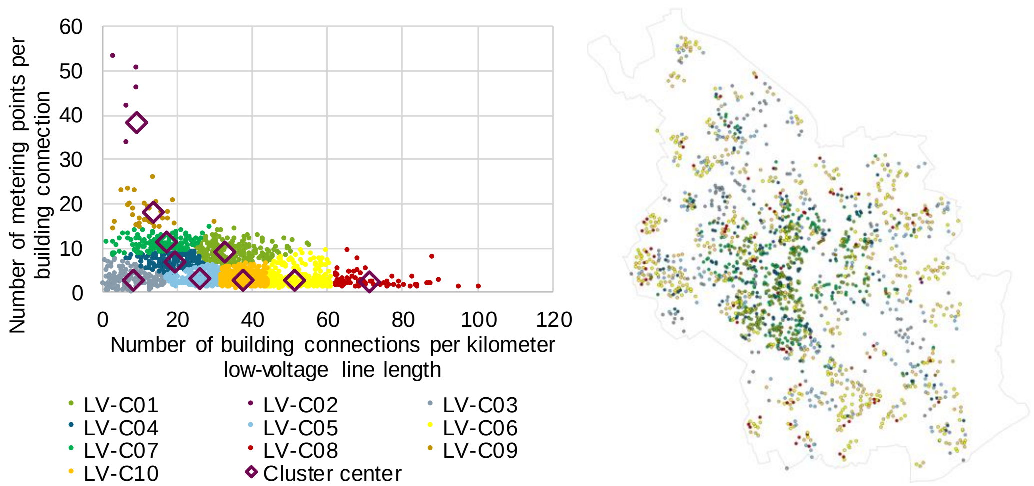


**Figure 4** Excerpt of clusters for Cologne (left) and visualization for the supply area (right)

## 4.2 Medium-voltage clustering

Using the above-mentioned dataset, the connection-density as number of distribution transformers (DTs) per LL, and the load-density as MPs per DTs, are determined as clustering parameters for the MV electricity grids.

Similar to section 4.1, **Figure 5** shows the results of the clustering of all 126 MV grids (left) and the corresponding mean values of the clusters (right). As in the previous section, **Figure 6** shows the excerpt of the clustering for Cologne on the left side and the geographical visualization on the right side. This also shows that MV grid clusters such as MV-C03 and MV-C04 are more likely to be found in the inner-city than in the surrounding suburban area.

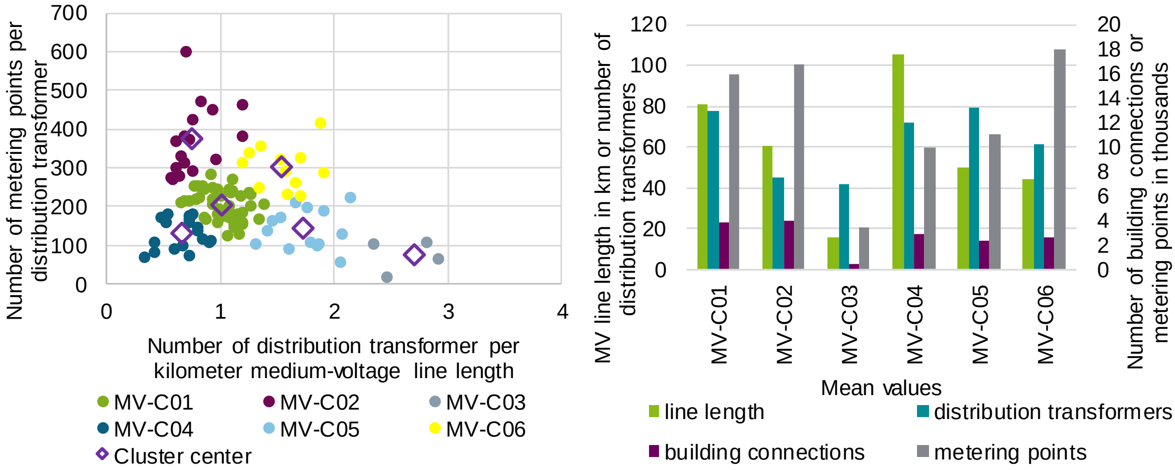


**Figure 5** Clustering of medium-voltage (MV) grids (left) and mean values per cluster (right) [6]

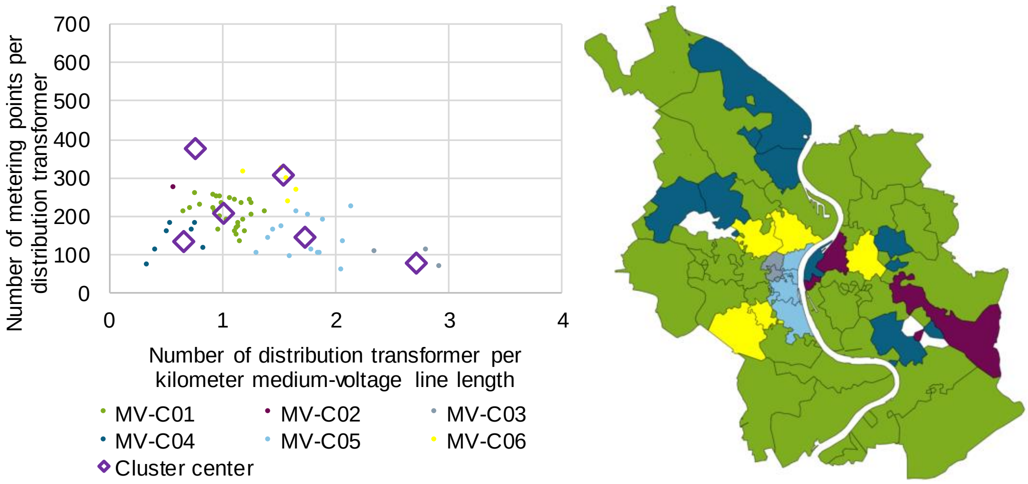


**Figure 6** Excerpt of clusters for Cologne (left) and visualization for the supply area (right)

## 4.3 Correlations of typical grids across voltage levels

The separate clustering per voltage level in section 4.1 and 4.2 are now combined with each other to identify possible correlations. **Figure 7** shows an evaluation in which the LV grids from six DSOs (including Cologne) as well as the LV grids separately for Cologne are shown per cluster. Regardless of whether the results are consolidated (left column in each case) or separately for the city of Cologne, it can be seen that the MV clusters are made up of specific LV clusters. In **Figure 8**, this finding can also be confirmed geographically.

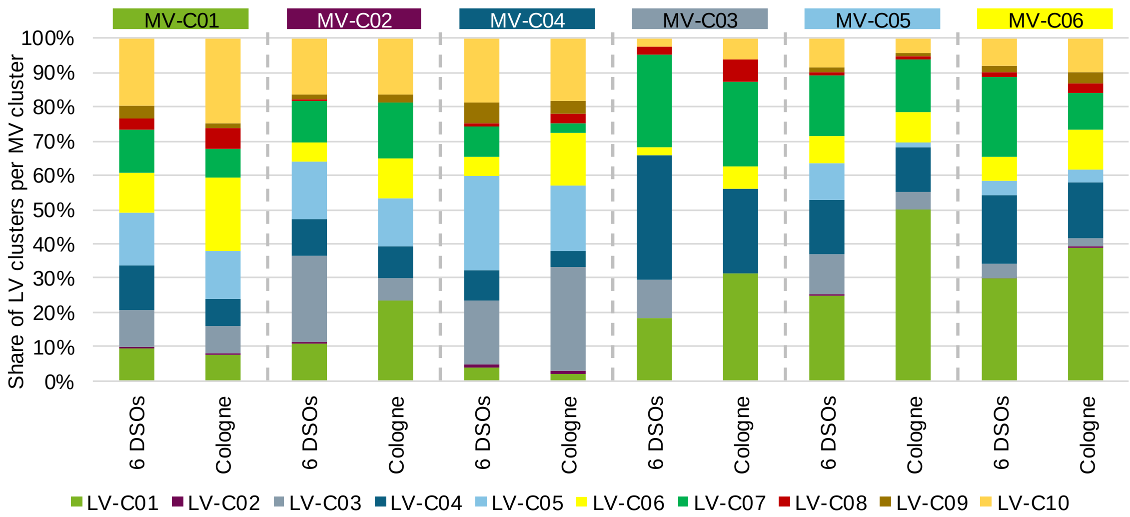


**Figure 7** Correlation between MV and LV clusters for all six DSOs and Cologne

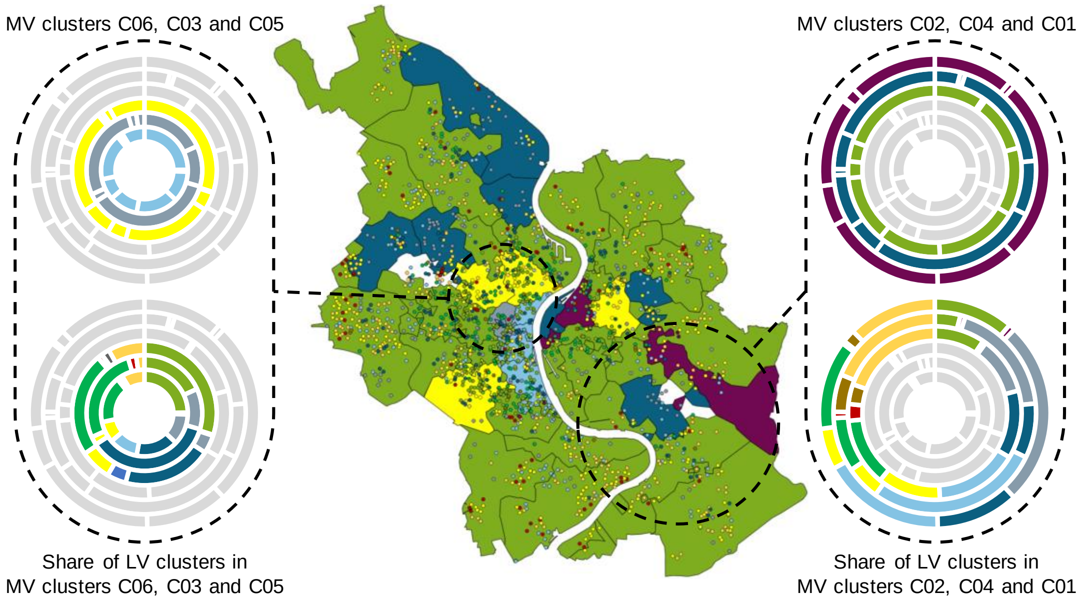


**Figure 8** Visualization of the correlation between low- and medium-voltage clusters

## 4.4 Low- and medium pressure gas clustering

This section aims to identify whether there are correlations between clustering in the gas grid and clustering in the electricity grid. It also attempts to determine whether typical energy grids can be derived from this correlation, or whether there are certain areas where certain grid structures can be found across electricity and gas grids.

### 4.4.1 Data preparation

In principle, the same clustering procedure as described in section 3 is used. However, for gas grids, especially in Cologne, the difficulty lies in separating the grid itself. Unlike electricity grids, gas grids are in most cases operated in a meshed manner across an entire supply area, hence no individual gas grids can be identified. Therefore, the supply areas are determined approximately by means of a geographic area analysis with Voronoi cells [12] in order to approximately assign a gas grid to each of the gas pressure regulating and metering stations.

### 4.4.2 Application

Similar to section 4.1 and section 4.2, **Figure 9** shows the results of the clustering of all 238 low- and medium pressure gas grids on the left side and the geographical visualization on the right side. As already evident for the electricity grids, city areas can also be identified for the gas grid in which certain clusters are primarily located. In the inner city, for example, the clusters G-C03 and G-C04 are mainly present.

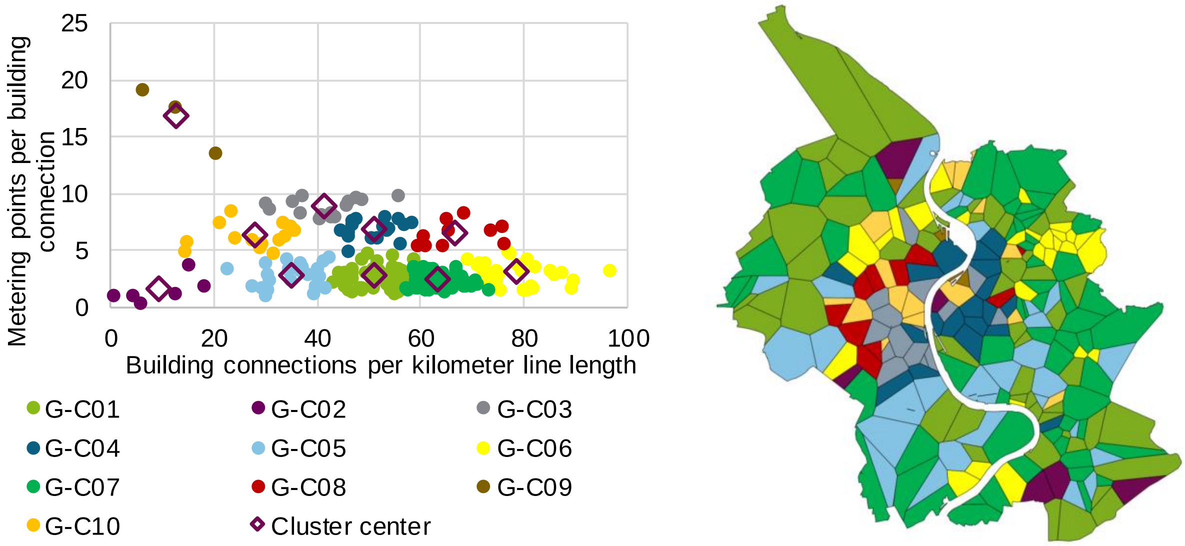


**Figure 9** Resulting clustering of low- and medium pressure gas grids (left) for Cologne and visualization for the supply area (right)

## 4.5 Correlations of typical grids across sectors

As a result of the clustering (see **Figure 9**), **Figure 10** shows how the LV grids colored by clusters (points) are distributed in the gas grids also colored by clusters (polygons). The LV clusters LV-C06 and LV-C10 are mainly located in the gas clusters G-C01, G-C06 and G-C07.

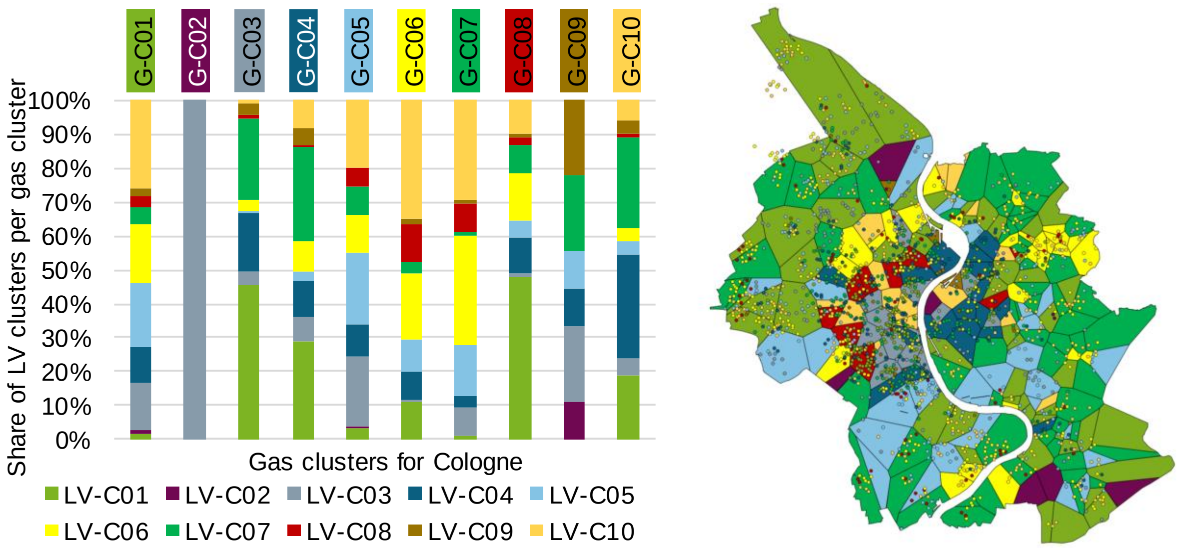


**Figure 10** Correlation between low-voltage and gas clusters (left) and visualization for the supply area (right) (points: local distribution substations, polygons: approximated supply area of the gas pressure regulating and metering stations)

The correlation between gas and electricity grids increases in importance over the upcoming years due to the increasing electrification trend of the heating sector mainly by HPs [13]. Since the integration of HPs happens predominantly in certain LV clusters, their presence on the correlated gas grid clusters needs to be investigated and analyzed.

The prospective use of HPs in electricity grids can substitute the gas heating, which in turn can lead to oversized gas grids in some areas and can even result in a significant decline in gas sales and - in the worst case - may become uneconomical. By the regionalization of HPs, the penetration of HPs in electricity grids can be analyzed within each gas grid cluster.

The regionalization method from [14] is used, in which the progressive scenario for HPs from [15] are first regionalized from country level down to street level. The focus of the distribution is primarily on detached and semi-detached houses because it can be assumed that the share of HPs for multi-family houses will be significantly lower in the future due to building restrictions.

In the top of **Figure 11,** the results for the regionalization of HPs for the year 2050 are displayed. Each data point from [7] contains information about the residential buildings. The lower part of the figure shows the number and share of HPs in the respective gas cluster. A comparison with **Figure 2** shows that the majority of HPs are found in the suburban grid areas. The largest share is found in the gas clusters G-C01 and G-C07, each representing one third of the total number of HPs in the city. Additionally, from the perspective of the LV clusters, the largest shares of HPs are 8 % for cluster LV-C05, 16 % for LV-C06 and 17 % for LV-C10.

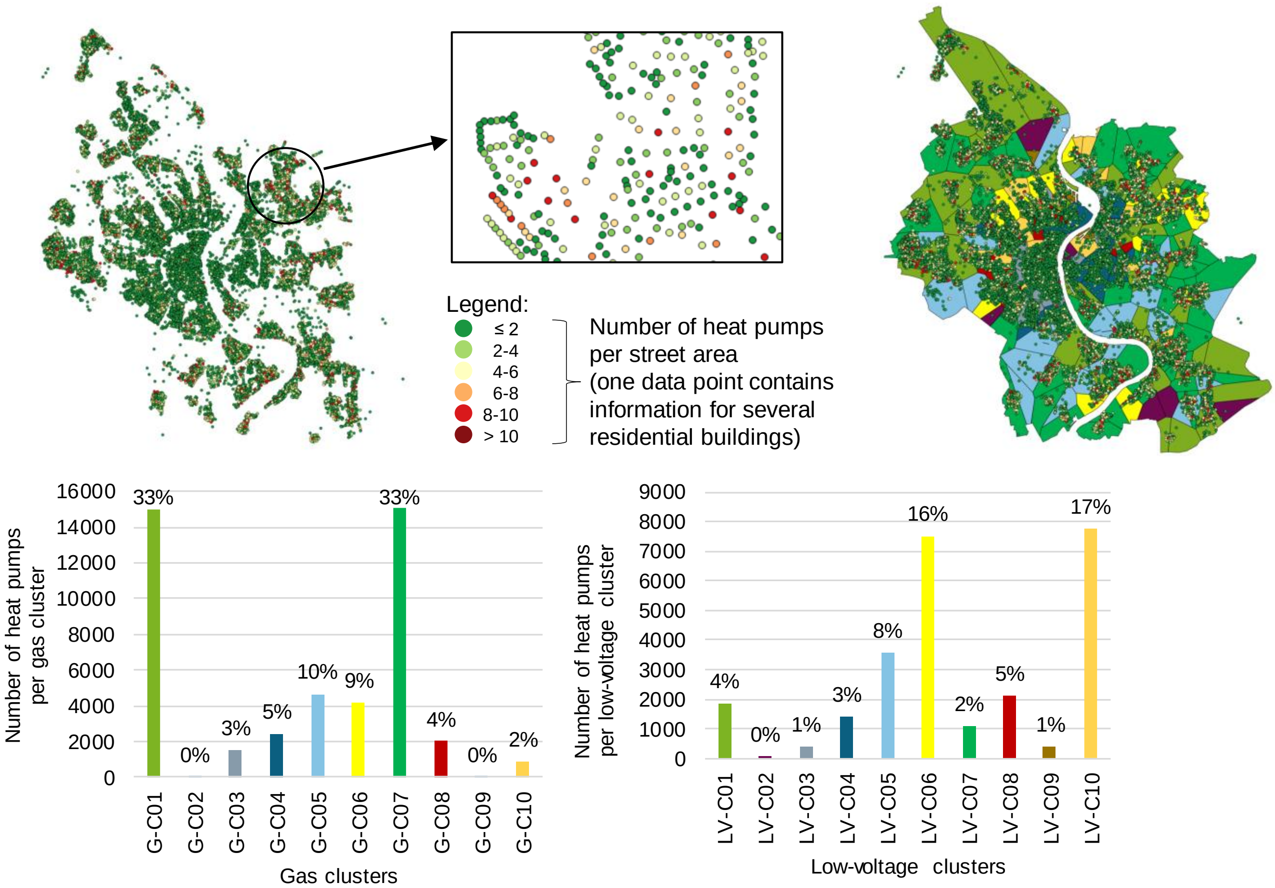


**Figure 11** Visualization of the resulting number of heat pumps (top left) with detailed view (top center) and with gas clusters (top right) as well as the number of heat pumps in the respective gas clusters (bottom left) and low-voltage clusters (bottom right)

## 5 Conclusion and recommendations

In summary, it can be stated that a clustering of the electricity and gas grids shows similar results in terms of grid structure (see **Figure 12**), with certain clusters being found predominantly in the inner-city or on the suburban areas. Consequently, the regionalization of HPs shows that they are concentrated in the respective suburban clusters.

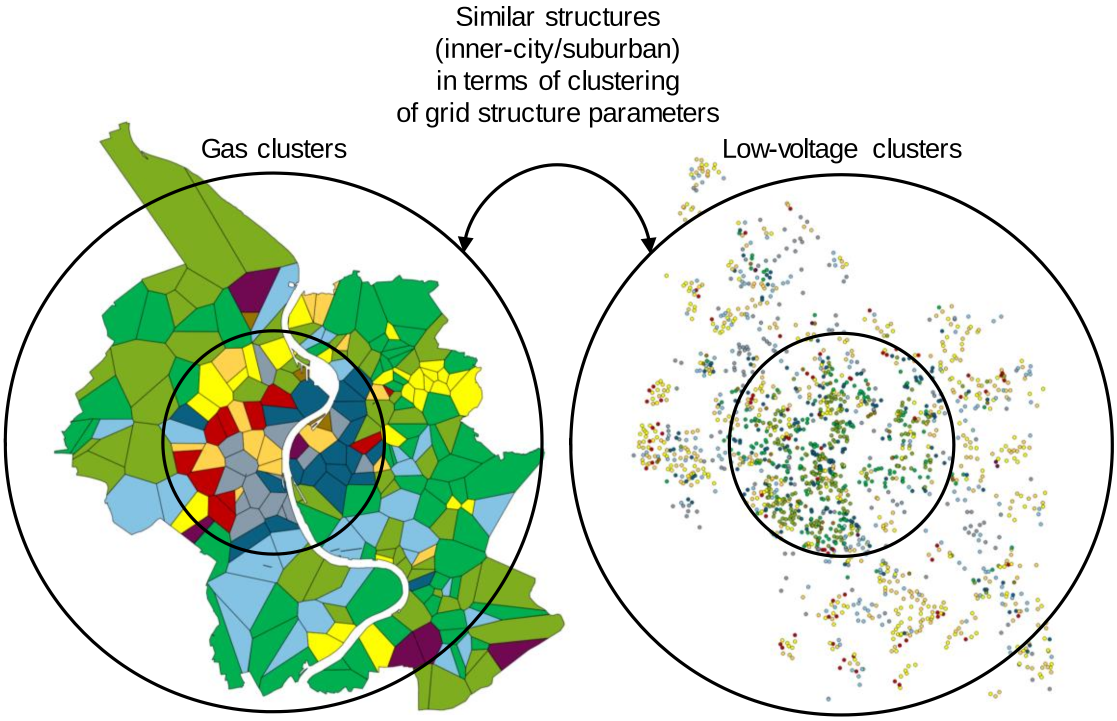


**Figure 12** Visual comparison of the clustered grid structure data for gas grids (left) and low-voltage grids (right)

Based on the analyzes carried out, it is therefore recommended that existing energy network structures be considered together and across sectors as part of the electrification of the heating sector. Especially, the integration of HPs in suburban areas can lead to a considerable load increase in the LV grids and thus also affect the upstream voltage levels. The joint clustering of electricity and gas grids as well as regionalization of HPs for other DSOs can support strategic grid planning. In addition, it can support in conjunction with asset management the company's own renewal strategy and indicate where changes in the supply structure are to be expected in the long term. [16] has indicated that gas grids are not always economical for new development areas where the penetration of a partially electrified heat supply is expected. Based on the present contribution and [16], it can be preliminarily deduced that declining gas sales in existing areas can lead to an inefficient operation of the gas grids. However, this still needs to be confirmed in further work and analyses.

## Acknowledgments

The detailed dataset employed was provided by Rheinische NETZGesellschaft mbH.

## References


[1] Strunz, K., Hatziargyriou, N., Andrieu, C. Benchmark systems for network integration of renewable and distributed energy resources, Cigre Task Force C6.04.02, April 2009.

[2] Schneider, K. P. et al. A taxonomy of North American radial distribution feeders in 2009 IEEE Power Energy Society General Meeting, July 2009, pp. 1-6.

[3] Dickert, J. et al. Benchmark low voltage distribution networks based on cluster analysis of actual grid properties, in PowerTech, 2013 IEEE Grenoble. IEEE, 2013, pp. 1-6.

[4] Höflich, B. et al. dena-Verteilnetzstudie, Ausbau- und Innovationsbedarf der Stromverteilnetze in Deutschland bis 2030, Berlin, 2012.

[5] Meinecke, S. et al. SimBench – Simulationsdatenbasis zum einheitlichen Vergleich von innovativen Lösungen im Bereich der Netzanalyse, Netzplanung und -betriebsführung, Final Report, Kassel, Aachen, Dortmund, 2019.

[6] Wintzek, P. et al. Identification of typical urban grid structures across voltage levels by clustering distribution grid big data, In Proceedings of the International ETG Congress, Wuppertal, Germany, 18-19 May 2021, ISBN 978-3-8007-5549-3.

[7] GfK Society for Consumer Research: Population structure data, 2018.

[8] Hartigan, J. A.; Wong, M. A. A K-Means Clustering Algorithm, Journal of the Royal Statistical Society, Series C (Applied Statistics), pp. 100-108, 1979.

[9] Floek, K. et al. Taksonomia wrocławska, Przegląd Antropol.17, pp. 193-211, 1951.

[10] Ester, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise, KDD-96 Proceedings, Vol. 96, pp. 226-231, 1996.

[11] Moehrke, F. et al. Clustering And Determination Of Relevant Network Operating Points In Analytical Reliability Calculations, CIRED 2019, Madrid, 2019.

[12] González, D. L.; Einstein, T. L. Voronoi cell patterns: Theoretical model and applications, Physical Review E 84, 051135, 2011, DOI: 10.1103/PhysRevE.84.051135.

[13] Sterchele, P. et al. Paths to a Climate-Neutral Energy System – The German Energy Transition in its Social Context, Fraunhofer ISE, Freiburg, 2020.

[14] Wintzek, P.; Ali, S. A.; Kotthaus, K.; Wruk, J.; Zdrallek, M.; Monscheidt, J.; Gemsjäger, B.; Slupinski, A. Application and Evaluation of Load Management Systems in Urban Low-Voltage Grid Planning, World Electric Vehicle Journal 2021, Special Issue "Electric Vehicles Integrated with Green Energy Sources", Bd. 12(3), 91.

[15] Monscheidt, J. et al. An energy demand model for electric load forecast in urban distribution networks, CIRED Workshop, Berlin, June 2020.

[16] Riedlinger, T. et al., Cross-sectoral planning of energy grids in development areas with different heating concepts; Power and Energy Student Summit (PESS), online, Germany, 5-7 October 2020.