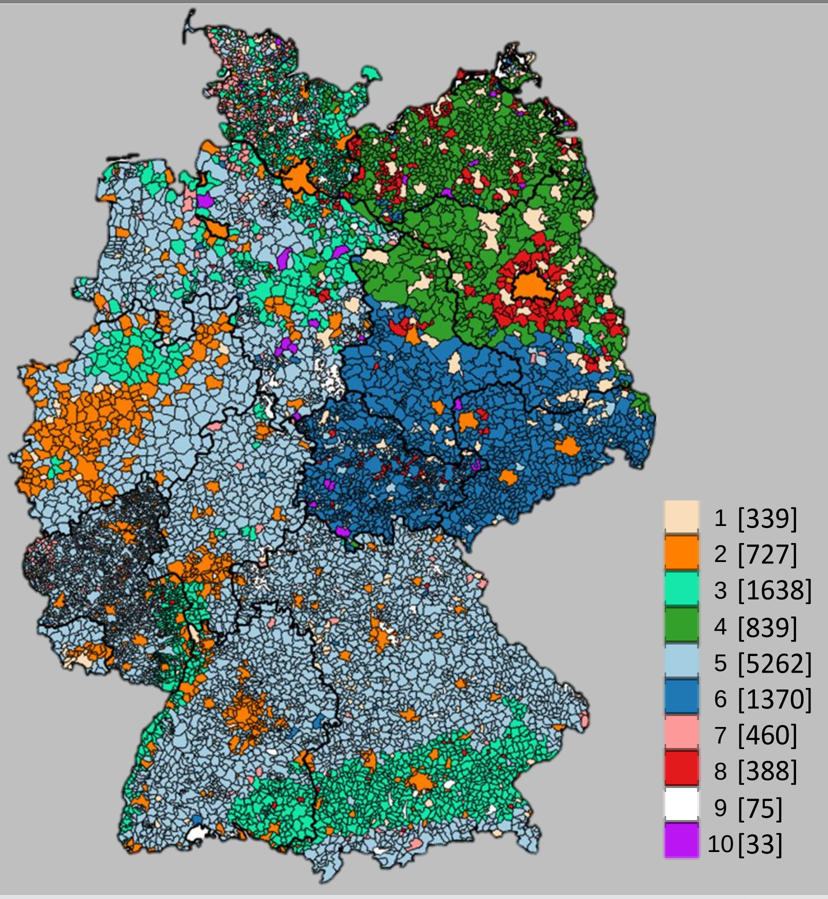
Karlsruhe Institute of Technology

Developing a municipality typology for modelling decentralised energy systems

By Jann Weinand, Russell McKenna and Wolf Fichtner

No. 26 | January 2018

WORKING PAPER SERIES IN PRODUCTION AND ENERGY



Developing a municipality typology for modelling decentralised energy systems

Jann Weinand*, Russell McKenna, Wolf Fichtner

Chair of Energy Economics, Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT), Hertzstraße 16, 76187 Karlsruhe, Germany

* Corresponding author. *E-mail address:* jann.weinand@kit.edu. *Tel.:*+49 721 60844 444

Abstract

The recent rapid expansion of renewable energy capacities in Germany has been dominated by decentralised wind, photovoltaic (PV) and bioenergy plants. The spatially disperse and partly unpredictable nature of these resources necessitates an increasing exploitation of integration measures such as curtailment, supply and demand side flexibilities, network strengthening and storage capacities. Indeed, one solution to the large-scale integration of renewable energies could be decentralised autonomous municipal energy systems. The achievement of grid parity for some renewable energy technologies has strengthened the desire of some communities to become independent from central markets. Whilst many communities in Germany already strive for so-called energy autonomy, the vast majority do so only on an annual basis. Several studies have already analysed the technical and economic implications of the mainly decentralised future energy system, but most are restricted in their insights by limited temporal and spatial resolution.

The large number (11,131) of German municipalities means that a national analysis at this resolution is not feasible. Hence, this study employs a cluster analysis to develop a municipality typology in order to analyse the techno-economic suitability of these municipalities for autonomous energy systems. A total of 34 socio-technical indicators are employed at the municipal level, with a particular focus on the sectors of Private Households and Transport, and the potentials for decentralised renewable energies. The first step is to scale the indicator values and reduce their number by using a factor analysis. Several alternative methods are weighed against each other, and the most suitable methods for the factor analysis are chosen. Secondly, selected quantitative cluster validation methods are employed alongside qualitative criteria to determine the optimal number of clusters. This results in a total of ten clusters, which show a large variation as well as some overlap with respect to specific indicators. For example, one cluster contains all major German cities and has a low potential for renewable energies. Another cluster, on the other hand, contains the municipalities with a higher potential for renewable energies due to their high hydrothermal potential for geothermal power.

An analysis of the municipalities from three German renewable energy projects “Energy Municipalities”, ”Bioenergy Villages” and “100% Renewable Energy Regions” shows that in eight of the ten clusters municipalities are aiming for energy autonomy (in varying degrees). It is challenging to differentiate between the clusters regarding readiness for energy autonomy projects, however, especially if the degree of social acceptance and engagement for such projects is to be considered. To answer the more techno-economical part of this question, future work will employ the developed clusters in the context of an energy system optimisation. Insights gained at the municipal level will then be qualitatively transferred to the national context to assess the implications for the whole energy system.

Working Paper Series in Production and Energy

recent issues

No. 1 Alexandra-Gwyn Paetz, Lisa Landzettel, Patrick Jochem, Wolf Fichtner: Eine netnografische Analyse der Nutzererfahrungen mit E-Rollern

No. 2 Felix Teufel, Michael Miller, Massimo Genoese, Wolf Fichtner: Review of System Dynamics models for electricity market simulations

No. 3 Patrick Jochem, Thomas Kaschub, Wolf Fichtner: How to integrate electric vehicles in the future energy system?

No. 4 Sven Killinger, Kai Mainzer, Russell McKenna, Niklas Kreifels, Wolf Fichtner: A regional simulation and optimization of renewable energy supply from wind and photovoltaics with respect to three key energy-political objectives

No. 5 Kathrin Dudenhöffer, Rahul Arora, Alizée Diverrez, Axel Ensslen, Patrick Jochem, Jasmin Tücking: Potentials for Electric Vehicles in France, Germany, and India

No. 6 Russell McKenna, Carsten Herbes, Wolf Fichtner: Energieautarkie: Definitionen, Für- bzw. Gegenargumente, und entstehende Forschungsbedarfe

No. 7 Tobias Jäger, Russell McKenna, Wolf Fichtner: Onshore wind energy in Baden-Württemberg: a bottom-up economic assessment of the socio-technical potential

No. 8 Axel Ensslen, Alexandra-Gwyn Paetz, Sonja Babrowski, Patrick Jochem, Wolf Fichtner: On the road to an electric mobility mass market - How can early adopters be characterized?

No. 9 Kai Mainzer, Russell McKenna, Wolf Fichtner: Charakterisierung der verwendeten Modellansätze im Wettbewerb Energieeffiziente Stadt

No. 10 Hannes Schwarz, Valentin Bertsch, Wolf Fichtner: Two-stage stochastic, large-scale optimization of a decentralized energy system – a residential quarter as case study

No. 11 Leon Hofmann, Russell McKenna, Wolf Fichtner: Development of a multi-energy residential service demand model for evaluation of prosumers‘ effects on current and future residential load profiles for heat and electricity

No. 12 Russell McKenna, Erik Merkel, Wolf Fichtner: Energy autonomy in residential buildings: a techno-economic model-based analysis of the scale effects

No. 13 Johannes Schäuble, Silvia Balaban, Peter Krasselt, Patrick Jochem, Mahmut Özkan, Friederike Schnellhas-Mende, Wolf Fichtner, Thomas Leibfried, Oliver Raabe: Vergleichsstudie von Systemansätzen für das Schnelladen von Elektrofahrzeugen

Working Paper Series in Production and Energy

recent issues

No. 14	Marian Hayn, Valentin Bertsch, Anne Zander, Stefan Nickel, Wolf Fichtner: The impact of electricity tariffs on residential demand side flexibility
No. 15	Erik Merkel, Robert Kunze, Russel McKenna, Wolf Fichtner: Modellgestützte Bewertung des Kraft-Wärme-Kopplungsgesetzes 2016 anhand ausgewählter Anwendungsfälle in Wohngebäuden
No. 16	Russell McKenna, Valentin Bertsch, Kai Mainzer, Wolf Fichtner: Combining local preferences with multi-criteria decision analysis and linear optimisation to develop feasible energy concepts in small communities
No. 17	Tilman Apitzsch, Christian Klöffer, Patrick Jochem, Martin Doppelbauer, Wolf Fichtner: Metaheuristics for online drive train efficiency optimization in electric vehicles
No. 18	Felix Hübner, Georg von Grone, Frank Schultmann: Technologien zur Zerlegung und zur Dekontamination von kerntechnischen Anlagen
No. 19	Felix Hübner, Jennifer Jana Jung, Frank Schultmann: Gefahren ionisierender Strahlung für Mensch und Umwelt in Bezug auf kerntechnische Anlagen
No. 20	Juri Lüth, Tobias Jäger, Russell McKenna, Wolf Fichtner: Photovoltaik auf Gebäuden: eine GIS-gestützte Ermittlung des Potenzials in Baden-Württemberg
No. 21	Felix Hübner, Jennifer Jana Jung, Frank Schultmann: Auswirkungen nuklearer Unfälle auf den Menschen und die Umwelt
No. 22	Felix Hübner, Ulli Schellenbaum, Christian Stürck; Patrick Gerhards, Frank Schultmann: Evaluation von Schedulingproblemen für die Projektplanung von Großprojekten am Beispiel des kerntechnischen Rückbaus
No. 23	Martin Hain, Hans Schermeyer, Marliese Uhrig-Homburg, Wolf Fichtner: An Electricity Price Modeling Framework for Renewable-Dominant Markets
No. 24	Hannes Schwarz, Lars Kotthoff, Holger Hoos, Wolf Fichtner and Valentin Bertsch: Using automated algorithm configuration to improve the optimization of decentralized energy systems modeled as large-scale, two-stage stochastic programs
No. 25	Felix Hübner, Tobias Hünlich, Florian Frost, Rebekka Volk, Frank Schultmann Analyse des internationalen Marktes für den Rückbau kerntechnischer Anlagen Stand und Ausblick

Developing a municipality typology for modelling decentralised energy systems

Weinand*, J. M., McKenna, R., Fichtner, W.

Chair of Energy Economics, Institute for Industrial Production (IIP), Karlsruhe Institute for Technology (KIT), Hertzstraße 16, 76187 Karlsruhe, Germany

*Corresponding author: Email: jann.weinand@kit.edu, Tel.: 0721 6084 4444

Abstract:

The recent rapid expansion of renewable energy capacities in Germany has been dominated by decentralised wind, photovoltaic (PV) and bioenergy plants. The spatially disperse and partly unpredictable nature of these resources necessitates an increasing exploitation of integration measures such as curtailment, supply and demand side flexibilities, network strengthening and storage capacities. Indeed, one solution to the large-scale integration of renewable energies could be decentralised autonomous municipal energy systems. The achievement of grid parity for some renewable energy technologies has strengthened the desire of some communities to become independent from central markets. Whilst many communities in Germany already strive for so-called energy autonomy, the vast majority do so only on an annual basis. Several studies have already analysed the technical and economic implications of the mainly decentralised future energy system, but most are restricted in their insights by limited temporal and spatial resolution.

The large number (11,131) of German municipalities means that a national analysis at this resolution is not feasible. Hence, this study employs a cluster analysis to develop a municipality typology in order to analyse the techno-economic suitability of these municipalities for autonomous energy systems. A total of 34 socio-technical indicators are employed at the municipal level, with a particular focus on the sectors of Private Households and Transport, and the potentials for decentralised renewable energies. The first step is to scale the indicator values and reduce their number by using a factor analysis. Several alternative methods are weighed against each other, and the most suitable methods for the factor analysis are chosen. Secondly, selected quantitative cluster validation methods are employed alongside qualitative criteria to determine the optimal number of clusters. This results in a total of ten clusters, which show a large variation as well as some overlap with respect to specific indicators. For example, one cluster contains all major German cities and has a low potential for renewable energies. Another cluster, on the other hand, contains the municipalities with a higher potential for renewable energies due to their high hydrothermal potential for geothermal power.

An analysis of the municipalities from three German renewable energy projects “Energy Municipalities”, ”Bioenergy Villages” and “100% Renewable Energy Regions” shows that in eight of the ten clusters municipalities are aiming for energy autonomy (in varying degrees). It is challenging to differentiate between the clusters regarding readiness for energy autonomy projects, however, especially if the degree of social acceptance and engagement for such projects is to be considered. To answer the more techno-economical part of this question, future work will employ the developed clusters in the context of an energy system optimisation. Insights gained at the municipal level will then be qualitatively transferred to the national context to assess the implications for the whole energy system.

1. Introduction

Ambitious national targets in energy policy are leading to a radical change in the energy industry, which is particularly marked by the expansion of renewable energies. Germany already generates 30% of electricity with renewable energy technologies in 2016 (Statistisches Bundesamt 2017a), including around 50 GW of wind (on- and offshore), about 7 GW of bioenergy and 40 GW photovoltaic (PV) plants (BMWi 2016a), of which around 98% are connected to the low voltage distribution networks (Wirth 2016). Regions are often referred to as the driving force behind the energy transition since renewable energies alongside energy efficiency are exploited on a decentralised basis due to their characteristics. Hence the characteristics of the energy system are changing towards a more decentralised structure, which also applies to the owners and operators of energy plants. In Germany, private individuals are increasingly investing in renewable energy systems or forming so-called citizen-energy cooperatives for this purpose. In fact, the majority of renewable plants in Germany are owned and operated by private individuals, farmers and communities (Klaus Novy Institut e.V. & trend:research 2011). This development is based on various socio-economic motives: among other things, citizens have the desire to play an active role in energy supply and to be more independent of central markets and structures (e.g. Boon & Dieperink 2014; Volz 2012).

In this context, the concept of municipal energy autonomy (Deutschle et al. 2015; Rae & Bradley 2012; McKenna et al. 2014b, 2015, 2017b) has become established, which is employed here to also include energy autarky (Müller et al. 2011), self-sufficiency (Deutschle et al. 2015; Balcombe et al. 2015) and integrated community energy systems (Koirala et al. 2016). Alone the number of terms for this concept illustrates the diversity within the literature, which also extends to its definition. Three rough distinctions can be made between complete energy autonomy (i.e. off-grid), net or balanced energy autonomy, whereby local generation equals or exceeds demand on an annual basis, and a tendency towards higher energy autonomy through decentralised renewables (McKenna et al. 2015). The extensive survey of Engelken et al. (2016) shows that the overwhelming number of municipalities with energy autonomy aspirations strive for the state of balanced energy autonomy and that the focus is usually on electrical energy.

The feasibility of municipal energy autonomy has been investigated in several case studies. In Scheffer (2008), a rural model region with 10,000 inhabitants and agriculture as well as trade and commerce, but without large-scale industry, is considered. The suitability of a rural settlement structure for energy autonomy is also investigated in Peter (2013), who shows that renewable energies could cover the electricity requirements of the “example village” with 3,850 inhabitants, but with immense storage costs. Jenssen et al. (2014) conclude that the complete energy autonomy in an “average” German municipality is technically attainable through the “bioenergy village” approach, albeit at high costs. Schmidt et al. (2012) examine the advantages and disadvantages of energy autonomy compared to conventional energy supply in Sauwald, Austria. Woyke & Forero (2014) evaluate complete energy autonomy in Pellworm, a municipality with 1,100 inhabitants, which has already been a model location for the construction of renewable energies. Although the supply of energy exceeds the demand, a complete energy autonomy is not possible with the current energy system in Pellworm due to grid constraints. Finally, the study by Burgess et al. (2012) examines the Marston Vale region in the UK, which would have to import heat energy and fuel in particular, while a large proportion of the demand for electricity could be met by energy supplied by the region itself.

Despite some general conclusions from these studies, such as a tendency to focus on balanced energy autonomy and electricity in more rural municipalities, there is until now no general framework within which to assess the feasibility of energy autonomy for a specific municipality. In addition, the high spatial and temporal resolutions required to satisfactorily model decentralized energy systems with large fractions of renewable energies makes approaches to information reduction indispensable. This paper goes some way towards filling these gaps by developing a typology of Germany’s 11,131 municipalities to support the selection of municipalities for future decentralised energy autonomy projects. With the help of a cluster analysis, these municipalities are divided into homogeneous clusters by socio-energetic indicators. The objective is to identify municipalities where energy autonomy aspirations could make technical and economic sense, and thereby to support the transferal of successful projects to other municipalities within the same cluster. In addition, a foundation for energy system models is developed which enables large-scale modelling of decentralized energy systems without the requirement for high spatial resolutions, which is often a central limitation in such models at the national scale and above (Keles et al. 2017). Finally, representatives of municipalities can be encouraged to initiate energy autonomy projects themselves if they have already been successfully implemented in a similar municipality.

The paper is structured as follows. Section 2 presents a literature review and more clearly locates this paper in context. Section 3 then presents the methodology, before section 4 presents and section 5 discusses the results. The paper closes with a summary and conclusions in section 6.

2. Literature review

Several areas of energy research are relevant to this contribution, including those relating to the analysis of decentralised and centralised energy systems, the field of urban morphology, and the application of cluster analysis to energy systems in order to reduce information quantity whilst retaining quality.

Characterising and contrasting centralised and decentralised energy systems is a relevant area of research for this paper because it strongly relates to the suitability of decentralised energy systems to become energy autonomous. Examples of contributions in this area include Funcke & Bauknecht (2016), who develop typologies for both of these types of energy infrastructure, by focussing on infrastructure location and operation. Further, Schmid et al. (2016) analyse the actor types, motives and conceivable roles within today's centralised and tomorrow's decentralised energy systems from the perspectives of technology, actors and institutions. Others raise the question of the optimal "degree of centralisation" (Zentralisierungsgrad), first coined by Jensch (1989), i. e. the level at which decentralised energy systems should be aggregated and balanced (Bauknecht et al. 2015). Currently, most energy autonomous regions rely on the overarching centralised energy system for their flexibility and controllability (Funcke & Bauknecht 2016). For example, Wimmer et al. (2014) compare centralised with decentralised wind expansion scenarios, concluding that the overall flexibility requirements are similar in both cases. Reiner Lemoine Institut (2013) finds that a decentralised renewables expansion would be economically favourable, largely due to higher required network expansion costs in the centralised case. Others reach the opposite conclusion, however, that centralised and hybrid energy systems are more economically efficient than purely decentralised ones (acatech 2016). Although it is clear that a completely renewable energy supply based on decentralised, autonomous regions does not seem economical due to very large storage requirements (Peter 2013), there is no clear consensus about the optimal degree of centralisation. Especially the related question of the technical feasibility of decentralised energy autonomy is addressed in this paper whilst the micro- and macroeconomic assessment is left to future work.

Urban morphology is the second relevant research area. It focusses on the form of the urban environment, including building types, ages and forms, and (amongst other things) its implications for the energy system. The field is well established, as demonstrated by the earlier contribution of Steemers (2003), who analysed the relationship between urban morphology and energy use in buildings and transport, the two main sectors (other than industry) that are relevant for urban planning. Also Ratti et al. (2005) explored the effects of urban textures on building energy consumption with digital elevation models, with case studies in three European cities. Similar methods were also more recently employed in the LSECities project (Rode et al. 2014a, 2014b), which analysed the effects of different types of urban forms on heat energy demand and derived generalised insights into these relationships in larger European cities. In the context of her PhD thesis, Miller (2013) approaches the connection between urban form and building energy use with a multi-scale approach and using the Metro Vancouver region in Canada as an example. All of these studies demonstrate the diversity amongst the urban building stock, leading to a substantial variation in heat demand. Others within this field have examined the relationship between solar energy potential and urban morphology in London, concluding that by optimising combinations of eight variables of urban form the solar irradiation of roofs and facades could be increased by around 9% and 45% (Sarralde et al. 2015). More recently, Urquizo et al. (2017) explored different urban morphology metrics and their impact on energy consumption in four districts of Newcastle, UK. In a more detailed analysis, Hargreaves et al. (2017) investigate the most cost-effective decarbonisation options for regions with different urban forms in a UK context, showing for example how low-density urban areas are more suited to exploit ground-source heat pumps. Summarizing, then, the field of urban morphology offers insights into the connection between energy demand and urban structures, but does not provide a transferable typology for the whole decentralised energy system.

The third and most relevant research field for this paper is that of cluster analysis. Despite being a common method in energy studies more widely, it has not yet been often employed in the analysis of decentralised energy systems. One example is Chévez et al. (2017), who examine the single region "Great La Plata" in Argentina at the administrative level. The region is clustered into eight census area types with a k-means cluster analysis according to the consumption of electrical energy and other socioeconomic variables. The most important result is that electricity consumption increases strongly with the household sizes, which could, for example, support the construction of distribution networks. In addition, Unternährer et al. (2017) cluster 6224 buildings not yet connected to the local heating network at the administrative level. Depending on indicators such as the demand for space heating and domestic hot water, as well as georeferenced drilling costs for deep geothermal energy, the cluster analysis results in 16 clusters. Clusters and typologies have often been applied at the district scale, in identifying the most cost-effective low carbon energy solution for different types of districts (Hargreaves et al. 2017; McKenna et al. 2016, 2017a; Su et al. 2017), as well as at the building scale, for example in the context of residential heat demand studies (McKenna et al. 2016, 2017a). In addition, Marquant et al. (2017) present a holistic approach for optimisation of multi-scale distributed energy systems, by employing clusters of similar buildings at the district level.

There are some examples of applications of cluster analysis at higher levels of spatial aggregation. For example, Kaundinya et al. (2013) employ a k-medoid clustering method to divide a region in India into clusters of villages for supply with decentralised biomass power plants, and the value of k is chosen to minimise the total system costs. For Austria, Bramreiter et al. (2016) divide all of the 82 Austrian "Climate and Energy Model Regions (CEMs)", which aim for energy autonomy, into three clusters by ten indicators (e. g. population density, employment figures, energy consumption). In a subsequent step, all other Austrian municipalities are examined by cluster analysis, with the aim of identifying municipalities with characteristics similar to those of the CEMs. It is shown that large parts of Austria could also become CEMs and thus have the potential to become energy self-sufficient, at least on an annual basis. In the study of Requia et al. (2017), all 5570 municipalities in Brazil are divided into five clusters. However, the analysis does not focus on socio-energetic indicators, but on six types of pollutant emissions in the Transport sector such as CO_2 and NO_x. To transfer the results to energy systems of municipalities, indicators for the other consumption sectors Private Households, Industry and Commercial would have to be included in the cluster analysis. The investigations are not always limited to one country. For example, Noiva et al. (2016) investigate 142 cities, spread across all continents. Indicators for the analysis of the cities divided into six clusters are the parameters of supply and demand for water.

Another relevant example in the present case is the PhD dissertation of Wall (2016), who conducted a cluster analysis with the German county-free cities as objects and based on 41 socio-energetic indicators. The cluster analysis in Wall (2016) differs from this study not only in the choice of indicators but also in the choice of the research objects. The survey objects are not the municipalities, but only the 107 county-free cities in Germany. Other studies have employed cluster analysis to German regions, but most of these neither have a high spatial resolution nor focus on energy aspects. For example, in Kronthaler (2003) Germany was divided into 97 regions, which were then assigned to ten clusters in a cluster analysis. The study looked at 13 socioeconomic indicators, including employment figures and investment in industry. The research showed that the economic power of the regions in eastern Germany is still significantly lower than that of the western German regions. Heinritz (2000) also came to a similar conclusion, by evaluating the economic strength of the 441 counties in Germany, and dividing the counties into five clusters by socio-economic indicators such as gross domestic product per inhabitant. In three other studies, German municipalities are investigated, but none of the studies considers all 11,131 municipalities. Geyler et al. (2008) only analyse 240 municipalities in the core region of Central Germany. The delimitation into six clusters is based on local development trends. These include 16 indicators such as the development of the settlement and traffic area or business tax revenue per inhabitant. In Schultz & Brandt (2016) 2,916 of the 11,131 German municipalities are divided into nine clusters by demographic indicators (among other things, the “share of single-person households” or “share of under-18s”). Finally, the 1102 municipalities in the federal state of Baden-Württemberg are investigated in Statistisches Landesamt Baden-Württemberg (2009). The goal was not to place the municipalities in clusters but to identify the two municipalities that are closest to each other. Indicators such as population density or cars per 1,000 inhabitants were used. Hence, although several German regions have been analysed with cluster analysis, a classification with energy indicators has not yet been carried out at the municipal level. This is the research gap addressed in this paper, as outlined in the following section.

3. Methodology

This section describes the data collection and standardisation (cf. section 3.1) as well as the execution of the factor analysis (cf. section 3.2) and cluster analysis (cf. section 3.3). The vast majority of cluster analyses evaluated in section 2 perform a hierarchically agglomerative cluster analysis with the Ward algorithm. 17 of the 23 analyses evaluated in Wall (2016) also apply hierarchically agglomerative cluster analysis. Hierarchical cluster analysis generates high-quality clusters therefore this is also used in this paper. To support the traceability of the cluster analysis, the most important information according to Bacher et al. (2010) is listed in Table 1.

Table 1: Overview of the most important aspects of traceability of a cluster analysis.

Objects	11,131 German municipalities
Variables/Indicators	59 Indicators (see section 3.1)
Algorithm	Ward
Cluster analysis method	Hierarchical-agglomerative, k-means
Criteria used to determine the number of clusters	26 different methods and elbow criteria (see section 3.3.2)
Software used	R

3.1. Data collection and standardisation

Many of the indicators used in the studies mentioned in section 2 are also used in the cluster analysis presented in this paper, as well as newly selected indicators. This study uses the indicators in a comprehensive analysis and for the first time clusters all 11,131 municipalities in Germany. The 59 indicators used in the cluster analysis include data on the energy consumption sectors “Private Households”, ”Transport”, ”Industry” and “Commercial” as well as data to estimate the potential for renewable energies (see Table 2). The indicators whose data is only available at the county level are shown in italics in Table 2. In the following, the “X” values in brackets are used as abbreviations for the indicators. For the last three groups of indicators in the Private Household sector, the specific allocations of the “X” values will be described later in the text. Only the indicators used in the final analysis are assigned to “X” values. The question of why not all indicators are used is answered in section 3.1. A complete list of all indicators and their references is given in Table 8 in the Appendix.

Table 2: Overview of the indicators used in the cluster analysis. Italics means that the data of the indicators were only available at the county level.

Consumption sector Private Households (29)	**Consumption sector Transport (11)**	**Consumption sector Industry and Commercial (12)**	**Potential for renewable energies (7)**
Population development between 2010 and 2015 (X1) [%]	Number of motor vehicles per 1,000 inhabitants (X27)	*Share of employment in the industrial sector* [%]	Achievable hydrothermal temperature (X32) [°C]
Living space per person (X2) [m²]	Number of cars per 1,000 inhabitants (X28)	*Share of employment in the commercial sector* [%]	Necessary hydrothermal drilling depth (X33) [m]
Share of single-person households (X3) [%]	*Share of diesel vehicles* [%]	*Energy productivity of manufacturing industry* [€/GJ]	Technical PV potential per inhabitant (X34) [kWh/y]
Average household size (X4) [Persons]	*Share of petrol vehicles* [%]	*Energy intensity of manufacturing industry* [MJ/€]	Technical PV potential per km² (X35) [MWh/y]
Household density (X5) [Housholds per km²]	*Share of gas vehicles* [%]	*Productivity level of manufacturing industry* [€/GJ]	Technical wind potential per inhabitant (X36) [MWh/y]
Share of owner-occupied apartments (X6) [%]	*Share of hybrid vehicles* [%]	*Specific energy consumption of manufacturing industry* [MJ/€]	Technical wind potential per km² (X37) [MWh/y]
Income per household (X7) [k€]	*Share of electric vehicles* [%]	*Share of industrial sales tax* [%]	Share of forest and agricultural land (X38) [%]
Share of over 65-year-olds (X8) [%]	*Share of other vehicle types* [%]	*Share of commercial sales tax* [%]	
Unemployment rate (X9) [%]	Population density (X29) [Inhabitants per km²]	*Development of employment share in the industrial sector* [%]	
Share of settlement and traffic area (X10) [%]	Share of 18-64-year-olds (X30) [%]	*Development of employment share in the commercial sector* [%]	
Heating days	Share of commuters in the workforce [%]	*Development of energy intensity in the manufacturing sector* [%]	
Heating degree days		Number of manufacturing enterprises per 1,000 households (X31)	
Degree day number			
Share of heating types (3 indicators) (X11-X13) [%]			
Share of building age class (9 indicators) (X14-X22) [%]			
Share of building type (4 indicators) (X23-X26) [%]			

27 of the 59 indicators also used Wall (2016) in his analysis of county-free cities. In the following, the reasons for selecting the additional indicators are explained.

3.1.1. Indicators of the consumption sector Private Households

Private households account for 26% of Germany's final energy consumption and should not be neglected in the energetic classification of municipalities. The majority of the final energy (69%) is used in households for space heating (Umweltbundesamt & BMWi 2017).

Share of heating types

For the shares of heating types, the available data have been grouped into three groups:

1) Share of buildings with heating systems based on district heating (X11)
2) Share of buildings with heating systems not based on district heating (X12)
3) Share of buildings without heating system (X13)

This segmentation allows conclusions to be drawn as to whether and to what extent there is a district heating network, a gas network or both in the municipalities. The existing infrastructures influence the selection decision of technologies which are suitable in the municipalities. As an example, power-to-heat plants and power-to-gas plants offer great opportunities for future flexibility in power generation. However, to store the energy from these plants, various networks are required, such as a district heating network for power-to-heat plants or a gas network for power-to-gas plants (Böttger et al. 2014). Furthermore, district heating systems are suitable for the integration of heat from renewable energies such as geothermal power plants (Durst 2015).

Shares of building age classes

The insulation condition of the building envelope has a significant influence on the space heating requirement in buildings (Braun 2010). The building age class influences the insulation condition of the building envelope and is, therefore, an essential indicator for estimating the heat demand (Schuler et al. 2000). In the cluster analysis applied here, the building ages were divided into nine groups (see Table 8 in the Appendix).

Shares of building types

The type of building also has a significant influence on the demand for space heating in private households (Wei et al. 2014). Shipworth et al. (2010), for example, showed that the operating hours of the heating system in English homes are statistically dependent on the type of building. The biggest difference was found between detached houses, in which the heating is much longer, and terraced houses. This study distinguishes between detached houses (X23), semi-detached houses (X24), terraced houses (X25) and “other types of buildings” (X26).

3.1.2. Indicators of the consumption sector Transport

For the indicators representing the Transport sector, the shares of hybrid, gas and other vehicles in the vehicle stock have been added (compared to Wall 2016). Hybrid vehicles also include an internal combustion engine in addition to the electric motor. The combustion engine can compensate for the disadvantage of the limited range of electric vehicles (Høyer 2008). The number of gas vehicles in Germany is around 100,000, and they can contribute to a significant reduction in pollutants and, in some cases, CO_2 emissions. If biomethane or synthetic methane is added to the fuel, gas vehicles can be as climate-friendly as electric vehicles (BMWi 2016a).

3.1.3. Indicators of the consumption sector Industry and Commercial

The Industry consumption sector accounts for almost 50% of the electricity supplied in Germany (Javied et al. 2016). Most of the data from the Industry consumption sector are only available for the manufacturing sector. These data are suitable for estimating the energy consumption of industry, as manufacturing accounts for the largest share of energy consumption (27.4% of Germany's total primary energy demand) (Umweltbundesamt 2016).

Number of manufacturing enterprises per 1,000 households

The number of manufacturing enterprises is the only indicator of this sector provided at the municipal level. The indicator is based on 1,000 households to compare the values for the different municipalities.

3.1.4. Indicators of the potential for renewable energies

In most practical examples of municipal energy autonomy, renewable energies are used to establish a sustainable energy system (Schmidt et al. 2012). Therefore, the potentials of renewable energies in a region are important indicators. The potentials of renewable energies applied in the cluster analysis are explained below.

Achievable hydrothermal temperature

In Germany, an increase in deep geothermal power stations is expected by 2030 (installed capacity in 2030: 850 MW_{el}) (Hechler & Bredel-Schürmann 2011). From 2003 to 2013, the annual supply of thermal energy by deep geothermal energy plants has increased from 60 GWh_{th} to 530 GWh_{th}, the supply of electrical energy has increased from 0 GWh_{el} to 36 GWh_{el} (Agemar et al. 2014). In this study, the focus is on hydrothermal systems, because petrothermal systems are not yet used in Germany (Hechler & Bredel-Schürmann 2011). Electricity from geothermal energy currently receives a subsidy of 25.2 €-cents/kWh (Deutscher Bundestag 2017).

Hydrothermal power plants have two main advantages: on the one hand, unlike many other renewable energy plants, they are capable of providing energy as base load. On the other hand, they show the lowest emissions of pollutants after hydroelectric power plants during the life cycle of the plant (Purkus & Barth 2011). At the municipal level, several geothermal power plants are already being used to supply local and district heating (Hechler & Bredel-Schürmann 2011). Therefore, the use of this technology should also be considered in future energy autonomy efforts.

An important indicator for estimating the economic potential of a geothermal plant is the achievable hydrothermal temperature. Hydrothermal temperatures above 110 °C are required for the economical operation of a geothermal plant to generate electricity (Agemar et al. 2014). Figure 1 shows that the achievable hydrothermal temperatures strongly depend on the region. This means that municipalities have different hydrothermal potentials. Therefore, the indicators for hydrothermal energy are included in the cluster analysis.

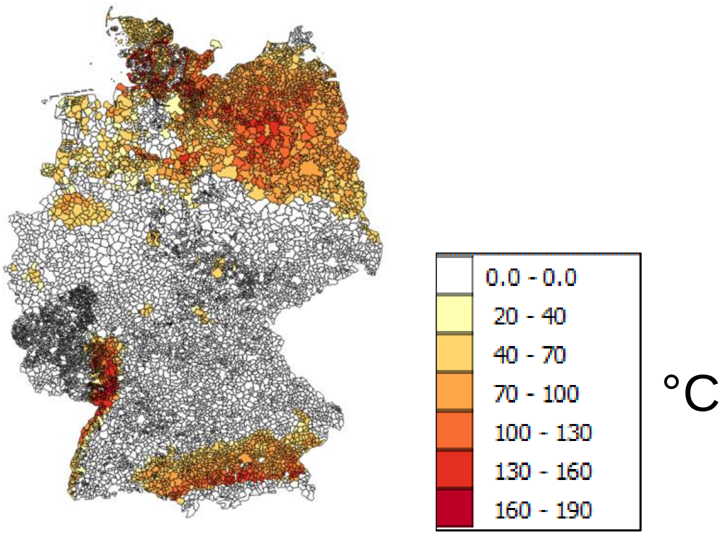


Figure 1: Achievable average hydrothermal temperature (°C) at a depth of up to 5000 meters in German municipalities according to Agemar (2017).

Necessary hydrothermal drilling depth

The depth of drilling to the water reservoirs mainly determines the amount of investment for a geothermal plant. The depth of the well depends on the local temperature gradient. In Germany, the average temperature gradient is 32 K/km, in some regions (Upper Rhine valley) up to 100 K/km are reached (Agemar et al. 2014).

Technical photovoltaic (PV) potential per inhabitant/per km²

The data from Mainzer et al. (2014) were used to estimate the PV potential in municipalities. However, this data must be made comparable for the cluster analysis. The indicator is therefore determined by dividing the PV potential in kWh by the number of inhabitants. This means that the technical PV potential per inhabitant can now be used for each municipality. However, this does not yet complete the estimation of the PV potential, as this indicator does not allow a statement to be made about the potential of the PV systems in the relation to the area. For this purpose, the energy density of the PV systems is determined by dividing the PV potential in MWh by the area in km².

Technical wind potential per inhabitant/per km²

The indicators for the technical wind potential in MWh were determined in analogy to the technical PV potential per inhabitant and per km². The data from McKenna et al. (2014a) was used for this. This data is available at postcode level and could be assigned to the municipalities using the geoinformation system QGIS.

Share of forest and agricultural land in total area

Land areas are required for the construction and operation of many technologies based on renewable energies such as wind power plants, ground-mounted photovoltaics and biogas plants (Marx Gómez et al. 2014). While wind power plants may only be built in certain areas such as forests or agricultural land, biogas plants require entire areas for the cultivation of energy crops (mainly maize) (Lüker-Jans et al. 2017; McKenna et al. 2014a). The proportion of woodland and agricultural land in the total area of the municipalities can, therefore, be used as an indicator to estimate the potential for these renewable energies.

3.1.5. Data in the investigation

In the final study, only 38 of the 59 indicators were used. The indicators used are marked in Table 2 with a compounded X-value. The first cluster analysis with all indicators showed that there was too much dependence on the indicators based on county data for them to accurately represent values for municipalities. For this reason, the indicators shown in italics in Table 2 were excluded from further analysis. In addition, the indicator "Share of commuters in employment" could not be included in the study, as the data only exist for 2014 and are incomplete.

Furthermore, one indicator for each of the proportions of heating types, building age classes and building types were eliminated for a reason described below. Many calculation steps of a factor and cluster analysis require a positive semi-definite data matrix (Lorenzo-Seva & Ferrando 2006). A symmetrical matrix is positive semi-definite if all eigenvalues are nonnegative (Zhang 2011). In this study, the matrices were not positive semi-definite. This problem was solved by eliminating linear dependent variables. Since the proportions of the indicator groups heating types, building age classes and building types can be added up to 100% in each case, an indicator value can always be calculated with the other indicator values. Therefore, X13, X22 and X26 are deleted from the dataset. More information about positive semi-definite matrices can be found in Zhang (2011).

3.1.6. Standardisation

Once the data is complete, it must be standardised for factor analysis. Standardisation serves to make the indicators comparable in their range of values (Milligan & Cooper 1988). This prevents indicators with larger values from being weighted more strongly. Many studies use the Z-transformation to standardise the data. The standardised values Z are calculated using the original indicator value X, the arithmetic mean $\bar{X}$ and the standard deviation s (Heyde 1990):

$$Z = (X - \bar{X})/s \tag{1}$$

However, Milligan & Cooper (1988) showed that this traditional Z-value method leads to poorer results in cluster analyses than other standardisation methods. In most cases, the Z-value method works well only with normally distributed data (Office for National Statistics 2015). The following calculation has proved to be the best method, which is also used in the present study (Milligan & Cooper 1988):

$$Z = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{2}$$

3.2. Exploratory Factor Analysis

An exploratory factor analysis serves to examine the data and reduce the number of (required) indicators. The j-th factor F_j can be determined using the k indicators X_1, X_2,..., X_k and the weights or factor loadings W_{ji}:

$$F_j = W_{j1}X_1 + W_{j2}X_2 + \cdots + W_{jk}X_k \tag{3}$$

The larger the factor load W_{ji}, the stronger the value of the factor F_j is determined by the indicator X_i (Aljandali 2017). The R-function “fa” from the package “psych” is used here for factor analysis (Revelle 2017). Factor analysis was conducted following the steps proposed in Osborne (2014), which are explained below.

3.2.1. Selection of the extraction method

An extraction method is used to investigate the correlation between all indicators with the aim of extracting the latent variables. A latent variable, here a factor, is a variable that cannot be measured directly but is the basis of the observed variables. If the data is predominantly normally distributed, then the maximum likelihood method is best suited as an extraction method, if it is not normally distributed, the principal axis factor method should be used (Osborne 2014). Figure 7 in the Appendix shows the distributions of the standardised indicators. The data were checked for normal distribution with the Kolmogorov-Smirnov test since the Shapiro-Wilk test is only suitable for data records with up to 5000 datasets (Shapiro & Wilk 1965; Lopes 2011). The p-values were smaller than $2.2 * 10^{-16}$, so all data series are not normally distributed. Therefore, the principal axis factor method seems to be suitable as an extraction method.

However, in the factor analysis using the principal axis factor method, so-called Heywood cases occurred. A Heywood case occurs when variances are negative, or correlations (in this case some factor loadings) are greater than one. Due to the Heywood cases, the solution of the factor analysis is inadmissible. In addition, the causes of Heywood cases are difficult to distinguish (Dillon et al. 1987). With the recommended extraction method in Revelle (2017), the "minimum residual" method, almost the same result was obtained as with the principal axis factor method, since only one indicator was assigned to a different factor. However, Heywood cases also occurred when using this extraction method. The Heywood cases were not discussed in Osborne (2014), so no other method was recommended for this case. In Revelle (2017), it is pointed out that in contrast to other methods, the "Minimum Rank Factor Analysis" (MRFA) does not include Heywood cases. Therefore, the MRFA method is selected below as the extraction method. The MRFA method is described in Lorenzo-Seva & Ferrando (2006) as the only method that calculates the part of the variance explained by each factor. This is also the only difference between this extraction method and the "minimum residual" method (Shapiro & Berge 2002).

3.2.2. Selection of the number of factors

In his study, Osborne (2014) points out that no criterion for selecting the number of factors is better than another, the suitability of the criteria varies depending on the case. Therefore, several methods should be used. In this paper, the Kaiser criterion from Kaiser (1960) combined with a "Scree-Plot" and the "Parallel Analysis" from Horn (1965) are applied. Ten factors are recommended with the Kaiser method, and nine with the Parallel Analysis (cf. Figure 8 in the Appendix). In the following, ten factors are assumed according to the Kaiser criterion (cf. curve "Eigenvalues > 0" in Figure 8 in the Appendix).

3.2.3. Selection of the rotation method

The rotation was invented shortly after the factor analysis to facilitate the interpretation of the results of the factor analysis (Osborne 2014). The goal is a simple structure in which each indicator describes as few factors as possible (or "loads onto them"). In addition, rotation creates groups of factors containing related indicators (Yong & Pearce 2013). This analysis uses the "Varimax" method, which is widely used in practice, to maximise the variance of factor loadings and minimise the number of factors (Eckstein 2016).

3.2.4. Results of the Factor Analysis

Table 3 shows the allocation of the indicators to the factors resulting from the factor analysis with the extraction method MRFA. The indicator X31 is the only one of the indicators not described by the factors, as its factor loading is very low for each factor. This means that X31 is no longer included in the further analysis. Figure 9 in the Appendix shows the size of the factor loadings of all remaining indicators for each factor. The results can be assessed as plausible since each factor describes a specific issue (see column “Factor name” in Table 3). Figure 10 in the Appendix also shows a correlation diagram of the indicator values. As an example, a high correlation between X29, X5 and X10 is shown there. These indicators are therefore all assigned to Factor 1 (see Table 3).

Table 3: Assignment of the indicators with their factor loadings to the ten factors and naming of the factors.

Factor	Indicators	Abbre-vations	Factor loading	Factor name
1	1) Household density 2) Share of settlement and transport area 3) Population density 4) Technical PV potential per km² 5) Share of forest and agricultural area	X5 X10 X29 X35 X38	0.917 0.918 0.934 0.921 -0.768	Area factor (all indicators refer to the area of the municipality)
2	1) Income per household 2) Unemployment rate 3) Share of buildings built before 1919 4) Share of buildings built between 1919 and 1949 5) Share of buildings built between 1960 and 1969 6) Share of buildings built between 1970 and 1979 7) Share of buildings built between 1980 and 1989	X7 X9 X14 X15 X17 X18 X19	0.464 -0.503 -0.629 -0.791 0.587 0.791 0.560	East/West Factor (this factor reflects the inequalities between West and East Germany)
3	1) Achievable hydrothermal temperature 2) Necessary hydrothermal drilling depth	X32 X33	0.949 0.937	Hydrothermal factor
4	1) Number of motor vehicles per 1,000 inhabitants 2) Number of cars per 1,000 inhabitants	X27 X28	0.857 0.882	Traffic factor
5	1) Share of over 65-year-olds 2) Share of buildings built between 1990 and 1999 3) Share of buildings built between 2000 and 2005 4) Share of 18-64-year-olds	X8 X20 X21 X30	-0.726 0.737 0.529 0.626	Age factor
6	1) Share of buildings with heating systems based on district heating 2) Share of buildings with heating systems not based on district heating	X11 X12	-0.939 0.934	Heating system factor
7	1) Population development between 2010 and 2015 2) Living space per person 3) Average household size 4) Technical PV potential per person	X1 X2 X4 X34	0.459 -0.660 0.811 -0.555	Population Factor (all indicators depend on population size)
8	1) Share of buildings built between 1950 and 1959	X16	0.891	-
9	1) Share of single-person households 2) Share of owner-occupied apartments 3) Share of detached houses 4) Share of semi-detached houses 5) Share of terraced houses	X3 X6 X23 X24 X25	0.436 -0.550 -0.863 0.663 0.719	Building factor
10	1) Technical wind potential per inhabitant 2) Technical wind potential per km²	X36 X37	0.823 0.754	Wind factor

3.3. Cluster analysis

As already described above, high quality clusters are generated with hierarchical agglomerative cluster analysis. However, this method requires high computing times (Bouguettaya et al. 2015). In the cluster analysis carried out here, the high computing times were mainly due to the complex determination of the number of clusters. Similar to Wall (2016), the results of the factor analysis were used as input for the cluster analysis.

3.3.1. Ward algorithm

The clusters can be classified using distance metrics. To determine the distance matrix, the distance or similarity between all objects is determined (Johnson 1967). The Ward algorithm is the only method among the agglomerative cluster methods that is based on the classical sum of squares and determines groups, minimising dispersion within the groups at each step. The sum of the squares is determined with the help of the distance matrix (Murtagh & Legendre 2014). In this study, the distance matrix is calculated using the Euclidean distance, since it should be the basis for the Ward method (Miyamoto et al. 2015).

To use the Ward algorithm, the R function “hclust” has been executed (Müllner 2016). Within this function, two different algorithms Ward1 or Ward2 can be selected. Murtagh & Legendre (2014) showed that only the algorithm Ward2 minimises the Ward criterion and should, therefore, be used. For more information about the mathematical differences of Ward1 and Ward2, the authors refer to Murtagh & Legendre (2014). The difference d^2 of two clusters R and Q is calculated with the help of the cluster foci $\bar{x}$ using the following formula (Gentle et al. 1991):

$$d^2(R,Q) = \frac{2|R|\,|Q|}{|R|+|Q|}\|\bar{x}(R) - \bar{x}(Q)\|^2 \qquad (4)$$

3.3.2. Determining the number of clusters

In hierarchical agglomerative cluster analysis, the number of clusters is not known in advance but must be determined using suitable methods (Salvador & Chan 2004). The more clusters selected, the more similar the objects within the clusters are. At the same time, the clusters are more difficult to distinguish between each other as the number of clusters increases.

In some studies such as Wall (2016) or Yang et al. (2017), the number of clusters is estimated using the common but often inaccurate “elbow” method. Alternatively, the R-function “NbClust” from Charrad et al. (2014) offers 30 methods for determining the optimal number of clusters. None of the criteria studied so far can predict the optimal number of clusters in any case (Albatineh & Niewiadomska-Bugaj 2011). Therefore, all 30 methods were implemented. More information about the mathematical description of the methods can be found in Charrad et al. (2014). The results of the procedures in the context of this study are shown in Table 9 in the Appendix. Only 26 of the 30 methods are listed in the table since the computationally intensive methods such as "gamma" had to be aborted after almost two months of computing time. As can be seen in Table 9 in the Appendix, the 26 methods yielded quite different values for the number of clusters. It is therefore necessary to examine more closely whether the methods should be used at all in this particular case. 22 of the 30 procedures are already explained and evaluated in Milligan & Cooper (1985). For Example, the “ch” procedure of Calinski & Harabasz (1974) was rated as the best procedure. However, Islam et al. (2016), showed that “ch” is poor with a high number of clusters and usually prefers - as in this study - a 2 cluster solution. Almost all 26 methods have poor functionality with a high number of clusters (cf. Table 9 in the Appendix). The only algorithm for which a good functionality with high cluster numbers could be found in the relevant literature is “duda”, which suggests ten clusters in this study. However, it should be further examined whether the ten clusters represent the optimal number of clusters in this study.

Therefore, several cluster solutions with different numbers of clusters are compared to determine an appropriate number of clusters. Table 4 shows how the structure of the clusters changes from five clusters up to 15. For example, Cluster 1 from the 5 cluster solution divides into two further clusters at 14 cluster solution.

Table 4: Development of cluster composition for solutions with 5 to 15 clusters.

<table>
<tr><th>Number</th><th colspan="15">Cluster</th></tr>
<tr><td>5</td><td colspan="2" rowspan="9">339</td><td colspan="2" rowspan="8">727</td><td colspan="4" rowspan="2">2898</td><td colspan="5" rowspan="3">5722</td><td colspan="2">1445</td></tr>
<tr><td>6</td><td rowspan="10">1370</td><td rowspan="10">75</td></tr>
<tr><td>7</td><td colspan="2" rowspan="2">1671</td><td colspan="2" rowspan="3">1227</td></tr>
<tr><td>8</td><td colspan="4" rowspan="3">5262</td><td rowspan="8">460</td></tr>
<tr><td>9</td><td rowspan="7">1638</td><td rowspan="7">33</td></tr>
<tr><td>10</td><td rowspan="6">839</td><td rowspan="6">388</td></tr>
<tr><td>11</td><td colspan="2">3927</td><td colspan="2" rowspan="4">1335</td></tr>
<tr><td>12</td><td rowspan="4">1899</td><td rowspan="4">2028</td></tr>
<tr><td>13</td><td rowspan="3">181</td><td rowspan="3">546</td></tr>
<tr><td>14</td><td rowspan="2">11</td><td rowspan="2">328</td></tr>
<tr><td>15</td><td>726</td><td>609</td></tr>
</table>

Figure 2 shows the course of the within-cluster sum of squares as a function of the number of clusters. The within-cluster sum of squares describes the squared distance of an object to the cluster centre, i. e. how similar the object is to the other objects of the group (Anderson 2001). The smaller the within-cluster sum of squares is, the more similar the objects in the clusters are.

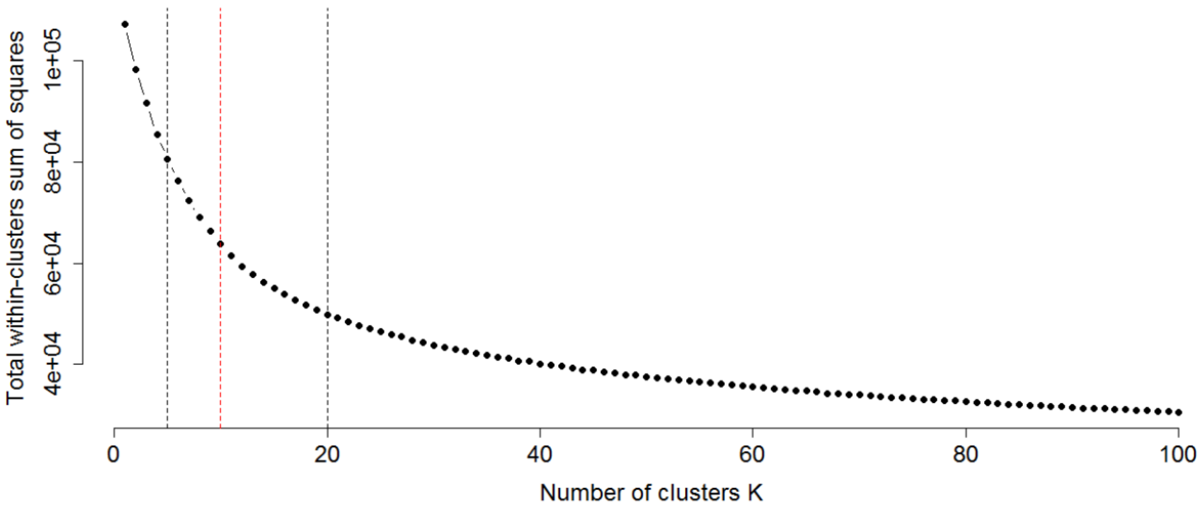


Figure 2: Sum of squares within clusters as a function of the number of clusters.

The aforementioned elbow method is based on the within-cluster sum of the squares, as shown in Figure 2, where the elbow represents the point of decreasing marginal returns. This means that right behind the elbow, with an increase in the number of clusters, the increase in information is very small. However, the region of the elbow is often not as clearly visible, as in Figure 2 (Kodinariya & Makwana 2013), so this method alone could not be used. The elbow method is after all only a heuristic one (Tibshirani et al. 2001). The elbow could be between five and 20 clusters in the area delimited by black dotted lines. The 10 cluster solution proposed by the “duda” method (see Table 9 in the Appendix), is also in this area (red dotted lines).

Therefore, the clusters need to be analysed further. It turned out that the new clusters formed in the 11 cluster solution differed significantly less from each other than the clusters formed in the previous steps. The upper diagram of Figure 3 illustrates the deviation in the mean values of all indicators for the two new clusters in the 11 cluster solution. The values have been scaled to values between 0 and 1 to improve the comparability. The two new clusters are clusters 5 with 3,927 municipalities and 8 with 1,335 municipalities, as the cluster numbers change in each step (cf. Table 4). The diagram shows that the mean values for each indicator are approximately the same.

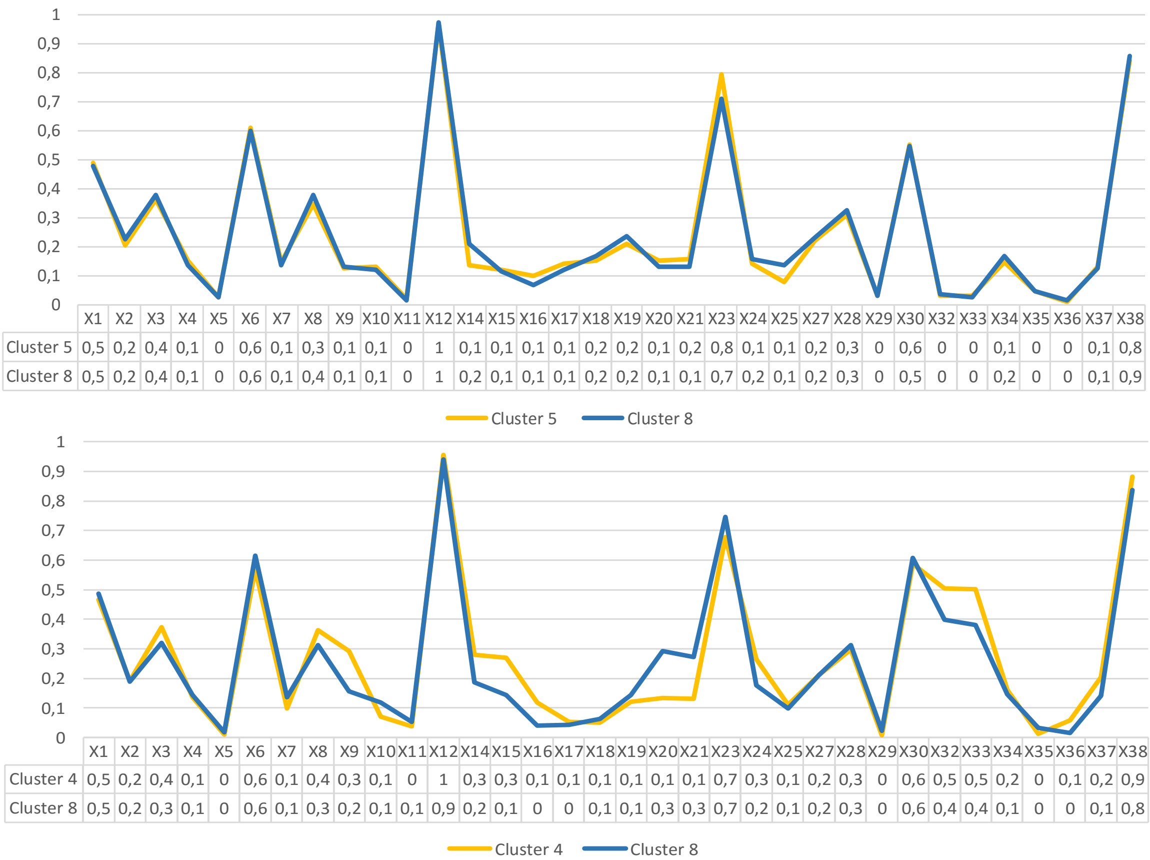


Figure 3: The mean values of the two newly formed clusters in the 11 cluster solution (upper figure) and the 10 cluster solution (bottom figure) over the 34 indicators.

This means that a further separation of the clusters from ten clusters onwards creates only a low added value. As a comparison, the curves of the mean values of the two newly created clusters in the 10 cluster solution are shown in the bottom diagram of Figure 3. In this case, the mean values vary significantly, so the number of clusters should be increased from nine to ten. In the following, ten clusters will be selected as the appropriate number of clusters, since this number can be justified by the “duda” method, the elbow method and further analysis.

4. Results of the cluster analysis

Figure 4 shows all German municipalities with a colour assignment to the clusters of the 10 cluster solution. The broader outlines separate the 16 federal states in Germany. Especially in Rhineland-Palatinate and Schleswig-Holstein, some municipalities seem to be dark to black. This is due to the small size of the municipalities; in Rhineland-Palatinate, the municipalities have by far the smallest size. Due to the poor visibility of these municipalities, the map is magnified in Figure 11 to Figure 13 in the Appendix.

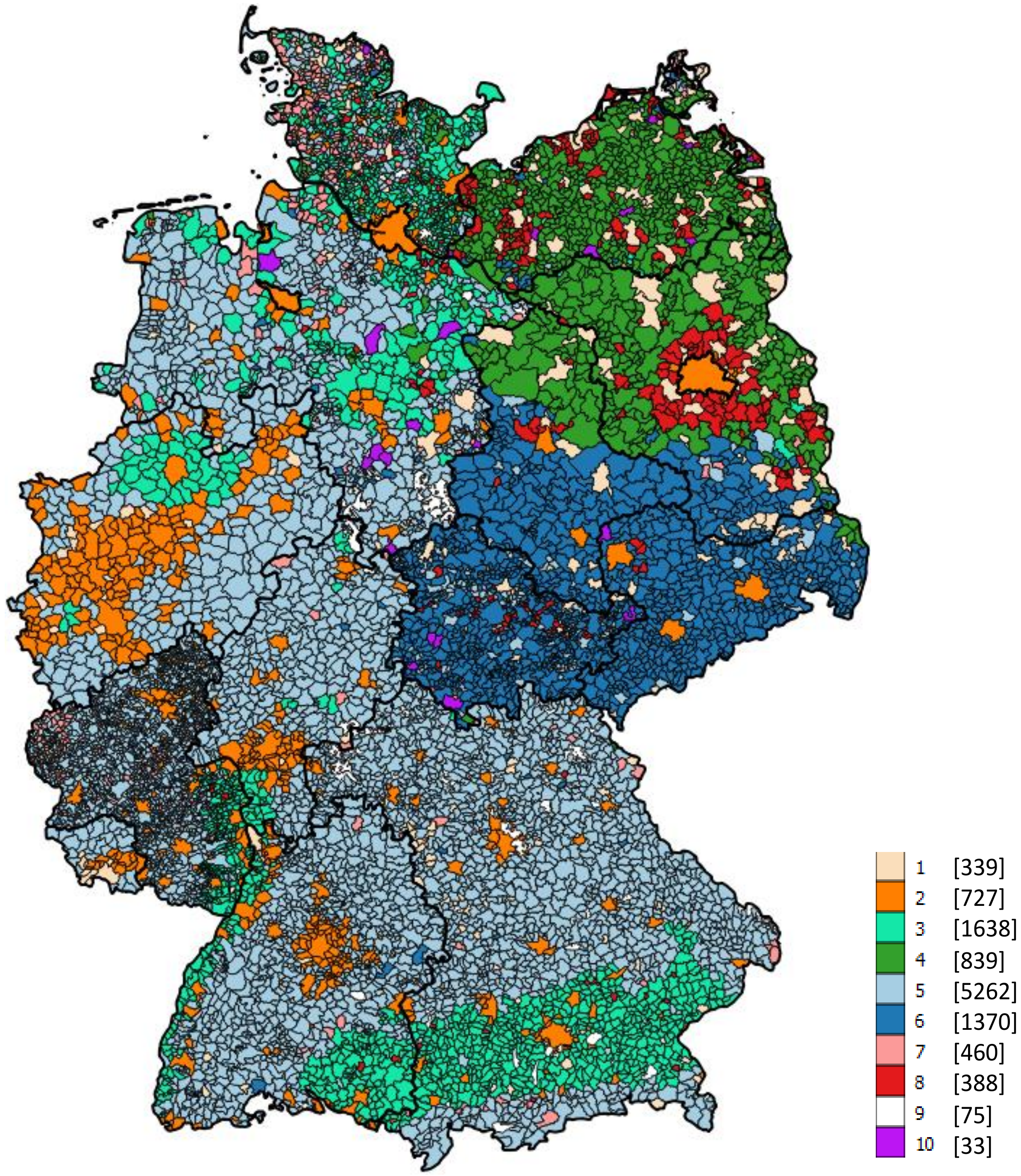


Figure 4: Illustration of all German municipalities with their allocation in the 10 cluster solution. The numbers of municipalities in the clusters are in parentheses.

The mean values of all 34 indicators were determined for all clusters and every single cluster (see Table 5). The different colours are chosen to distinguish between the sectors Private Households (blue/red) and Transport (yellow) as well as potential for renewable energies (green).

Table 5: Mean values of the indicators X1-X38 for the ten Clusters and all Clusters.

Indicator	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Mean value	-1,0	51,2	27,1	2,4	81,1	63,0	34,2	20,1	3,7	13,0	2,7	96,3
Cluster 1	-1,3	47,6	31,2	2,3	118,9	51,4	30,3	21,4	5,6	16,6	20,0	78,9
Cluster 2	1,1	46,7	34,1	2,3	417,5	49,9	37,8	21,1	4,4	34,9	3,8	95,8
Cluster 3	1,1	50,5	25,8	2,5	71,3	63,3	41,3	18,8	2,9	12,7	2,1	97,2
Cluster 4	-4,3	49,5	27,4	2,2	23,5	61,7	24,9	20,2	6,9	6,7	3,3	95,9
Cluster 5	-0,8	53,3	26,7	2,4	59,9	65,2	35,7	19,7	3,0	12,3	1,5	97,4
Cluster 6	-3,7	47,2	27,2	2,3	54,6	62,4	26,2	22,5	4,7	10,1	2,5	96,5
Cluster 7	-2,1	58,0	24,3	2,4	21,4	69,7	32,8	20,2	3,3	7,7	1,9	96,9
Cluster 8	-0,9	48,6	23,4	2,4	40,7	65,9	33,9	17,3	3,7	11,4	4,9	94,6
Cluster 9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3,9	2,0	90,7
Cluster 10	33,5	21,3	28,3	6,2	37,7	62,5	57,6	20,9	4,6	9,5	2,9	96,5

Indicator	X14	X15	X16	X17	X18	X19	X20	X21	X23	X24	X25
Mean value	18,9	11,4	8,5	11,3	13,4	11,1	14,6	7,5	77,2	10,2	7,8
Cluster 1	17,5	14,0	9,0	10,9	12,3	10,0	14,6	8,1	68,0	12,3	15,1
Cluster 2	10,0	10,9	12,1	16,4	16,3	11,8	12,0	6,8	58,3	17,5	20,8
Cluster 3	14,3	7,6	7,6	12,4	16,8	12,6	15,9	8,6	76,7	11,5	6,2
Cluster 4	28,1	20,5	11,8	5,4	5,1	7,0	13,4	6,6	72,6	15,7	7,6
Cluster 5	15,2	8,9	9,0	13,5	15,7	12,4	14,4	7,5	80,9	8,5	6,2
Cluster 6	35,2	19,9	5,4	4,7	6,5	8,5	12,6	5,0	75,8	9,1	9,7
Cluster 7	28,5	9,4	6,4	11,4	12,8	8,8	12,2	7,6	85,2	4,7	2,9
Cluster 8	18,6	10,9	4,0	4,3	6,4	8,2	29,2	13,6	78,6	10,5	6,7
Cluster 9	15,4	8,3	8,5	13,2	14,8	11,0	12,8	6,5	74,0	9,4	6,0
Cluster 10	26,0	16,6	8,2	10,4	10,3	8,8	12,0	5,4	79,5	11,4	5,1

Indicator	X27	X28	X29	X30	X32	X33	X34	X35	X36	X37	X38
Mean value	832,7	634,5	183,3	62,8	29,6	842,8	2482,7	398,4	41,4	2000,9	83,6
Cluster 1	785,7	598,5	273,0	62,4	44,5	1342,1	2173,0	501,4	44,9	2062,8	77,2
Cluster 2	697,2	588,3	928,8	62,0	21,6	552,4	1987,0	1703,7	1,1	605,7	60,9
Cluster 3	855,0	649,7	171,0	62,4	90,8	2438,4	2390,4	389,8	33,0	2327,2	83,7
Cluster 4	819,9	618,2	33,5	65,2	88,1	2740,0	2755,9	88,0	120,4	2759,8	88,2
Cluster 5	860,6	652,5	141,4	62,5	5,6	156,9	2570,9	342,4	24,0	1737,4	84,8
Cluster 6	800,5	615,0	114,7	63,6	3,4	106,9	2477,5	260,0	34,1	1704,4	86,7
Cluster 7	957,1	661,5	41,7	61,3	25,3	789,4	2818,0	113,7	222,6	5863,6	89,9
Cluster 8	819,9	645,4	104,5	66,7	69,8	2086,4	2486,1	247,2	32,0	1897,8	83,8
Cluster 9	0,0	0,0	0,0	0,0	19,4	592,5	0,0	7,0	0,0	320,1	90,0
Cluster 10	711,9	560,1	86,3	63,4	41,5	1297,3	1285,5	211,1	46,8	2123,2	84,1

The following description of the clusters is based on the mean values in Table 5. To help classify the clusters, the proportions of municipalities per cluster are assigned to the seven municipality types of the BBSR typology in Figure 5 (BBSR 2015). The criteria for classifying the municipalities are the population and the central function of the municipality. The evaluation of the central function is based on the central place theory of Christaller (1980). A municipality is defined as a rural municipality if either the population is less than 5,000 inhabitants or if the municipality has no basic central function. The cities in the BBSR typology are classified according to population size with the lower limits of 5,000, 10,000, 20,000, 50,000, 100,000 and 500,000 inhabitants.

Cluster 1 contains an above-average number of larger towns (see Figure 5). This cluster is characterised by the highest share of district heating systems by far. This is obvious since district heating networks are particularly suitable in towns and conurbations with high heat demand densities (Connolly et al. 2014). The high proportion of over 65-year-olds is also typical of German cities (Lauf et al. 2016). The population density is above average, while vehicles per 1,000 inhabitants are the second lowest. The potential for renewable energies is below average except for the mediocre wind power potential.

The largest share of cities is in **Cluster 2** (see. Figure 5). In this cluster, the rural municipalities account for the smallest share compared to the other clusters, and the cities from the larger small town to the larger cities take the highest share. Figure 4 shows that the largest cities in Germany, such as Berlin, Hamburg, Munich and Cologne, are all part of this cluster. For this reason, the indicators household density, population density as well as the shares of terraced houses and semi-detached houses are particularly high in this cluster, and the share of detached houses is particularly low. Furthermore, buildings built between 1950 and 1979 dominate the municipalities in this cluster. This is due to the destruction of many cities during the Second World War. In the city of Dresden, for example, large areas of prefabricated concrete slab buildings were created in the 1970s due to a shortage of houses (Wurm et al. 2009). The number of vehicles per 1,000 inhabitants is the lowest, as there are more transport alternatives in cities and the average distances travelled are shorter because of the high population density (Woldeamanuel et al. 2009). Due to the high building density, the technical PV potential per km² is the highest here. On the other hand, the technical PV potential per inhabitant is the lowest after Cluster 10 due to the high population density. As expected, the proportion of forest and agricultural land in this cluster is the smallest, so the technical wind power potential is also very low. The geothermal potential is below average.

In **Cluster 3**, the hydrothermal potential is very high; an average hydrothermal temperature of 90°C at a depth of 2,400 metres can be used in the municipalities. Figure 4 also shows that the municipalities of this cluster are predominantly located in the three large German hydrothermal regions “North German Basin”, ”Upper Rheine Graben” and “South German Molasse Basin” (Agemar et al. 2014). The potentials for the other renewable energies are average. Furthermore, there are more modern detached houses in the municipalities of the cluster, the income per household is high, and the unemployment rate is particularly low. From this cluster onwards, the share of rural municipalities in each cluster is more than 45%, and larger cities from the small midtown onwards are only very little represented (see. Figure 5).

In Figure 4, **Cluster 4** is represented by dark green coloured municipalities and occupies a large, almost continuous area. A closer look reveals that the western border of the area corresponds to the border of the former German Democratic Republic (GDR). The municipalities from Cluster 6 and 8 are also predominantly located in the territory of the former GDR. Cluster 4 is characterised by a high proportion of old houses, and the proportion of buildings built between 1919 and 1949 reaches its maximum here. Buildings built between 1970 and 1989 are very scarce in these municipalities. Also, the unemployment rate is particularly high. In line with this, the population in these municipalities has been declining the most in recent years, and income per household is the lowest. The sharp decline in the population is due to the growing childlessness in eastern Germany since German reunification (Bernardi & Keim 2017). At the same time, population density and average household size are the lowest in this cluster. These two latter indicators also determine the high value of photovoltaic potential per inhabitant. On the other hand, the photovoltaic potential per km² in this cluster is the second lowest after cluster 9, due to the small share of settlement areas in the total area. In contrast to this, the wind power potential in this cluster is the second highest. In addition, the municipalities of the cluster could exploit the second highest hydrothermal potential in Germany, but this would require drilling 300 metres deeper on average than in Cluster 3.

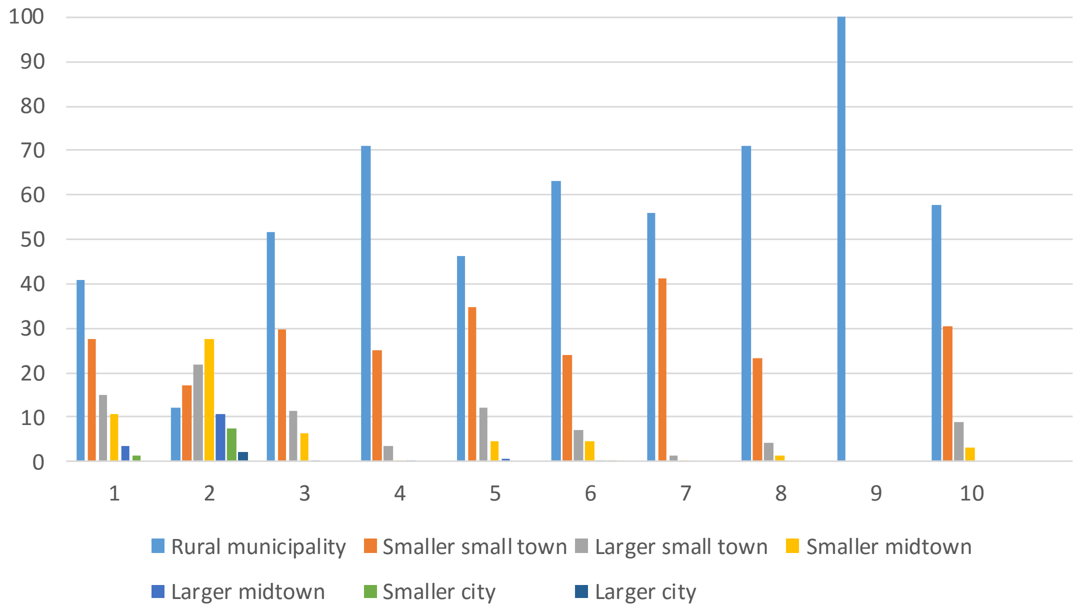


Figure 5: Classification of municipalities according to the BBSR municipality typology.

With 47% of all German municipalities, **Cluster 5** contains the largest number of municipalities. In contrast to cluster 1, district heating systems are the least widespread in this case, whereas the proportion of heating types that are not based on district heating is the highest. The number of cars and motor vehicles per 1,000 inhabitants is also very high. The cluster has the second lowest hydrothermal potential. In addition, the potential for photovoltaics in this cluster is only mediocre, and the potential for wind power is low. Based on the indicators selected in this study, Cluster 5 represents the “average” municipalities in Germany.

Like Cluster 4, **Cluster 6** is also characterised by a high building age because of its location in eastern Germany. Due to the increasing childlessness, the population is declining and the proportion of people over 65 years of age is steadily increasing. The values of the indicators representing the transport sector are rather average. In contrast to the low to average wind power and photovoltaic potential, the hydrothermal potential in this cluster is particularly low.

Cluster 7 contains almost exclusively rural municipalities and small towns (see Figure 5). The proportion of apartments occupied by the owner and the living space per person are at their maximum, while at the same time the household density is minimal. Due to the low density of households and population, the number of cars and motor vehicles per 1,000 inhabitants reaches its maximum here. In addition, the detached houses reach the largest share in this cluster. Furthermore, this cluster has the highest potential for renewable energies, despite its very low geothermal potential. The high living space per person and the low density of households mean that the highest photovoltaic and wind power potentials per person are achieved. Also, the wind power potential per km² is at its maximum, as most of the municipalities in the cluster are located in Northern Germany and thus in areas with high wind speeds and have a high proportion of forest and agricultural land.

The building age in **Cluster 8** is unusually low, although these municipalities are mainly located in Eastern Germany. This can be explained by an example: in the description of clusters 4 and 6, the decline in population in eastern Germany has already been discussed. Although this development applies to all the new federal states, the decline in Brandenburg between 1990 and 2008 was significantly lower. This was mainly due to new settlements in the surrounding area of Berlin, the so-called “commuter belt” (Jesse et al. 2014). Municipalities from cluster 8 almost exclusively form this commuter belt (cf. Figure 4). Due to the rising rent in Berlin, more and more young families are moving into the commuter belt. This also explains the maximum proportion of 18-64-year-olds and the minimum proportion of 65 year-olds in this cluster (Bünger 2017). The proportion of cars per 1,000 inhabitants is also above average here, presumably because most people have to drive to work in the city. A closer examination of the red municipalities shown in Figure 4 reveals that most of the municipalities are located in the surrounding area of major cities in clusters 1 and 2. Thus, the conclusions mentioned above on the Berlin “commuter belt” can also be transferred to the other municipalities in Cluster 8. Also, this cluster has the third highest geothermal potential, while the potential of the other renewable energies is below average.

Cluster 9 contains all areas in which there are no inhabitants. Therefore, all indicators that depend on the population have a value of zero. These areas are municipality-free (in German: “gemeindefrei”), and therefore 100% of them are rural municipalities (see Figure 5). Settlement and traffic area is present in these municipalities, because of roads leading through these areas. However, this indicator has the smallest value here. At the same time, the proportion of forest and agricultural land reaches its highest level. It is interesting to note that the technical wind power potential per km² is nevertheless at its minimum in this cluster. The reason for this could be, among other things, nature reserves in which no wind turbines may be installed. The technical photovoltaic potential in this cluster is also approaching zero since only a few buildings are located here. Despite the buildings, no residents are assigned to these areas, as the buildings in the municipal areas belong to military training areas or similar (Goderbauer 2016). This cluster has the lowest potential for renewable energies, as the geothermal potential is also below average.

With only 33 municipalities, **Cluster 10** represents the smallest cluster in this study. This cluster is characterised by the highest population growth between 2010 and 2015. Due to the largest average household size by far, the income per household is also reaching its maximum value and the technical PV potential per inhabitant its minimum value. In addition, the number of vehicles per 1,000 inhabitants in this cluster is below average. The cluster must be evaluated as an outlier since many of the characteristics of this cluster are due to the high population growth. The population figures from 2015 have been used in the calculation of many indicators to establish a uniform reference. However, the most recent household data are available for 2014 and have only been roughly updated since the last survey in 2011. As a result, the high population growth leads to, among other things, high values for the average household size, as the number of households is no longer up to date. This cluster, therefore, includes outliers. Nevertheless, the heterogeneity and independence of the cluster can be justified by the significantly higher population growth as in the other clusters.

5. Discussion

5.1. Critical appraisal of the methodology

Wall (2016) shows that factor analysis is an important step ahead of cluster analysis. However, most studies describe cluster analyses without prior factor analysis. For this reason, the cluster analysis was repeated again without the factor analysis. The results were worse than those of the cluster analysis with the values from the factor analysis. For example, with the factor values, the 75 municipality-free areas without population (Cluster 9) were already divided into a cluster in the 6 cluster solution (cf. Table 4). In the analysis with the raw data, these municipalities were not separated, at least up to the 20 cluster solution.

Whilst cluster analysis provides a good basis for transferring the results of energy autonomy studies to other municipalities, results cannot always be completely transferred and an examination of the individual case will be necessary. This is illustrated by the following example: Figure 6 shows the violin plot of the indicator “Share of buildings with heating systems based on district heating (X11)” for all clusters. In a violin plot, the density trace and the box plot are combined into one diagram (for more information see: Hintze & Nelson 1998). The red plus signs indicate the position of the mean value and the green boxes indicate the position of the median.

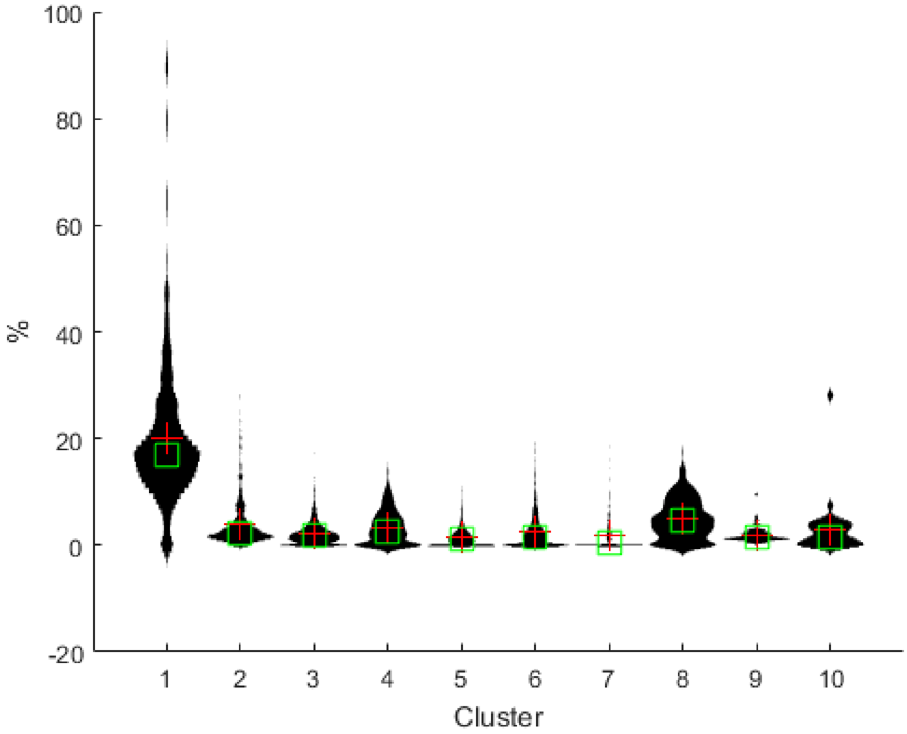


Figure 6: Violin Plot of indicator X11 in % for the ten clusters.

Cluster 1 is characterised by a high average share of district heating. However, Figure 6 shows that this cluster also contains a few municipalities with very low shares of district heating. These municipalities are then more similar to the cluster focus in the other indicators.

The following explains in more detail why it is necessary to carry out a cluster analysis to appropriately transfer the results of energy system analyses in municipalities by comparing the model or average municipalities from the studies described in Section 1 with the mean values of all municipalities in this study (cf. Table 6).

Table 6 shows that the average municipality differs from the municipalities/regions of the studies of Jenssen et al. (2014), Scheffer (2008) and Peter (2013). A comparison of the values reveals that none of the surveyed municipalities represents an average municipality in Germany. Even though it was not the intention to select an average municipality in Germany in some of the studies, the results are difficult to transfer to other municipalities or regions. Rather, the choice of municipalities and data appears to be influenced by many other than technical factors in some cases. For example, Scheffer (2008) tried to use the indicator values to describe a rural municipality. In the classification of German municipalities in BBSR (2015), the municipality would be placed in the category “Smaller Small Town”. In addition, only a few data on the municipalities are described in the studies. This could give the impression that the results from Jenssen et al. (2014), for example, can be transferred to municipalities with 3,000 inhabitants, 800 buildings and a household size of 2.2 persons. Instead, a transferability depends on how precisely the municipality is represented, i. e. how many indicators are used to describe the municipality. Since the cluster analysis carried out here uses considerably more indicators to describe the municipalities, the result can be used as a basis for transferring appropriate energy systems to other municipalities.

Table 6: Comparison of the model, example and average municipalities/regions from relevant literature with the average municipality from this study.

Municipality from	Number of inhabitants	Number of buildings	Average house-hold size [people]	Share of settlements and traffic areas [%]	Population density [Inhabitants /km²]	Number of vehicles per 1,000 inhabitants
Jenssen et al. (2014)	3,000	800	2.2	Not specified	Not specified	Not specified
Scheffer (2008)	10,000	Not specified	Not specified	Not specified	Not specified	630
Peter (2013)	3,850	1,224	3.1	8	106	Not specified
Burgess et al. (2012)	25,550	Not specified	2.4	8	310	Not specified
Schmidt et al. (2012)	20,619	Not specified	Not specified	< 11	68	Not specified
Woyke & Forero (2014)	1,100	Not specified	Not specified	Not specified	Not specified	Not specified
Average munici-pality (cf. Table 5)	7,380	1,670	2.4	13	183	830

5.2. Suitability of municipalities for energy autonomy

To investigate the suitability of individual municipalities and clusters for energy autonomy, precise calculations must be carried out. Nevertheless, in this section an attempt is made to determine an initial assessment of this suitability by analysing the clusters in which municipalities are already aiming for energy autonomy. For this, 165 municipalities from the energy projects “Energy Municipalities”, “Bioenergy Villages” and “100% Renewable Energy Regions” are assigned to the ten clusters. These projects aim to achieve the goal of an autonomous energy supply in the municipalities. However, the municipalities have defined different objectives in the projects about autonomous supply, including “100% heat”, ”100% electricity” or “100% renewable energies” (McKenna et al. 2014b). Some of the municipalities take part in several of the projects mentioned above. Districts and counties involved in the projects were not included in the analysis. The result of the assignment is shown in Table 7.

Table 7: Assignment of the municipalities from the energy projects "Energy Municipalities", "Bioenergy Villages" and "100% Renewable Energy Regions" to the ten clusters.

Cluster	1	2	3	4	5	6	7	8	9	10
Number	25	21	20	6	76	12	2	3	0	0
Fraction in the cluster	7,4%	2,9%	1,2%	0,7%	1,4%	0,9%	0,4%	0,8%	0%	0%
Example	Jühnde	München	Furth	Barth	Brilon	Jena	Hürup	Pleß	-	-

First of all, it is noticeable that no municipalities from clusters 9 and 10 participate in the energy projects. This fact is quickly explained since there is no population in Cluster 9 and Cluster 10 is very small and is also more of an outlier cluster. The first two clusters, on the other hand, have the largest proportion of municipalities that are members of the energy projects. As shown above, these clusters contain most of the cities (see Figure 5). One reason for the high proportions in these clusters could be the existence of a critical mass of innovators (cf. Deutschle et al. 2015), but such aspects could not be included in this analysis. On the other hand, however, achieving the goal of energy autonomy is all the more difficult, the more inhabitants a municipality has. While in rural municipalities the focus is often on the expansion of renewable energies, development in large cities depends to a large extent on development outside the city borders. Discussions in major cities are mainly focused on increasing energy efficiency, creating smart grids and providing storage capacity. For example, the City of Munich aims to halve per capita emissions by 2030 (Gailing et al. 2013). In the other clusters, the municipalities' participation in energy projects is not so pronounced. However, in each of the clusters 3 to 8, at least two municipalities are involved in energy projects. This is likely to be due to the non-technical reasons mentioned above rather than a suitability of the municipalities per se. If the potential for renewable energies is used as a basis for the assessment, the municipalities from Cluster 3 and Cluster 7 could be particularly suitable for energy autonomy: Cluster 3, among other things, due to its high hydrothermal potential and the associated potential base-load energy supply; Cluster 7 because of the highest potential for renewable energies.

6. Summary and Conclusion

In the context of the trend towards decentralised energy systems, both high temporal and spatial resolutions are required in order to adequately consider their interactions with the centralised system. This is a central challenge in energy modelling, as compromises must inevitably be made between model resolution, scope and computational feasibility. This paper makes a significant contribution to complexity reduction in this area by clustering the 11,131 German municipalities using 34 pre-identified socio-energetic indicators, mainly based on freely available data relating to the consumption sectors of Private Households and Transport, as well as indicators relating to the potentials for renewable energies. The method involves two main steps, namely a factor analysis and a cluster analysis. For the former, different methods are weighed against each other, and the most effective methods for allocating the indicators to factors are chosen. Selected cluster validation methods are then used to determine an appropriate number of 10 clusters to which the 11,131 municipalities are distributed. Due to the high number and differentiation of indicators, clusters overlap with each other for different indicators, but the results also show significant differences between the clusters. For example, Cluster 2 contains all major German cities and most of the other cities in Germany and has a low potential for renewable energies. Cluster 9, on the other hand, describes all German municipalities in which there are no inhabitants.

The methodology used in this study could be improved for more accurate results in future work. On the one hand, other indicators should be included in the study, including, if possible, indicators from the Industry and Commercial sector as well as indicators relating to the local climate. However, this is challenging due to the lack of available data at this spatial resolution. If available, data should also refer to the same year, as some of the results might be distorted because of different reference years, as shown by the average household size in cluster 10. Furthermore, weights for the indicators should be determined with the help of expert interviews. If it is known which indicators have the most considerable influence on the suitability for energy autonomy, these can be weighted more strongly in the cluster analysis and a new set of clusters generated based on these weights. In addition, the employed cluster methodology should be scrutinised more closely. Although the selection methods can be adequately justified in this study, others (e.g. Chicco 2012) have shown that the Ward algorithm is not always the best choice for cluster analysis. Further work is also required to analyse the economic effects of municipal energy autonomy on the overarching energy system (for a discussion see Jägemann et al. 2013, McKenna 2017).

A comparison of the average municipality from the dataset used here with the average municipalities from other energy autonomy studies is difficult due to a lack of data at the level of detail employed here. Based on the available data in these studies, a comparison shows few similarities, which means the results of the studies relating to their transferability to other municipalities should be questioned. Assigning the municipalities from the three German energy projects “Energy Municipalities”, ”Bioenergy Villages” and “100% Renewable Energy Regions” to the 10 clusters further shows that in eight of the ten clusters municipalities are aiming for energy autonomy (in varying degrees). As a result, it is not possible to differentiate between the clusters regarding readiness for such energy projects, which is most likely due to the influence of non-technical factors on the emergence of these initiatives. However, the results of the cluster analysis show that some of the municipalities could be technically more suitable for energy autonomy, for example Cluster 7 is characterised by a high potential for renewable energies. A comparison of the ten clusters with the average municipality from the data set also demonstrates their benefit, with a large variation across clusters in terms of energy demand structure, renewables potentials and overall size. The results therefore reduce the effort of subsequent studies, as only a few municipalities from the clusters need to be examined regarding their suitability for energy autonomy to be able to make statements for all municipalities of the cluster. However, this study also makes it clear that not every result can be transferred to all the other municipalities within a cluster, instead an individual examination is required for each municipality. Nevertheless, the results help to identify municipalities in which already successful measures from other municipalities could be applied, and provide a basis for further energy analyses at the national level.

Acknowledgements

The authors gratefully acknowledge funding by the German Federal Ministry of Education and Research (BMBF) within the Kopernikus Project ENSURE 'New ENergy grid StructURes for the German Energiewende' (funding reference: FKZ 03SFK1N0).

Publication bibliography

acatech (2016): Flexibility concepts for the German power supply 2050. Ensuring stability in the age of renewable energies. Available online at https://www.leopoldina.org/uploads/tx_leopublication/2016_02_Stellungnahme_Flexibility_concepts.pdf, checked on 4/4/2017.

Agemar, Thorsten; Weber, Josef; Schulz, Rüdiger (2014): Deep Geothermal Energy Production in Germany. In Energies 7 (7), pp. 4397–4416. DOI: 10.3390/en7074397.

Agemar, T. (2017): Geothermische Informationssystem für Deutschland. Available online at https://www.geotis.de/homepage/project.php?loc=de_DE, checked on 8/14/2017.

Albatineh, Ahmed N.; Niewiadomska-Bugaj, Magdalena (2011): MCS. A Method for Finding the Number of Clusters. In J Classif 28 (2), pp. 184–209. DOI: 10.1007/s00357-010-9069-1.

Aljandali, Abdulkader (2017): Factor Analysis. In Abdulkader Aljandali (Ed.): Multivariate Methods and Forecasting with IBM® SPSS® Statistics. Cham, s.l.: Springer International Publishing (Statistics and Econometrics for Finance), pp. 97–106.

Anderson, Marti J. (2001): A new method for non-parametric multivariate analysis of variance. In Austral Ecology 26 (1), pp. 32–46. DOI: 10.1111/j.1442-9993.2001.01070.pp.x.

Arbelaitz, Olatz; Gurrutxaga, Ibai; Muguerza, Javier; Pérez, Jesús M.; Perona, Iñigo (2013): An extensive comparative study of cluster validity indices. In Pattern Recognition 46 (1), pp. 243–256. DOI: 10.1016/j.patcog.2012.07.021.

Bacher, Johann; Pöge, Andreas; Wenzig, Knut (2010): Clusteranalyse. Anwendungsorientierte Einführung in Klassifikationsverfahren. 3., erg., vollst. überarb. und neu gestaltete Aufl. München: Oldenbourg.

Balcombe, Paul; Rigby, Dan; Azapagic, Adisa (2015): Energy self-sufficiency, grid demand variability and consumer costs. Integrating solar PV, Stirling engine CHP and battery storage. In Applied Energy 155, pp. 393–408. DOI: 10.1016/j.apenergy.2015.06.017.

Ball, G. H.; Hall, D. J. (1965): ISODATA: a novel method of data analysis and pattern classification. Springfield: NTIS. Available online at http://worldcatlibraries.org/wcpa/oclc/67143991.

Bauknecht, D.; Vogel, M.; Funcke, S. (2015): Energiewende – Zentral oder dezentral? Diskussionspapier im Rahmen der Wissenschaftlichen Koordination des BMBF Förderprogramms „Umwelt- und Gesellschaftsverträgliche Transformation des Energiesystems“. Öko-Institut. Freiburg.

BBSR (2015): Referenz Gemeinden und Gemeindeverbände, Stadt- und Gemeindetyp. Available online at http://www.bbsr.bund.de/BBSR/DE/Raumbeobachtung/Raumabgrenzungen/StadtGemeindetyp/downloadangebote.html?nn=443182, checked on 8/29/2017.

Beale, E. M. (1969): Euclidean Cluster Analysis.: Scientific Control Systems Ltd. Available online at http://worldcatlibraries.org/wcpa/oclc/500100942, checked on 8/14/2017.

Bernardi, L.; Keim, S. (2017): Childless at Age 30. A Qualitative Study of the Life Course Plans of Working Women in East and West Germany. In Michaela Kreyenfeld, Dirk Konietzka (Eds.): Childlessness in Europe. Contexts, causes, and consequences. Berlin: Springer Open, pp. 253–267.

BMWi (2016a): Erneuerbare Energien in Zahlen: Nationale und internationale Entwicklung im Jahr 2015, BMWi, Berlin.

BMWi (2016b): Fünfter Monitoring-Bericht zur Energiewende. Die Energie der Zukunft. Edited by Bundesministerium für Wirtschaft und Energie. Available online at https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/fuenfter-monitoring-bericht-energie-der-zukunft.pdf?__blob=publicationFile&v=24, checked on 9/8/2017.

Boon, Frank Pieter; Dieperink, Carel (2014): Local civil society based renewable energy organisations in the Netherlands. Exploring the factors that stimulate their emergence and development. In Energy Policy 69, pp. 297–307. DOI: 10.1016/j.enpol.2014.01.046.

Böttger, Diana; Götz, Mario; Lehr, Nelly; Kondziella, Hendrik; Bruckner, Thomas (2014): Potential of the Power-to-Heat Technology in District Heating Grids in Germany. In Energy Procedia 46, pp. 246–253. DOI: 10.1016/j.egypro.2014.01.179.

Bouguettaya, Athman; Yu, Qi; Liu, Xumin; Zhou, Xiangmin; Song, Andy (2015): Efficient agglomerative hierarchical clustering. In Expert Systems with Applications 42 (5), pp. 2785–2797. DOI: 10.1016/j.eswa.2014.09.054.

Bramreiter, Rafael; Truger, Barbara; Schinko, Thomas; Bednar-Friedl, Birgit (2016): Identification of economic and energy framework conditions of the Austrian climate and energy model regions. LINKS Working Paper 1.1. Available online at http://pure.iiasa.ac.at/14590/1/LINKS_Working_Paper_1-1_final.pdf, checked on 8/2/2017.

Braun, Frauke G. (2010): Determinants of households’ space heating type. A discrete choice analysis for German households. In Energy Policy 38 (10), pp. 5493–5503. DOI: 10.1016/j.enpol.2010.04.002.

Bünger, Reinhardt (2017): Arbeiten in Berlin – Wohnen in Brandenburg. In Der Tagesspiegel, 5/29/2017. Available online at http://www.tagesspiegel.de/wirtschaft/immobilien/immowelt-studie-arbeiten-in-berlin-wohnen-in-brandenburg/19855150.html, checked on 9/5/2017.

Burgess, Paul J.; Rivas Casado, Monica; Gavu, Jerry; Mead, Andrew; Cockerill, Tim; Lord, Richard et al. (2012): A framework for reviewing the trade-offs between, renewable energy, food, feed and wood production at a local level. In Renewable and Sustainable Energy Reviews 16 (1), pp. 129–142. DOI: 10.1016/j.rser.2011.07.142.

Calinski, T.; Harabasz, J. (1974): A dendrite method for cluster analysis. In Comm. in Stats. - Theory & Methods 3 (1), pp. 1–27. DOI: 10.1080/03610927408827101.

Charrad, Malika; Ghazzali, Nadia; Boiteau, Véronique; Niknafs, Azam (2014): NbClust. An R Package for Determining the Relevant Number of Clusters in a Data Set. In J. Stat. Soft. 61 (6). DOI: 10.18637/jss.v061.i06.

Chévez, Pedro; Barbero, Dante; Martini, Irene; Discoli, Carlos (2017): Application of the k-means clustering method for the detection and analysis of areas of homogeneous residential electricity consumption at the Great La Plata region, Buenos Aires, Argentina. In Sustainable Cities and Society 32, pp. 115–129. DOI: 10.1016/j.scs.2017.03.019.

Chicco, Gianfranco (2012): Overview and performance assessment of the clustering methods for electrical load pattern grouping. In Energy 42 (1), pp. 68–80. DOI: 10.1016/j.energy.2011.12.031.

Christaller, Walter (1980): Die zentralen Orte in Süddeutschland. E. ökonomisch-geographische Untersuchung über d. Gesetzmäßigkeit d. Verbreitung u. Entwicklung d.

Siedlungen mit städtischen Funktionen. 3., unveränd. Aufl., reprograf. Nachdr. d. 1. Aufl., Jena 1933. Darmstadt: Wissenschaftliche Buchges.

Connolly, D.; Lund, H.; Mathiesen, B. V.; Werner, S.; Möller, B.; Persson, U. et al. (2014): Heat Roadmap Europe. Combining district heating with heat savings to decarbonise the EU energy system. In Energy Policy 65, pp. 475–489. DOI: 10.1016/j.enpol.2013.10.035.

Davies, David L.; Bouldin, Donald W. (1979): A Cluster Separation Measure. In IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1 (2), pp. 224–227. DOI: 10.1109/TPAMI.1979.4766909.

Deutscher Bundestag (2017): Gesetz für den Ausbau erneuerbarer Energien (ErneuerbareEnergien-Gesetz - EEG 2017). EEG 2017. Available online at https://www.gesetze-im-internet.de/eeg_2014/EEG_2017.pdf, checked on 8/14/2017.

Deutschle, Jürgen; Hauser, Wolfgang; Sonnberger, Marco; Tomaschek, Jan; Brodecki, Lukasz; Fahl, Ulrich (2015): Energie-Autarkie und Energie-Autonomie in Theorie und Praxis. In Z Energiewirtsch 39 (3), pp. 151–162. DOI: 10.1007/s12398-015-0160-5.

Dillon, William R.; Kumar, Ajith; Mulani, Narendra (1987): Offending estimates in covariance structure analysis. Comments on the causes of and solutions to Heywood cases. In Psychological Bulletin 101 (1), pp. 126–135. DOI: 10.1037/0033-2909.101.1.126.

Duda, Richard O.; Hart, Peter E. (1973): Pattern classification and scene analysis. New York NY u.a.: Wiley (A Wiley Interscience publication).

Dunn, J. C. (1974): Well-Separated Clusters and Optimal Fuzzy Partitions. In Journal of Cybernetics 4 (1), pp. 95–104. DOI: 10.1080/01969727408546059.

Durst, Anna (2015): Fernwärme effizient nutzen. Edited by BINE Informationsdienst (14/2015). Available online at http://www.bine.info/fileadmin/content/Publikationen/Projekt-Infos/2015/Projekt_14-2015/ProjektInfo_1415_internetx.pdf, checked on 9/8/2017.

Eckstein, Peter P. (Ed.) (2016): Statistik für Wirtschaftswissenschaftler. Eine realdatenbasierte Einführung mit SPSS. 5., aktualisierte und erweiterte Auflage. Wiesbaden: Springer Gabler.

Engelken, Maximilian; Römer, Benedikt; Drescher, Marcus; Welpe, Isabell (2016): Transforming the energy system. Why municipalities strive for energy self-sufficiency 98, pp. 365–377. DOI: 10.1016/j.enpol.2016.07.049.

Frey, T.; van Groenewoud, H. (1972): A Cluster Analysis of the D 2 Matrix of White Spruce Stands in Saskatchewan Based on the Maximum-Minimum Principle. In The Journal of Ecology 60 (3), p. 873. DOI: 10.2307/2258571.

Friedman, H. P.; Rubin, J. (1967): On Some Invariant Criteria for Grouping Data. In Journal of the American Statistical Association 62 (320), pp. 1159–1178. DOI: 10.1080/01621459.1967.10500923.

Funcke, Simon; Bauknecht, Dierk (2016): Typology of centralised and decentralised visions for electricity infrastructure. In Utilities Policy 40, pp. 67–74. DOI: 10.1016/j.jup.2016.03.005.

Gailing, Ludger; Hueesker, Frank; Kern, Kristine; Röhring, Andreas (2013): Die räumliche Gestaltung der Energiewende zwischen Zentralität und Dezentralität. Explorative Anwendung einer Forschungsheuristik. In IRS Working Paper 151.

Gentle, J. E.; Kaufman, L.; Rousseuw, P. J. (1991): Finding Groups in Data. An Introduction to Cluster Analysis. In Biometrics 47 (2), p. 788. DOI: 10.2307/2532178.

Geyler, Stefan; Warner, Barbara; Brandl, Anja; Kuntze, Martina; Forschungsverbund KoReMi (Eds.) (2008): Clusteranalyse der Gemeinden in der Kernregion Mitteldeutschland. Universität Leipzig. Leipzig (02).

Goderbauer, Sebastian (2016): Daten von Deutschland. In Sebastian Goderbauer (Ed.): Mathematische Optimierung der Wahlkreiseinteilung für die Deutsche Bundestagswahl. Modelle und Algorithmen für eine bessere Beachtung der gesetzlichen Vorgaben. Wiesbaden: Springer Spektrum (BestMasters), pp. 121–143.

Halkidi, M.; Vazirgiannis, M.; Batistakis, Y. (2000): Quality Scheme Assessment in the Clustering Process. In Djamel A. Zighed, Jan Komorowski, Jan Żytkow (Eds.): Principles of Data Mining and Knowledge Discovery. 4th European Conference, PKDD 2000 Lyon, France, September 13-16, 2000 Proceedings. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science, 1910), pp. 265–276.

Halkidi, M.; Vazirgiannis, M. (2001): Clustering validity assessment: finding the optimal partitioning of a data set. In Nick Cercone, Tsau Y. Lin, Xindong Wu (Eds.): 2001 IEEE International Conference on Data Mining. Proceedings: 29 November-2 December, 2001, San Jose, California. 2001 IEEE International Conference on Data Mining. San Jose, CA, USA, 29 Nov.-2 Dec. 2001. IEEE Computer Society; IEEE International Conference on Data Mining. Los Alamitos, Calif: IEEE Computer Society, pp. 187–194.

Hargreaves, Anthony; Cheng, Vicky; Deshmukh, Sandip; Leach, Matthew; Steemers, Koen (2017): Forecasting how residential urban form affects the regional carbon savings and costs of retrofitting and de centralised energy supply. In Applied Energy 186, pp. 549–561. DOI: 10.1016/j.apenergy.2016.02.095.

Hartigan, John A. (1975): Clustering algorithms. New York NY u.a.: Wiley (Wiley series in probability and mathematical statistics: Applied probability and statistics).

Hechler, H.; Bredel-Schürmann, Stefan (2011): Tiefe Geothermie. Edited by ASUE Arbeitsgemeinschaft für sparsamen und umweltfreundlichen Energieverbrauch e.V. Berlin. Available online at
http://asue.de/sites/default/files/asue/themen/umwelt_klimaschutz/2011/broschueren/07_05_11_asue-tiefe-geothermie.pdf, checked on 8/14/2017.

Heinritz, Günter (Ed.) (2000): Nationalatlas Bundesrepublik Deutschland. Leibniz-Institut für Länderkunde. München: Elsevier Spektrum Akadem. Verl.

Heyde, Thomas P. E. auf der (1990): Analyzing chemical data in more than two dimensions. A tutorial on factor and cluster analysis. In J. Chem. Educ. 67 (6), p. 461. DOI: 10.1021/ed067p461.

Hintze, Jerry L.; Nelson, Ray D. (1998): Violin Plots. A Box Plot-Density Trace Synergism. In The American Statistician 52 (2), pp. 181–184. DOI: 10.1080/00031305.1998.10480559.

Horn, John L. (1965): A rationale and test for the number of factors in factor analysis. In Psychometrika 30 (2), pp. 179–185. DOI: 10.1007/BF02289447.

Høyer, Karl Georg (2008): The history of alternative fuels in transportation. The case of electric and hybrid cars. In Utilities Policy 16 (2), pp. 63–71. DOI: 10.1016/j.jup.2007.11.001.

Hubert, Lawrence J.; Levin, Joel R. (1976): A general statistical framework for assessing categorical clustering in free recall. In Psychological Bulletin 83 (6), pp. 1072–1080. DOI: 10.1037/0033-2909.83.6.1072.

Hubert, Lawrence; Arabie, Phipps (1985): Comparing partitions. In Journal of Classification 2 (1), pp. 193–218. DOI: 10.1007/BF01908075.

Institut für Wohnen und Umwelt (2017): Gradtagszahlen in Deutschland. Available online at http://t3.iwu.de/fileadmin/user_upload/dateien/energie/werkzeuge/Gradtagszahlen_Deutschland.xls, checked on 9/7/2017.

Islam, Md. Atiqul; Alizadeh, Behrooz Z.; van den Heuvel, Edwin R.; Bruggeman, Richard; Cahn, Wiepke; Haan, Lieuwe de et al. (2016): A comparison of indices for identifying the number of clusters in hierarchical clustering. A study on cognition in schizophrenia patients. In Communications in Statistics: Case Studies, Data Analysis and Applications 1 (2), pp. 98–113. DOI: 10.1080/23737484.2015.1103670.

Jägemann, C., Hagspiel, S., Lindenberger, D., 2013. The economic inefficiency of grid parity: The case of German photovoltaics, EWI Working Paper No 13/19, December 2013.

Javied, Tallal; Rackow, Tobias; Stankalla, Roland; Sterk, Christian; Franke, Jörg (2016): A Study on Electric Energy Consumption of Manufacturing Companies in the German Industry with the Focus on Electric Drives. In Procedia CIRP 41, pp. 318–322. DOI: 10.1016/j.procir.2015.10.006.

Jensch, W. (1989): Energieversorgung: Ein Vergleich zentraler und dezentraler Systeme. In Energiewirtschaftliche Tagesfragen 39 (4), pp. 190-195.

Jenssen, Till; König, Andreas; Eltrop, Ludger (2014): Bioenergy villages in Germany. Bringing a low carbon energy supply for rural areas into practice. In Renewable Energy 61, pp. 74–80. DOI: 10.1016/j.renene.2012.08.014.

Jesse, Eckhard; Schubert, Thomas; Thieme, Tom (Eds.) (2014): Politik in Sachsen. Wiesbaden: Springer VS (SpringerLink).

Johnson, Stephen C. (1967): Hierarchical clustering schemes. In Psychometrika 32 (3), pp. 241–254. DOI: 10.1007/BF02289588.

Kaiser, Henry F. (1960): The Application of Electronic Computers to Factor Analysis. In Educational and Psychological Measurement 20 (1), pp. 141–151. DOI: 10.1177/001316446002000116.

Kaundinya, Deepak Paramashivan; Balachandra, P.; Ravindranath, N. H.; Ashok, Veilumuthu (2013): A GIS (geographical information system)-based spatial data mining approach for optimal location and capacity planning of distributed biomass power generation facilities. A case study of Tumkur district, India. In Energy 52, pp. 77–88. DOI: 10.1016/j.energy.2013.02.011.

Keles, D.; Jochem, P.; McKenna, R.; Ruppert, M.; Fichtner, W. (2017). Meeting the Modeling Needs of Future Energy Systems. In Energy technology 5 (7), pp. 1007-1025. doi:10.1002/ente.201600607.

Klaus Novy Institut e.V.; trend:research (2011): „Marktakteure: Erneuerbare-Energien-Anlagen in der Stromerzeugung“. Köln.

Kodinariya, T.; Makwana, P. (2013): Review on determining number of Cluster in K-Means Clustering. In International Journal of Advance Research in Computer Science and Management Studies 1 (6), pp. 90–95.

Koirala, Binod Prasad; Koliou, Elta; Friege, Jonas; Hakvoort, Rudi A.; Herder, Paulien M. (2016): Energetic communities for community energy. A review of key issues and trends shaping integrated community energy systems. In Renewable and Sustainable Energy Reviews 56, pp. 722–744. DOI: 10.1016/j.rser.2015.11.080.

Kraftfahrt-Bundesamt (2017a): Fahrzeugzulassungen (FZ). Bestand an Kraftfahrzeugen und Kraftfahrzeuganhängern nach Gemeinden. Edited by Kraftfahrt-Bundesamt. Available online at https://www.kba.de/SharedDocs/Publikationen/DE/Statistik/Fahrzeuge/FZ/2017/fz3_2017_pdf.pdf;jsessionid=031266C8B0663B938AA41410071E0075.live21303?__blob=publicationFile&v=3, checked on 9/7/2017.

Kraftfahrt-Bundesamt (2017b): Fahrzeugzulassungen (FZ). Bestand an Kraftfahrzeugen und Kraftfahrzeuganhängern nach Zulassungsbezirken. Edited by Kraftfahrt-Bundesamt. Available online at https://www.kba.de/SharedDocs/Publikationen/DE/Statistik/Fahrzeuge/FZ/2017/fz1_2017_pdf.pdf?__blob=publicationFile&v=3, checked on 9/7/2017.

Kronthaler, F. (2003): A Study of the Competitiveness of Regions based on a Cluster Analysis: The Example of East Germany. Institut für Wirtschaftsforschung Halle. Available online at http://www.iwh-halle.de/fileadmin/user_upload/publications/iwh_discussion_papers/179.pdf, checked on 11/17/2017.

Krzanowski, W. J.; Lai, Y. T. (1988): A Criterion for Determining the Number of Groups in a Data Set Using Sum-of-Squares Clustering. In Biometrics 44 (1), p. 23. DOI: 10.2307/2531893.

Lauf, Steffen; Haase, Dagmar; Kleinschmit, Birgit (2016): The effects of growth, shrinkage, population aging and preference shifts on urban development—A spatial scenario analysis of Berlin, Germany. In Land Use Policy 52, pp. 240–254. DOI: 10.1016/j.landusepol.2015.12.017.

Lebart, Ludovic; Morineau, Alain; Piron, Marie (2002): Statistique exploratoire multidimensionnelle. 3e éd., nouveau tirage révisé. Paris: Dunod (Sciences Sup. Mathématiques. 2e et 3e cycles).

Lopes, Raul H. C. (2011): Kolmogorov-Smirnov Test. In Miodrag Lovric (Ed.): International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 718–720.

Lorenzo-Seva, Urbano; Ferrando, Pere J. (2006): FACTOR. A computer program to fit the exploratory factor analysis model. In Behavior Research Methods 38 (1), pp. 88–91. DOI: 10.3758/BF03192753.

Lüker-Jans, Nicola; Simmering, Dietmar; Otte, Annette (2017): The impact of biogas plants on regional dynamics of permanent grassland and maize area—The example of Hesse, Germany (2005–2010). In Agriculture, Ecosystems & Environment 241, pp. 24–38. DOI: 10.1016/j.agee.2017.02.023.

Mainzer, Kai; Fath, Karoline; McKenna, Russell; Stengel, Julian; Fichtner, Wolf; Schultmann, Frank (2014): A high-resolution determination of the technical potential for residential-roof-mounted photovoltaic systems in Germany. In Solar Energy 105, pp. 715–731. DOI: 10.1016/j.solener.2014.04.015.

Marquant, Julien F.; Evins, Ralph; Bollinger, L. Andrew; Carmeliet, Jan (2017): A holarchic approach for multi-scale distributed energy system optimisation. In Applied Energy 208, pp. 935–953. DOI: 10.1016/j.apenergy.2017.09.057.

Marriott, F. H. C. (1971): Practical Problems in a Method of Cluster Analysis. In Biometrics 27 (3), p. 501. DOI: 10.2307/2528592.

Marx Gómez, Jorge; Sonnenschein, Michael; Vogel, Ute; Winter, Andreas; Rapp, Barbara; Giesen, Nils (Eds.) (2014): Information and communication technology for energy efficiency. Conference proceedings; EnviroInfo 2014 - 28th International Conference on Informatics for Environmental Protection; 10th - 12th September, 2014, Oldenburg, Germany. EnviroInfo; International Conference on Informatics for Environmental Protection. Oldenburg: BIS-Verl. der Carl-von-Ossietzky-Univ. Available online at http://nbn-resolving.de/urn:nbn:de:gbv:715-oops-20009, checked on 11/17/2017.

McClain, John O.; Rao, Vithala R. (1975): CLUSTISZ: A Program to Test for the Quality of Clustering of a Set of Objects. In Journal of Marketing Research 4 (12), pp. 456–460.

McKenna, R.; Hollnaicher, S.; Fichtner, W. (2014a): Cost-potential curves for onshore wind energy. A high-resolution analysis for Germany. In Applied Energy 115, pp. 103–115. DOI: 10.1016/j.apenergy.2013.10.030.

McKenna, Russell; Jäger, Tobias; Fichtner, Wolf (2014b): Energieautarkie – ausgewählte Ansätze und Praxiserfahrungen im deutschsprachigen Raum. In uwf 22 (4), pp. 241–247. DOI: 10.1007/s00550-014-0339-y.

McKenna, Russell; Jäger, Tobias; Fichtner, Wolf (2014c): Energieautarkie – ausgewählte Ansätze und Praxiserfahrungen im deutschsprachigen Raum. In uwf 22 (4), pp. 241–247. DOI: 10.1007/s00550-014-0339-y.

McKenna, Russell; Herbes, Carsten; Fichtner, Wolf (2015): Energieautarkie. Vorschlag einer Arbeitsdefinition als Grundlage für die Bewertung konkreter Projekte und Szenarien. In Z Energiewirtsch 39 (4), pp. 235–252. DOI: 10.1007/s12398-015-0164-1.

McKenna, R.; Hofmann, L.; Merkel, E.; Fichtner, W.; Strachan, N. (2016): Analysing socioeconomic diversity and scaling effects on residential electricity load profiles in the context of low carbon technology uptake. In Energy Policy 97, pp. 13–26. DOI: 10.1016/j.enpol.2016.06.042.

McKenna, R.; Djapic, P.; Weinand, J.; Fichtner, W.; Strbac, G. (2017a): Assessing the implications of socioeconomic diversity for low carbon technology uptake in electrical distribution networks. In Applied Energy. DOI: 10.1016/j.apenergy.2017.07.089.

McKenna, Russell; Merkel, Erik; Fichtner, Wolf (2017b): Energy autonomy in residential buildings. A techno-economic model-based analysis of the scale effects. In Applied Energy 189, pp. 800–815. DOI: 10.1016/j.apenergy.2016.03.062.

McKenna , R. (2018): The double-edged sword of decentralized energy autonomy. In Energy Policy, in press.

Miller, N. (2013): Urban Form and Building Energy: Quantifying Relationships using a Multi-Scale Approach. PhD Thesis in Resource Management and Environmental Studies at the Faculty of Graduate Studies. Vancouver.

Milligan, Glenn W. (1980): An examination of the effect of six types of error perturbation on fifteen clustering algorithms. In Psychometrika 45 (3), pp. 325–342. DOI: 10.1007/BF02293907.

Milligan, Glenn W. (1981): A monte carlo study of thirty internal criterion measures for cluster analysis. In Psychometrika 46 (2), pp. 187–199. DOI: 10.1007/BF02293899.

Milligan, Glenn W.; Cooper, Martha C. (1985): An examination of procedures for determining the number of clusters in a data set. In Psychometrika 50 (2), pp. 159–179. DOI: 10.1007/BF02294245.

Milligan, Glenn W.; Cooper, Martha C. (1988): A study of standardization of variables in cluster analysis. In Journal of Classification 5 (2), pp. 181–204. DOI: 10.1007/BF01897163.

Miyamoto, Sadaaki; Abe, Ryosuke; Endo, Yasunori; Takeshita, Jun-ichi (2015): Ward method of hierarchical clustering for non-Euclidean similarity measures. In Mario Köppen (Ed.): Proceedings of the 2015 Seventh International Conference of Soft Computing and Pattern Recognition (SoCPaR 2015). Fukuoka, Japan, November 13-15, 2015. 2015 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR). Fukuoka, Japan. Machine Intelligence Research Labs; International Conference of Soft Computing and Pattern Recognition; SoCPaR. Piscataway, NJ: IEEE, pp. 60–63.

Müller, Matthias Otto; Stämpfli, Adrian; Dold, Ursula; Hammer, Thomas (2011): Energy autarky. A conceptual framework for sustainable regional development. In Energy Policy 39 (10), pp. 5800–5810. DOI: 10.1016/j.enpol.2011.04.019.

Müllner, Daniel (2016): Package "fastcluster". Available online at https://cran.r-project.org/web/packages/fastcluster/fastcluster.pdf, checked on 8/16/2017.

Murtagh, Fionn; Legendre, Pierre (2014): Ward's Hierarchical Agglomerative Clustering Method. Which Algorithms Implement Ward's Criterion? In J Classif 31 (3), pp. 274–295. DOI: 10.1007/s00357-014-9161-z.

Noiva, Karen; Fernández, John E.; Wescoat, James L. (2016): Cluster analysis of urban water supply and demand. Toward large-scale comparative sustainability planning. In Sustainable Cities and Society 27, pp. 484–496. DOI: 10.1016/j.scs.2016.06.003.

Office for National Statistics (2015): Methodology Note for the 2011 Area Classification for Output Areas, checked on 8/16/2017.

Osborne, Jason W. (2014): Best practices in exploratory factor analysis. [Louisville, Ky.]: [CreateSpace Independent Publishing Platform].

Peter, Stefan (2013): Modellierung einer vollständig auf erneuerbaren Energien basierenden Stromerzeugung im Jahr 2050 in autarken, dezentralen Strukturen. Edited by Umweltbundesamt.

Purkus, Alexandra; Barth, Volker (2011): Geothermal power production in future electricity markets—A scenario analysis for Germany. In Energy Policy 39 (1), pp. 349–357. DOI: 10.1016/j.enpol.2010.10.003.

Rae, Callum; Bradley, Fiona (2012): Energy autonomy in sustainable communities—A review of key issues. In Renewable and Sustainable Energy Reviews 16 (9), pp. 6497–6506. DOI: 10.1016/j.rser.2012.08.002.

Ratkowsky, D. A.; Lance, G. N. (1978): A Criterion for determining the number of groups in a classification. In Australian Computer Journal 10 (3), pp. 115–117.

Ratti, Carlo; Baker, Nick; Steemers, Koen (2005): Energy consumption and urban texture. In Energy and Buildings 37 (7), pp. 762–776. DOI: 10.1016/j.enbuild.2004.10.010.

Reiner Lemoine Institut (2013): Vergleich und Optimierung von zentral und dezentral orientierten Ausbaupfaden zu einer Stromversorgung aus Erneuerbaren Energien in Deutschland. Berlin. Available online at https://www.bvmw.de/fileadmin/pdf-archiv/Studie_zur_dezentralen_Energiewende.pdf.pdf, checked on 11/8/2017.

Requia, Weeberb J.; Roig, Henrique L.; Koutrakis, Petros; Adams, Matthew D. (2017): Modeling spatial patterns of traffic emissions across 5570 municipal districts in Brazil. In Journal of Cleaner Production 148, pp. 845–853. DOI: 10.1016/j.jclepro.2017.02.010.

Revelle, William (2017): Package 'psych'. Available online at https://cran.r-project.org/web/packages/psych/psych.pdf, checked on 8/16/2017.

Rode, Philipp; Keim, Christian; Robazza, Guido; Viejo, Pablo; Schofield, James (2014a): Cities and Energy. Urban Morphology and Residential Heat-Energy Demand. In Environ Plann B Plann Des 41 (1), pp. 138–162. DOI: 10.1068/b39065.

Rode, P.; Burdett, R.; Robazza, G.; Schofield, J.; Keim, C.; Avci, N.; Dimitrova, S.; Garcia, P. V.; Murshed, S. M.; Bahu, J.-M. (2014b): Cities and Energy: Urban Morphology and Heat Energy Demand. Final Report. LSE, EIFER, London, February 2014.

Rousseeuw, Peter J. (1987): Silhouettes. A graphical aid to the interpretation and validation of cluster analysis. In Journal of Computational and Applied Mathematics 20, pp. 53–65. DOI: 10.1016/0377-0427(87)90125-7.

Salvador, S.; Chan, P. (2004): Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In 16th IEEE International Conference on Tools with Artificial Intelligence. Proceedings: ICTAI 2004: 15-17 November, 2004, Boca Raton, Florida. 16th IEEE International Conference on Tools with Artificial Intelligence. Boca Raton, FL, USA, 15-17 Nov. 2004. International Conference on Tools for Artificial Intelligence; IEEE Computer Society. Los Alamitos, Calif: IEEE Computer Society, pp. 576–584.

Sarralde, Juan José; Quinn, David James; Wiesmann, Daniel; Steemers, Koen (2015): Solar energy and urban morphology. Scenarios for increasing the renewable energy potential of neighbourhoods in London. In Renewable Energy 73, pp. 10–17. DOI: 10.1016/j.renene.2014.06.028.

Scheffer, K. (2008): Vom Bioenergiedorf zur autonomen Solarenergieregion. In Solarzeitalter 20 (4), pp. 23–30.

Schmid, Eva; Knopf, Brigitte; Pechan, Anna (2016): Putting an energy system transformation into practice. The case of the German Energiewende. In Energy Research & Social Science 11, pp. 263–275. DOI: 10.1016/j.erss.2015.11.002.

Schmidt, J.; Schönhart, M.; Biberacher, M.; Guggenberger, T.; Hausl, S.; Kalt, G. et al. (2012): Regional energy autarky. Potentials, costs and consequences for an Austrian region. In Energy Policy 47, pp. 211–221. DOI: 10.1016/j.enpol.2012.04.059.

Schuler, Andreas; Weber, Christoph; Fahl, Ulrich (2000): Energy consumption for space heating of West-German households. Empirical evidence, scenario projections and policy implications. In Energy Policy 28 (12), pp. 877–894. DOI: 10.1016/S0301-4215(00)00074-4.

Schultz, Annett; Brandt (2016): Demographietypisiserung für Kommunen. Methodisches Vorgehen und empirische Befunde - Berichtsjahr 2013. Bertelsmann Stiftung. Bochum. Available online at https://www.wegweiser-kommune.de/documents/10184/10615/Methodik_Clusterung.pdf/05a1b137-7dbf-4bf4-828d-9a097a4f3805, checked on 8/2/2017.

Scott, A. J.; Symons, M. J. (1971): Clustering Methods Based on Likelihood Ratio Criteria. In Biometrics 27 (2), p. 387. DOI: 10.2307/2529003.

Shapiro, S. S.; Wilk, M. B. (1965): An Analysis of Variance Test for Normality (Complete Samples). In Biometrika 52 (3/4), p. 591. DOI: 10.2307/2333709.

Shapiro, Alexander; Berge, Jos M. F. ten (2002): Statistical inference of minimum rank factor analysis. In Psychometrika 67 (1), pp. 79–94. DOI: 10.1007/BF02294710.

Shipworth, Michelle; Firth, Steven K.; Gentry, Michael I.; Wright, Andrew J.; Shipworth, David T.; Lomas, Kevin J. (2010): Central heating thermostat settings and timing. Building demographics. In Building Research & Information 38 (1), pp. 50–69. DOI: 10.1080/09613210903263007.

Statistisches Bundesamt (2000): Erwerbstätige nach Wirtschaftsbereichen - Jahresdurchschnitt - (WZ2008) regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=638-61-4&levelindex=0&levelid=1504792196820&index=13, checked on 9/7/2017.

Statistisches Bundesamt (2003a): Bruttoinlandsprodukt/Bruttowertschöpfung (WZ 2008) - Jahressumme - regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2, checked on 9/7/2017.

Statistisches Bundesamt (2003b): Energieverwendung der Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden - Jahressumme - regionale Tiefe: Kreise und krfr. Städte. Edited by Statistisches Bundesamt. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=060-41-4&levelindex=0&levelid=1504791708820&index=1, checked on 9/7/2017.

Statistisches Bundesamt (2011a): Gemeindeverzeichnis - Gebietsstand: 31.12.2010 (Jahr). Alle politisch selbständigen Gemeinden in Deutschland nach Fläche, Bevölkerung, Bevölkerungsdichte und der Postleitzahl des Verwaltungssitzes der Gemeinde. Edited by Statistische Ämter des Bundes und der Länder. Wiesbaden. Available online at https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122010_Auszug_GV.html, checked on 9/7/2017.

Statistisches Bundesamt (2011b): "Lohn- und Einkommensteuer - Jahressumme -regionale Tiefe: Gemeinden, Samt-/Verbandsgemeinden". Edited by Statistischen Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=62C84E17497FA378AD49A9004D9A6DF6.reg2?operation=statistikAbruftabellen&levelindex=0&levelid=1504787362480&index=2, checked on 9/7/2017.

Statistisches Bundesamt (2014a): Bevölkerung. Edited by Statistische Ämter des Bundes und der Länder. Wiesbaden. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=62C84E17497FA378AD49A9004D9A6DF6.reg2?operation=statistikAbruftabellen&levelindex=0&levelid=1504787709962&index=2, checked on 9/7/2017.

Statistisches Bundesamt (2014b): Bruttoinlandsprodukt/Bruttowertschöpfung (WZ 2008) - Jahressumme - regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=previous&levelindex=2&levelid=1504791473323&levelid=1504791439317&step=1, checked on 9/7/2017.

Statistisches Bundesamt (2014c): Energieverwendung der Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden - Jahressumme - regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=060-41-4&levelindex=0&levelid=1504791708820&index=1, checked on 9/7/2017.

Statistisches Bundesamt (2014d): Gebäude und Wohnungen. Edited by Statistische Ämter des Bundes und der Länder. Wiesbaden. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=62C84E17497FA378AD49A9004D9A6DF6.reg2?operation=statistikenVerzeichnisNextStep&levelindex=0&levelid=1504786087209&index=3&structurelevel=3, checked on 9/7/2017.

Statistisches Bundesamt (2014e): Haushalte und Familien. Edited by Statistische Ämter des Bundes und der Länder. Wiesbaden. Available online at https://www.regionalstatistik.de/genesis/online, checked on 9/7/2017.

Statistisches Bundesamt (2014f): Sozialversicherungspflichtig Beschäftigte am Arbeits- und Wohnort sowie Ein- und Auspendler über Gemeindegrenzen -Stichtag 30.06. - regionale Tiefe: Gemeinden. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=254-14-5&levelindex=1&levelid=1504790893929&index=5, checked on 9/7/2017.

Statistisches Bundesamt (2014g): Umsatzsteuer: Umsatzsteuerpflichtige, steuerbarer Umsatz für Lieferungen und Leistungen nach Wirtschaftsabschnitten - Jahr - regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=377-31-4&levelindex=0&levelid=1504791907481&index=7, checked on 9/7/2017.

Statistisches Bundesamt (2015a): Erwerbstätige nach Wirtschaftsbereichen-Jahresdurchschnitt - (WZ2008) regionale Tiefe: Kreise und krfr. Städte. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=638-61-4&levelindex=1&levelid=1504791242569&index=3, checked on 9/7/2017.

Statistisches Bundesamt (2015b): Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden: Betriebe, Beschäftigte - Stichtag: 30.09. - regionale Tiefe: Gemeinden, Samt-/Verbandsgemeinden. Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=DEE1DDC2D1089B9F9908854D96E648AC.reg2?operation=abruftabelleAbrufen&selectionname=001-03-5&levelindex=0&levelid=1504792334773&index=2, checked on 9/7/2017.

Statistisches Bundesamt (2016a): "Arbeitslose nach ausgewählten Personengruppen - Jahresdurchschnitt -regionale Ebenen". Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=62C84E17497FA378AD49A9004D9A6DF6.reg2?operation=abruftabelleAbrufen&selectionname=659-21-5&levelindex=1&levelid=1504788083861&index=2, checked on 9/7/2017.

Statistisches Bundesamt (2016b): "Bodenfläche nach Art der tatsächlichen Nutzung - Stichtag 31.12. - regionale Ebenen". Edited by Statistische Ämter des Bundes und der Länder. Available online at https://www.regionalstatistik.de/genesis/online/data;jsessionid=62C84E17497FA378AD49A9004D9A6DF6.reg2?operation=abruftabelleAbrufen&selectionname=449-01-5&levelindex=1&levelid=1504788598036&index=2, checked on 9/7/2017.

Statistisches Bundesamt (2017a): Bruttostromerzeugung in Deutschland für 2014 bis 2016. Available online at https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/Energie/Erzeugung/Tabellen/Bruttostromerzeugung.html, checked on 9/29/2017.

Statistisches Bundesamt (2017b): Gemeindeverzeichnis - Gebietsstand: 30.06.2017 (2. Quartal). Alle politisch selbständigen Gemeinden (mit Gemeindeverband) in Deutschland nach Fläche, Bevölkerung (auf Grundlage des Zensus 2011), Bevölkerungsdichte und der Postleitzahl des Verwaltungssitzes der Gemeinde. Edited by Statistische Ämter des Bundes und der Länder. Wiesbaden. Available online at https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugQ/AuszugGV2QAktuell.html, checked on 9/7/2017.

Statistisches Landesamt Baden-Württemberg (2009): Weit auseinander und doch beisammen. In Statistisches Monatsheft Baden-Württemberg (06/2009), pp. 51–53. Available online at https://www.statistik-bw.de/Service/Veroeff/Monatshefte/PDF/Beitrag09_06_11.pdf, checked on 8/2/2017.

Steemers, Koen (2003): Energy and the city. Density, buildings and transport. In Energy and Buildings 35 (1), pp. 3–14. DOI: 10.1016/S0378-7788(02)00075-0.

Su, Shi; Lu, Hai; Zhang, Liang; Alanne, Kari; Yu, Zitao (2017): Solar energy utilization patterns for different district typologies using multi-objective optimization. A comparative study in China. In Solar Energy 155, pp. 246–258. DOI: 10.1016/j.solener.2017.06.036.

Tibshirani, Robert; Walther, Guenther; Hastie, Trevor (2001): Estimating the number of clusters in a data set via the gap statistic. In J Royal Statistical Soc B 63 (2), pp. 411–423. DOI: 10.1111/1467-9868.00293.

Umweltbundesamt (2016): Branchenabhängiger Energieverbrauch des verarbeitenden Gewerbes. Available online at https://www.umweltbundesamt.de/daten/industrie/branchenabhaengiger-energieverbrauch-des#textpart-1, checked on 8/11/2017.

Umweltbundesamt; BMWi (2017): Energieeffizienz in Zahlen. Edited by Bundesministerium für Wirtschaft und Energie (BMWi). Available online at http://www.bmwi.de/Redaktion/DE/Publikationen/Energie/energieeffizienz-in-zahlen.pdf?__blob=publicationFile&v=10, checked on 11/2/2017.

Unternährer, Jérémy; Moret, Stefano; Joost, Stéphane; Maréchal, François (2017): Spatial clustering for district heating integration in urban energy systems. Application to geothermal energy. In Applied Energy 190, pp. 749–763. DOI: 10.1016/j.apenergy.2016.12.136.

Urquizo, Javier; Calderón, Carlos; James, Philip (2017): Metrics of urban morphology and their impact on energy consumption. A case study in the United Kingdom. In Energy Research & Social Science 32, pp. 193–206. DOI: 10.1016/j.erss.2017.03.011.

Vendramin, Lucas; Campello, Ricardo J. G. B.; Hruschka, Eduardo R. (2010): Relative clustering validity criteria. A comparative overview. In Statistical Analysis and Data Mining 3. DOI: 10.1002/sam.10080.

Volz, R. (2012): Bedeutung und Potenziale von Energiegenossenschaften in Deutschland. In Informationen zur Raumentwicklung 9 (10), pp. 515–524.

Wall, Wilhelm (2016): Energetisch vergleichbare Städtegruppen - eine gesamtheitliche Clusteranalyse und Clusterauswahl deutscher kreisfreier Städte auf Basis der typischen Verbrauchssektoren und sozio-energetischer Indikatoren. Bochum: Selbstverlag des Lehrstuhls Energiesysteme und Energiewirtschaft.

Wei, Shen; Jones, Rory; Wilde, Pieter de (2014): Driving factors for occupant-controlled space heating in residential buildings. In Energy and Buildings 70, pp. 36–44. DOI: 10.1016/j.enbuild.2013.11.001.

Wimmer, D.; Heinemann, C.; Bauknecht, D. (2014): Die Auswirkung räumlich verteilter Windstromproduktion auf den Flexibilitätsbedarf im deutschen Stromsystem. In Energiewirtschaftliche Tagesfragen 64 (12), pp. 32–35.

Wirth, H. (2016): Recent Facts about Photovoltaics in Germany. Fraunhofer ISE, Freiburg, October 14, 2016.

Woldeamanuel, Mintesnot G.; Cyganski, Rita; Schulz, Angelika; Justen, Andreas (2009): Variation of households' car ownership across time. Application of a panel data model. In Transportation 36 (4), pp. 371–387. DOI: 10.1007/s11116-009-9210-3.

Woyke, Wolfgang; Forero, Mario (2014): Methoden zur Bewertung regionaler Energieautarkie. Beitrag auf dem 13. Symposium Energieinnovation. Graz, Österreich.

Wurm, Michael; Taubenbock, Hannes; Roth, Achim; Dech, Stefan (2009): Urban structuring using multisensoral remote sensing data: By the example of the German cities Cologne and Dresden. In : Joint Urban Remote Sensing Event, 2009. 20 - 22 May 2009, Shanghai, China ; [the 5th GRSS/ISPRS Workshop on Data Fusion and Remote Sensing over Urban Areas (URBAN 2009) and the 7th International Urban Remote Sensing Conference (URS 2009)]. 2009 Joint Urban Remote Sensing Event. Shanghai, China. Geoscience and Remote Sensing Society; International Society for Photogrammetry and Remote Sensing; Joint Urban Remote Sensing Event; Urban Remote Sensing Joint Event; GRSS/ISPRS Workshop on Data Fusion and Remote Sensing over Urban Areas; URBAN; International Urban Remote Sensing Conference; URS. Piscataway, NJ: IEEE, pp. 1–8.

Yang, Junjing; Ning, Chao; Deb, Chirag; Zhang, Fan; Cheong, David; Lee, Siew Eang et al. (2017): k-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement. In Energy and Buildings 146, pp. 27–37. DOI: 10.1016/j.enbuild.2017.03.071.

Yong, An Gie; Pearce, Sean (2013): A Beginner's Guide to Factor Analysis. Focusing on Exploratory Factor Analysis. In Tutorials in Quantitative Methods for Psychology 09 (2), pp. 79–94.

Zhang, Fuzhen (2011): Positive Semidefinite Matrices. In Fuzhen Zhang (Ed.): Matrix Theory. Basic Results and Techniques. 2. ed. New York, NY: Springer Science+Business Media LLC (Universitext), pp. 199–252.

A. Appendix

Table 8: Indicators for cluster analysis and the associated units and references.

Indicator	References
Indicators of the Consumption Sector Private Households	
Population development between 2010 and 2015 [%]	(Statistisches Bundesamt 2011a, 2017b) and own calculation
Living space per person [m²]	(Statistisches Bundesamt 2017b, 2014d) and own calculation
Share of single-person households in total number of households [%]	(Statistisches Bundesamt 2014e) and own calculation
Average household size [Number of persons]	(Statistisches Bundesamt 2014e, 2017b) and own calculation
Household density [Households per km²]	(Statistisches Bundesamt 2017b, 2014e) and own calculation
Share of owner-occupied apartments in total number of apartments [%]	(Statistisches Bundesamt 2014d) and own calculation
Income per household [k€]	(Statistisches Bundesamt 2014e, 2011b) and own calculation
Share of over 65-year-olds in total population [%]	(Statistisches Bundesamt 2014a) and own calculation
Unemployment rate [%]	(Statistisches Bundesamt 2017b, 2014a, 2016a) and own calculation
Share of settlement and traffic area in total area [%]	(Statistisches Bundesamt 2016b) and own calculation
Heating days (long-term average)	(Institut für Wohnen und Umwelt 2017)
Heating degree days (long-term average)	(Institut für Wohnen und Umwelt 2017)
Degree day number (long-term average)	(Institut für Wohnen und Umwelt 2017)
Share of buildings with heating systems based on district heating [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings with heating systems not based on district heating [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings without heating system [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built before 1919 in total building stock (X14) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1919 and 1949 in total building stock (X15) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1950 and 1959 in total building stock (X16) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1960 and 1969 in total building stock (X17) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1970 and 1979 in total building stock (X18) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1980 and 1989 in total building stock (X19) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 1990 and 1999 in total housing stock (X20) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built between 2000 and 2005 in total building stock (X21) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of buildings built from 2006 onward in total building stock (X22) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of detached houses in total building stock (X23) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of semi-detached houses in total building stock (X24) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of terraced houses in total building stock (X25) [%]	(Statistisches Bundesamt 2014d) and own calculation
Share of “other building types” in total building stock (X26) [%]	(Statistisches Bundesamt 2014d) and own calculation

Indicators of the Transport consumption sector	
Number of motor vehicles per 1,000 inhabitants	(Statistisches Bundesamt 2017b; Kraftfahrt-Bundesamt 2017a) and own calculation
Number of cars per 1,000 inhabitants	(Statistisches Bundesamt 2017b; Kraftfahrt-Bundesamt 2017a) and own calculation
Share of diesel vehicles in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Share of petrol vehicles in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Share of gas vehicles in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Share of hybrid vehicles in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Share of electric vehicles in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Share of “other vehicle type” in vehicle fleet [%]	(Kraftfahrt-Bundesamt 2017b) and own calculation
Population density [Inhabitants per km²]	(Statistisches Bundesamt 2017b) and own calculation
Share of 18-64-year-olds in the total population [%]	(Statistisches Bundesamt 2014a) and own calculation
Share of commuters in total workforce [%]	(Statistisches Bundesamt 2014f) and own calculation
Indicators of the Consumption Sector Industry and Commercial	
Share of employment in the industrial sector [%]	(Statistisches Bundesamt 2015a) and own calculation
Share of employment in the commercial sector [%]	(Statistisches Bundesamt 2015a) and own calculation
Energy productivity of manufacturing industry [€/GJ]	(Statistisches Bundesamt 2014b, 2014c) and own calculation
Energy intensity of manufacturing industry [MJ/€]	(Statistisches Bundesamt 2014b, 2014c) and own calculation
Productivity level of manufacturing industry [€/GJ]	(Statistisches Bundesamt 2014b, 2014c) and own calculation
Specific energy consumption of manufacturing industry [MJ/€]	(Statistisches Bundesamt 2014b, 2014c) and own calculation
Share of industrial sales tax in total sales tax [%]	(Statistisches Bundesamt 2014g) and own calculation
Share of commercial sales tax in total sales tax [%]	(Statistisches Bundesamt 2014g) and own calculation
Development of employment share in the industrial sector [%]	(Statistisches Bundesamt 2015a, 2000) and own calculation
Development of employment share in the commercial sector [%]	(Statistisches Bundesamt 2015a, 2000) and own calculation
Development of energy intensity in manufacturing industry from 2003 to 2014 [%]	(Statistisches Bundesamt 2014b, 2014c, 2003a, 2003b) and own calculation
Number of manufacturing enterprises per 1,000 households	(Statistisches Bundesamt 2015b) and own calculation
Indicators of the potential for renewable energies	
Achievable hydrothermal temperature [°C]	(Agemar 2017) and own calculation
Necessary hydrothermal drilling depth [m]	Own calculation
Technical PV potential per inhabitant [kWh/y]	(Mainzer et al. 2014; Statistisches Bundesamt 2017b)
Technical PV potential per km² [MWh/y]	(Mainzer et al. 2014; Statistisches Bundesamt 2017b)
Technical wind potential per inhabitant [MWh/y]	(McKenna et al. 2014a; Statistisches Bundesamt 2017b) and own calculation
Technical wind potential per km² [MWh/y]	(McKenna et al. 2014a; Statistisches Bundesamt 2017b) and own calculation
Share of forest and agricultural land in total area [%]	(Statistisches Bundesamt 2016b) and own calculation

Table 9: Number of clusters resulting from 26 different procedures and evaluation of the procedures. A high number of clusters in the "evaluation of the procedure" column means a number of more than 4 clusters.

Index	Number	Evaluation of the procedure
"ch" (Calinski & Harabasz 1974)	2	Poor with a high number of clusters. Often prefers 2 cluster solutions (Islam et al. 2016; Arbelaitz et al. 2013; Vendramin et al. 2010).
"duda" (Duda & Hart 1973)	10	Good with a high number of clusters (Milligan, Cooper 1985; Islam et al. 2016). Second best procedure in (Milligan & Cooper 1985).
"pseudot2" (Duda & Hart 1973)	10	-
"cindex" (Hubert & Levin 1976)	6	Determines the optimum number of clusters +/-1 with a probability of only 50% (Islam et al. 2016). Poor with a high number of clusters (Arbelaitz et al. 2013).
"beale" (Beale 1969)	2	Poor with a high number of clusters (Arbelaitz et al. 2013).
"ptbiserial" (Milligan 1980, 1981)	10	A high number of clusters is often underestimated (Milligan & Cooper 1985).
"db" (Davies & Bouldin 1979)	10	In (Arbelaitz et al. 2013) the third-best index with a high number of clusters, but low success rate with a high number of clusters (Milligan & Cooper 1985; Arbelaitz et al. 2013).
"frey" (Frey & van Groenewoud 1972)	1	Result contradicts the cluster idea because of the number of clusters <2. The Number of clusters is rather underestimated with a high number of clusters (Milligan & Cooper 1985).
"hartigan" (Hartigan 1975)	5	Works well with a small number of indicators (Tibshirani et al. 2001; Albatineh & Niewiadomska-Bugaj 2011). Poor with a high number of clusters (Milligan & Cooper 1985).
"ratkowsky" (Ratkowsky & Lance 1978)	8	Poor with a high number of clusters (Milligan & Cooper 1985).
"scott" (Scott & Symons 1971)	3	Poor with a high number of clusters (Milligan & Cooper 1985).
"marriot" (Marriott 1971)	7	Poor with a high number of clusters (Milligan & Cooper 1985).
"ball" (Ball, Hall 1965)	3	Poor with a high number of clusters (Milligan & Cooper 1985).
"trcovw" (Milligan & Cooper 1985)	3	Poor with a high number of clusters (Milligan & Cooper 1985).
"tracew" (Milligan & Cooper 1985)	5	Poor with a high number of clusters (Milligan & Cooper 1985).
"friedman" (Friedman & Rubin 1967)	3	Poor with a high number of clusters (Milligan & Cooper 1985).
"mcclain" (McClain & Rao 1975)	2	Poor with a high number of clusters (Milligan & Cooper 1985).
"rubin" (Friedman & Rubin 1967)	8	Poor with a high number of clusters (Milligan & Cooper 1985).
"kl" (Krzanowski & Lai 1988)	3	Identifies only 40-50% of the clusters (Albatineh & Niewiadomska-Bugaj 2011; Islam et al. 2016).
"silhouette" (Rousseeuw 1987)	3	Poor with a high number of clusters (Islam et al. 2016; Arbelaitz et al. 2013).
"gap" (Tibshirani et al. 2001)	2	Poor with a high number of clusters (Islam et al. 2016).
"dindex" (Lebart et al. 2002)	5	-
"dunn" (Dunn 1974)	9	Poor with a high number of clusters (Arbelaitz et al. 2013).
"hubert" (Hubert & Arabie 1985)	4	-
"sdindex" (Halkidi et al. 2000)	2	-
"sdbw" (Halkidi & Vazirgiannis 2001)	7	Poor with a high number of clusters (Arbelaitz et al. 2013).

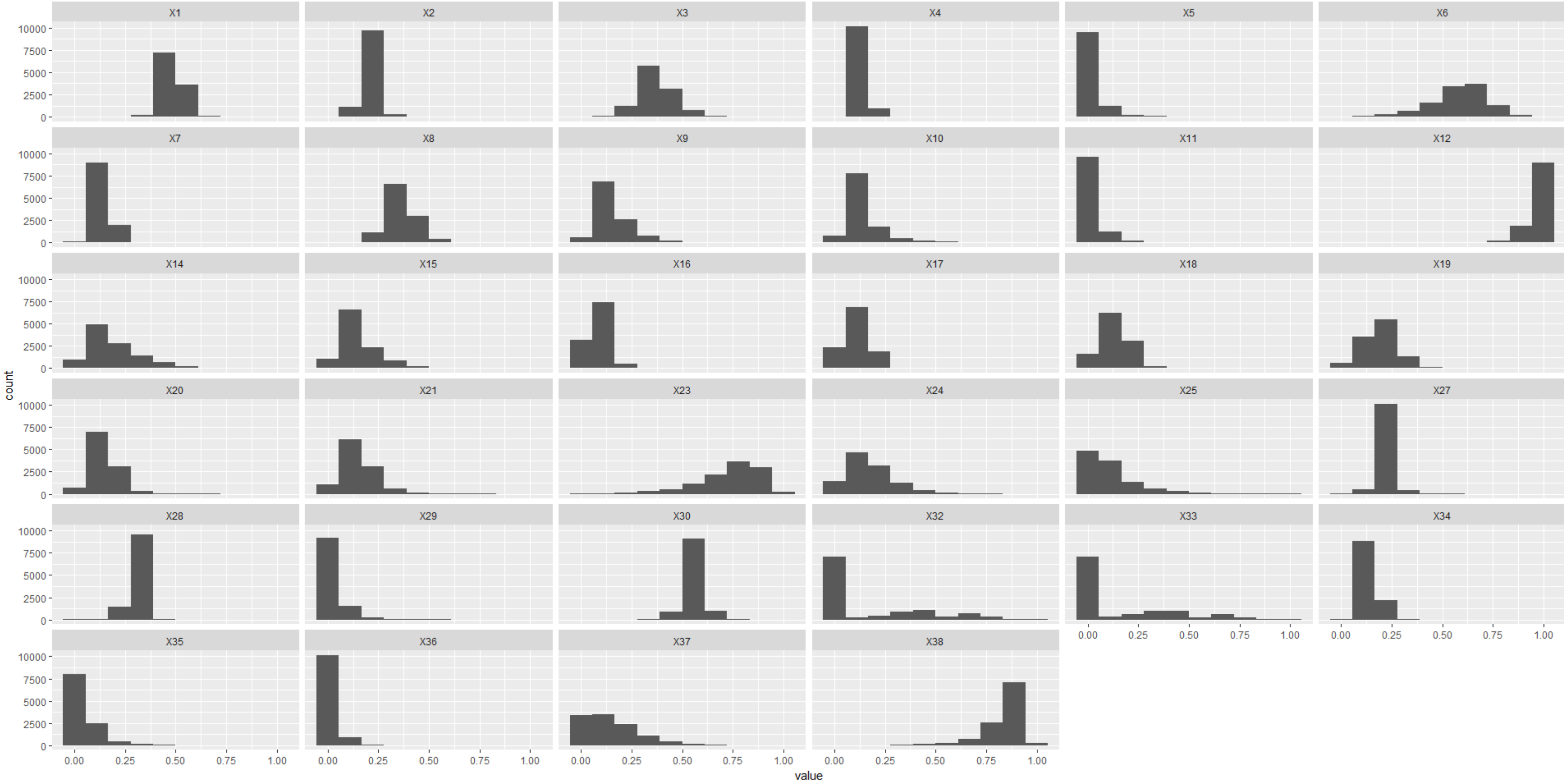


Figure 7: Distributions of indicator values.

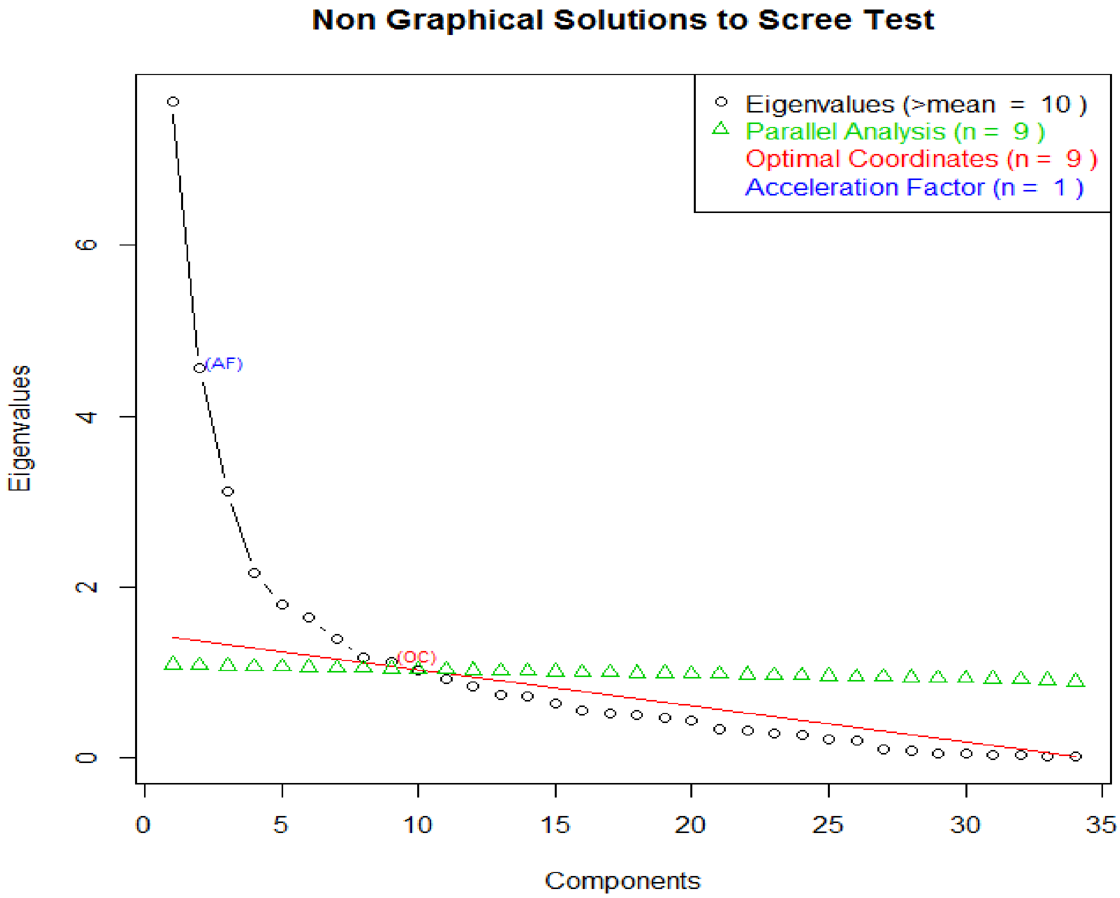


Figure 8: Results in determining the number of factors.

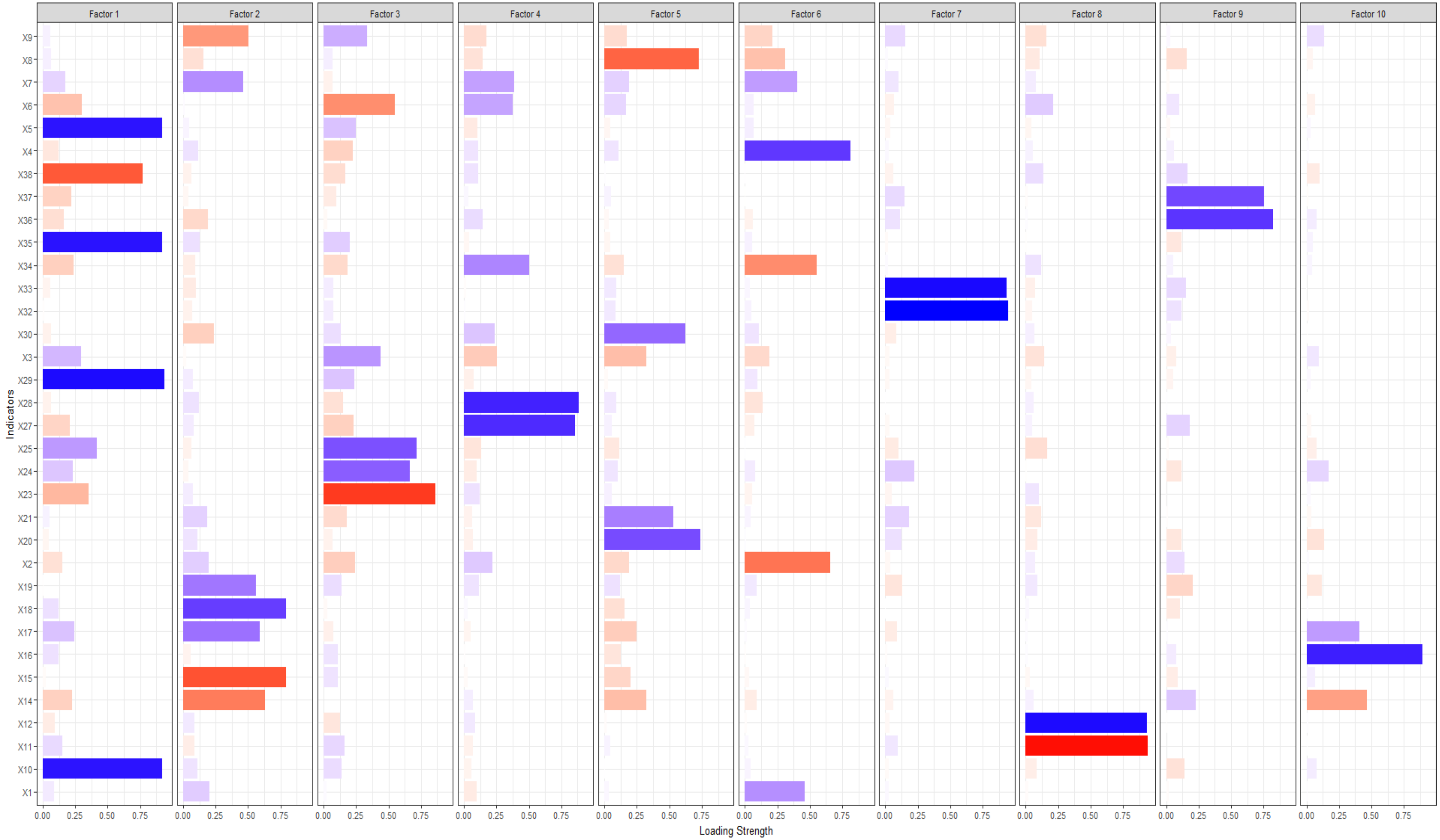


Figure 9: Factor loadings in each of the ten factors.

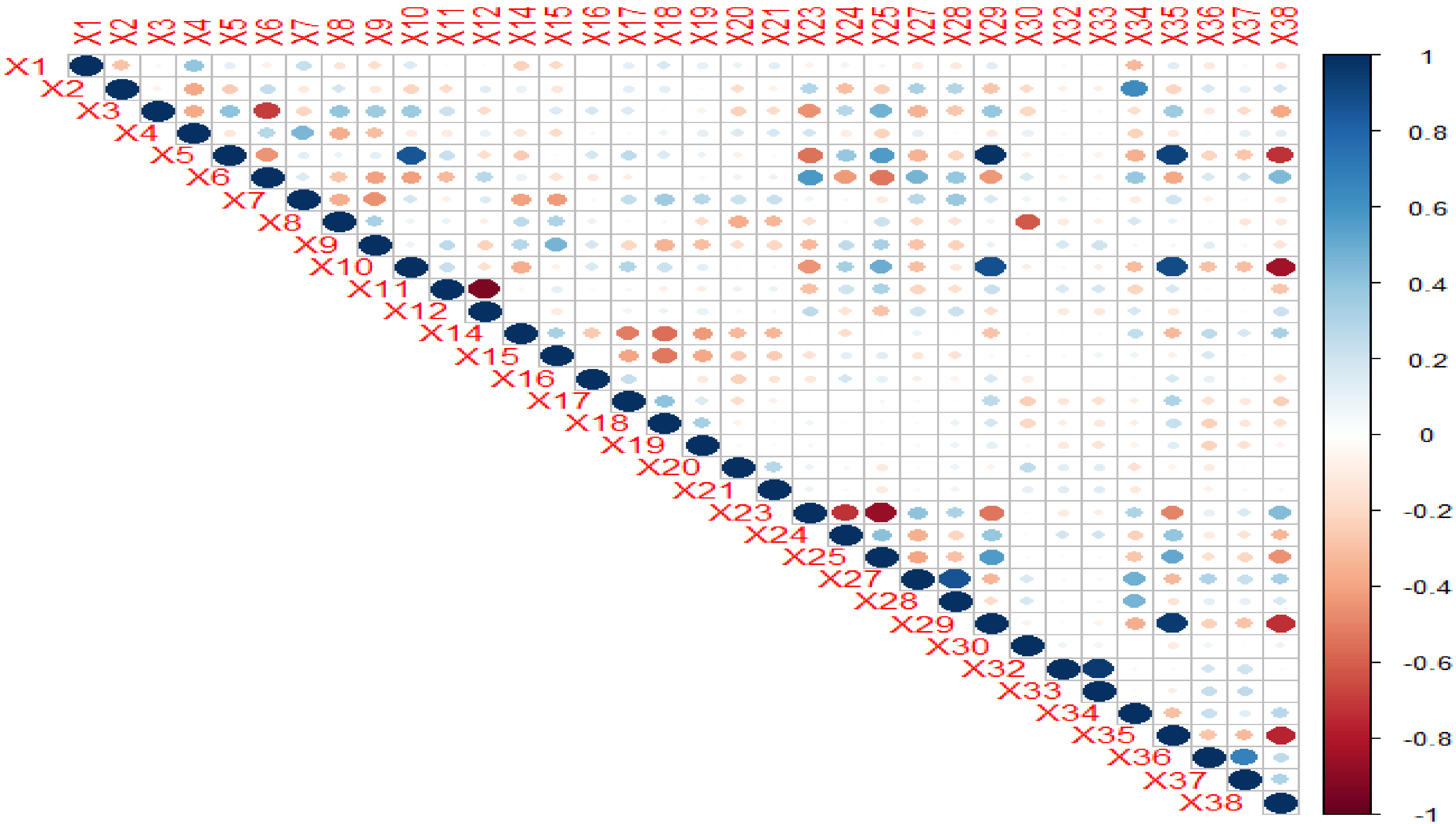


Figure 10: Correlation matrix of the indicator values.

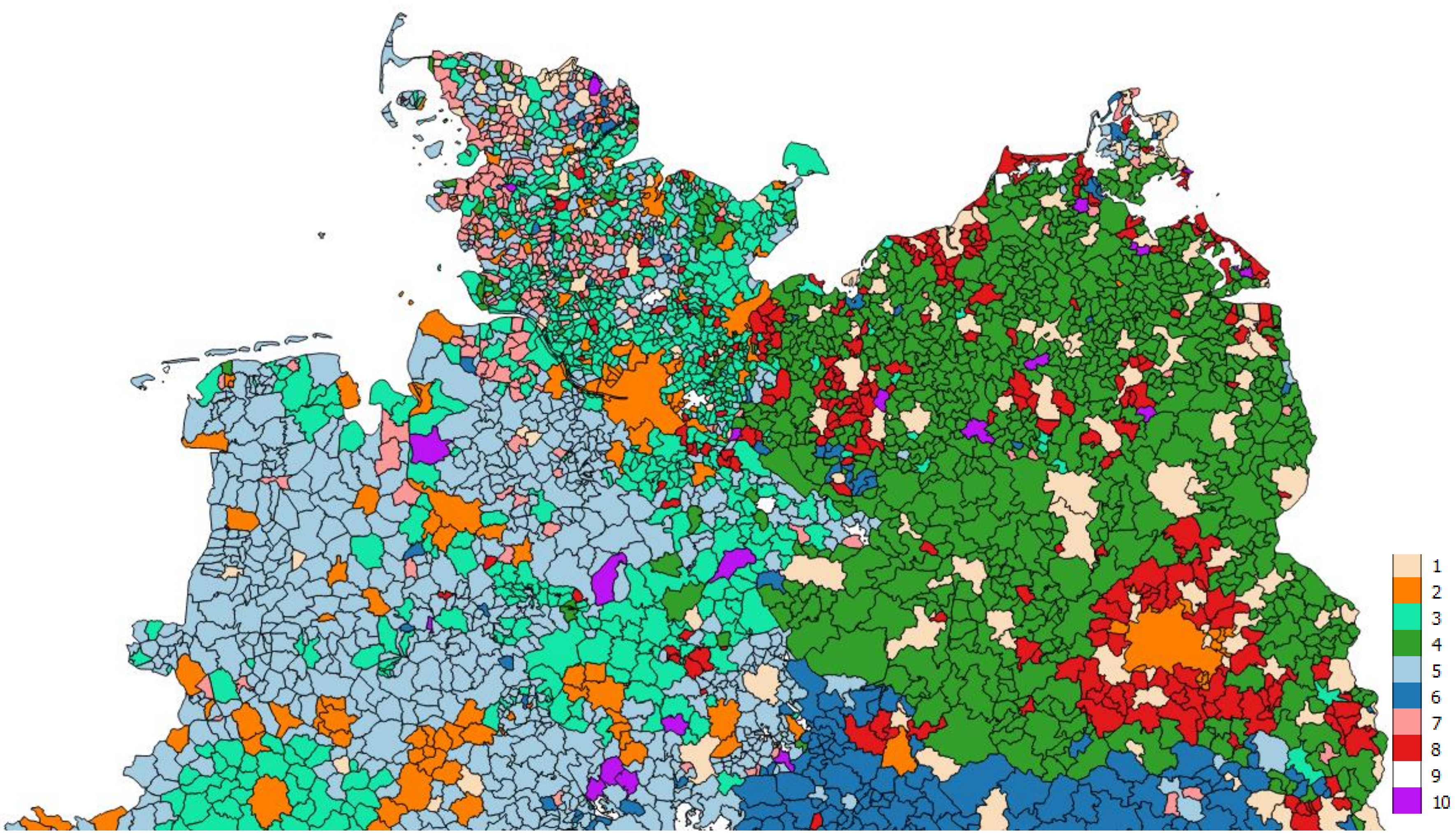


Figure 11: Illustration of the northern German municipalities with their allocation in the 10 cluster solution.

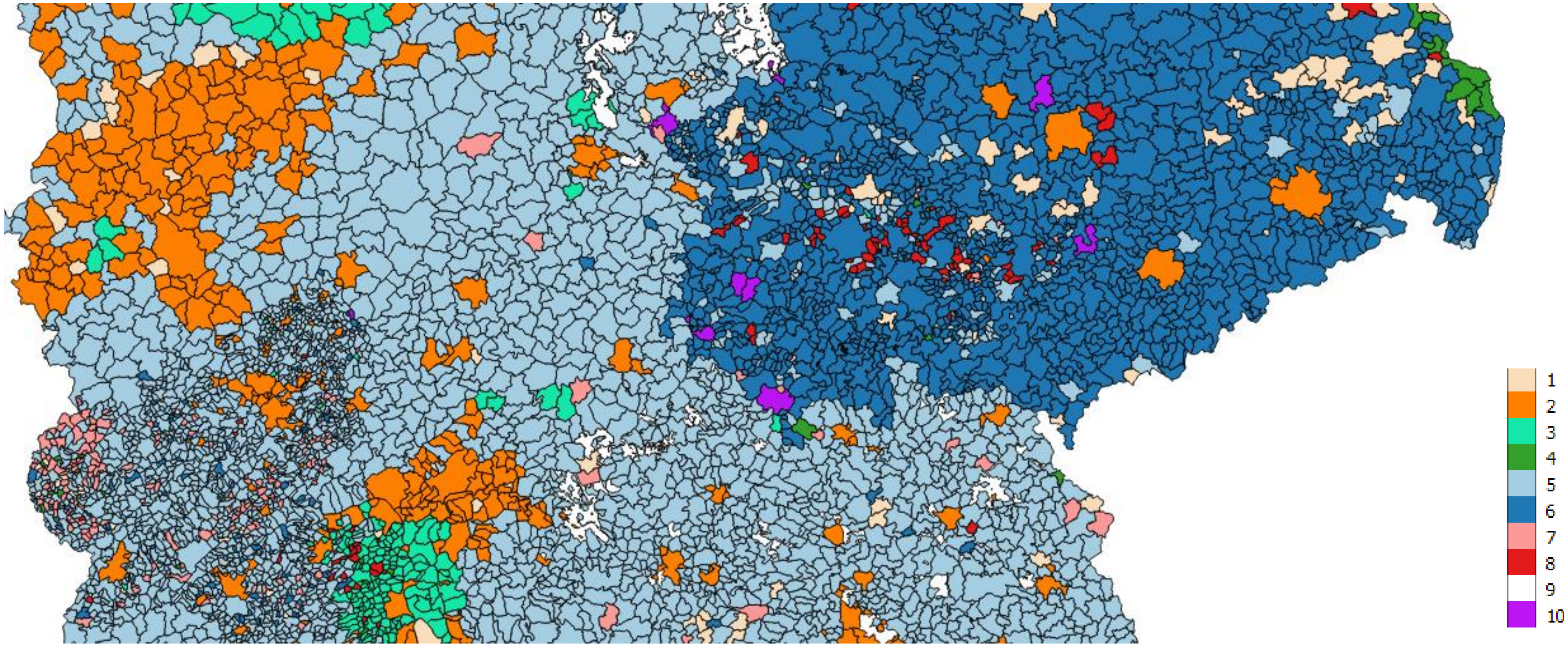


Figure 12: Illustration of the central German municipalities with their allocation in the 10 cluster solution.

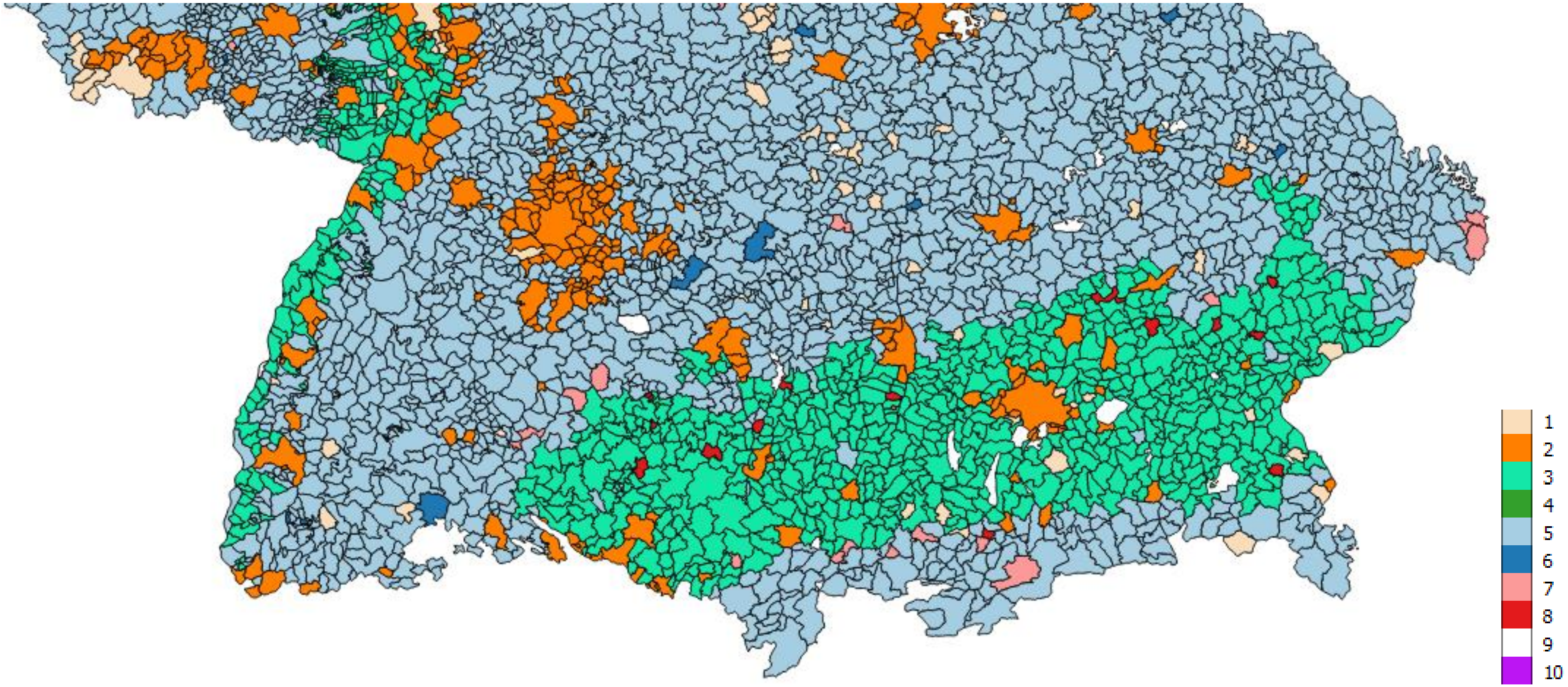


Figure 13: Illustration of the southern German municipalities with their allocation in the ten cluster solution

Impressum

Karlsruher Institut für Technologie

Institut für Industriebetriebslehre und Industrielle Produktion (IIP)
Deutsch-Französisches Institut für Umweltforschung (DFIU)

Hertzstr. 16
D-76187 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Production and Energy
No. 26, January 2018

ISSN 2196-7296