Programme Area: Distributed Energy

Project: Macro DE

Title: Energy Demand Analysis in Great Britain

Abstract:

This deliverable is numbers 2 and 3 of 3 in Work Package 2. It builds on the approach for calculating energy demand as described in deliverable D2.0. The outcome of the work described in this deliverable is the creation of 946 zones of demand which account for 55% of the GB energy demand across 4% of its geographical area. From these 946 zones of demand, 20 ‘characteristic zones of demand’ have been identified. Between them these zones have been shown to be suitable for characterising the energy demand (both heat and electricity) for GB and will be used in the development of a benefits case for DE in GB and the identification of technology development opportunities which will facilitate the up-take of macro scale DE. These two final points will be addressed in work package 5 of the project.

Context:

This project quantified the opportunity for Macro level Distributed Energy (DE) across the UK and accelerate the development of appropriate technology by 2020 for the purposes of significant implementation by 2030. The project studied energy demand such as residential accommodation, local services, hospitals, business parks and equipment, and is developing a software methodology to analyse local combinations of sites and technologies. This enabled the design of optimised distributed energy delivery solutions for these areas. The project identified a number of larger scale technology development and demonstration projects for the ETI to consider developing. The findings from this project is now being distilled into our Smart Systems and Heat programme. The ETI acknowledges that the project was undertaken and reports produced by Caterpillar, EDF, and the University of Manchester.

ETI Macro Distributed Energy Project

Project task 2.3

Energy Demand Analysis in GB

A report prepared for the Energy Technologies Institute

Philipp Meidl, EDF-EIFER
Maria Sipowicz, EDF-EIFER
Syed Monjur Murshed, EDF-EIFER
Stephanie Jumel, EDF R&D
Megan Jobson, University of Manchester
Gbemi Oluleye, University of Manchester
Ian O'Hanlon, EDF Energy
Kevin McKoen, EDF-EIFER
Niklas Griessbaum EDF-EIFER
Alexandru Nichersu, EDF-EIFER

revision history:

version 1.1
March 16, 2012

Executive Summary

The ETI Macro DE project is a feasibility study, funded by the Energy Technologies Institute, to broadly assess the potential for macro scale distributed energy in GB. The aim of work package 2 in this project was to calculate the temporal thermal and electrical energy demand for the residential and tertiary sectors in Great Britain (WP 2.1) together with a study about the potential use of excess heat of GB's industries (WP 2.2). As such a detailed study has not been executed for GB before, the method had to be developed first, which was presented in a trial on Harrogate (see (McKoen, Koch et al. 2011)), and the necessary data sets needed to be defined and searched for. Therefore, the basis for the calculations was set to the level of MLSOAs, because these are big enough to use standardised numbers and parameters and for which statistical data on housing, population, demand, etc. were available. The idea was to merge on average five MLSOAs together to create one distributed energy zone (DE zone). It was estimated that about 1000 zones will be created by this method, which should be grouped together into up to 25 classes afterwards. Out of each class one characteristic zone was to be chosen to represent the whole class in the benefit calculations for GB in this project. Generally, it can be concluded that the knowledge about the tertiary demand is less that about residential demand. Thus, the level of confidence is stronger on the calculations about the latter one. More research on the demand of the tertiary sector could lead to a refinement of the method developed in this project.

In the end, 946 DE zones have been created in WP 2.1, which were further grouped into 20 classes. This report describes the development of the zones and the corresponding classes step by step. These 5 tasks have been defined as:

1) Calculation of the energy demand of all 8429 MLSOAs (residential and tertiary, heat and electrical demand) (Energy Demand Calculation)
2) Creation of DE zones (Clustering)
3) Calculation of load curves and time bands for all DE Zones (Temporal Demand Calculation)
4) Grouping the DE zones into up to 25 classes (Classification)
5) Verification of results
    a) Distinctiveness between classes
    b) Homogeneity of zones within each class

These tasks plus the waste heat study (Chapter 6 —) form the frame of this report. The results of this work will be used in work package 5 for the GB benefits case studies. These 6 chapters of this report are briefly summarized on the following pages.

Chapter 1 — Energy Demand Calculation

Energy Demand Database

The demand database contains all of the data required for the calculation of the GB energy demand. It includes indicators, benchmarks and conversion factor data for residential and tertiary sectors[1] as well as administrative boundary data (e.g. postcode, MLSOA, GOR). That have all been carefully analysed and structured to prepare a demand database.

A consistent demand database is needed to work with the large volume (all of GB) and high resolution (MLSOA) of demand calculation.

Finally, the database output can be integrated into the GIS data catalogue for visualisation of the results and preparation of maps.

Table 0-1 explains the lists of data and their sources that are required to perform the annual energy demand calculation. Many of the datasets have already been described in the trial demand analysis report (see (McKoen, Koch et al. 2011)) whereas some additional data (especially the EDF Energy weather data, Marketsafe data) have been collected to improve the demand calculation method.

Sector	List of tables	Source
Geographic data	Country boundary	Office of National Statistic (ONS)
	GOR boundary	ONS
	GSP regions	EDF Energy
	MLSOA IGZ boundary	ONS
	Postcode boundary	NSPD
Residential sector	Housing stock	EDF Energy
	Age class	EDF Energy
	Housing type	EDF Energy
	Floor area	EDF Energy
	Residential Thermal Benchmark	EDF Energy
	Residential Electric Demand	(DECC 2008)
Tertiary sector	Company information	Marketsafe data (Creditsafe Business Solution Limited)
	Employee numbers	
	SIC code	
	Post code	
	Premise type	
	Standard Industrial Classification SIC2003	ONS
	TM46	Chartered Institution of Building Services Engineers (CIBSE)
	Conversion	Own Assumption

[1] Information on these data sets has been included in McKoen, K., A. Koch, et al. (2011). Development of a methodology to calculate energy demand. Karlsruhe, Paris, EIFER. **2.0.1**.

	Tertiary Thermal Benchmark	CIBSE
	Tertiary Electric Benchmark	CIBSE

Table 0-1: List of required datasets for the demand database

Improvement of energy demand calculation

The annual thermal and electricity energy demand calculation for the residential and tertiary sectors in GB are carried out based on the methodology explained in the trial report (see Equation 0-1 and (McKoen, Koch et al. 2011)).

$$Annual\ Energy\ Demand = \sum_{Sector} \sum_{Sites} (indicator \times factor \times benchmarks)$$

Equation 0-1: Annual energy demand calculation

However, in this report, energy demand calculation methods are refined with the availability of additional granular datasets. Compared to the trial report, the following improvements have been made:

- Incorporation of GB wide residential and tertiary indicator data to calculate the energy demand
- Improvement of thermal benchmarks by applying 'weather adjustment' with EDF Energy weather scenario data (see section 0)
- Modification of demand estimation method, especially with the redistribution of number of employees in the tertiary sector (see section 1.2.4)

Description of indicator data

The number of residential buildings within each MLSOA is used as an indicator to calculate residential thermal energy demand. The detailed housing stock data in GB was collected from the EDF Energy's customer research team. It is categorised according to five housing types and seven age classes (see (McKoen, Koch et al. 2011)).

The number of employees working at tertiary sites in GB is considered as the indicator data for tertiary energy demand. This data has been bought from the Creditsafe Business Solution Limited which offers credit rating related services and data on 4.4 million GB companies in all tertiary sectors. It also includes information on company address, SIC, link to parent company, etc. This data is categorised into two main business activities- limited company and non limited company. A detailed description of the dataset is given in (Creditsafe 2010).

Weather adjustment of benchmark data

Weather adjustment of benchmarks was performed using a formula specified in CIBSE's TM46 guideline (see (CIBSE 2008)).

In the trial demand analysis (see (McKoen, Koch et al. 2011)), weather adjustment to the 35 residential thermal energy benchmarks was based on a recent year. In this report, the weather compensation was carried out based on representative regional weather scenario data for a typical year provided by EDF Energy's meteorology team.

Residential energy demand calculation method

The residential energy demand is calculated following Equation 0-1. However, in this report, the residential thermal energy demand calculation method is improved, considering GB wide EDF Energy's housing stock data and weather adjusted benchmarks (see section 0). Residential electricity demand is extracted from DECC energy consumption data for standard and economy tariff (see (DECC 2008)). The detail description of the basic method is explained in Macro DE trial report (see (McKoen, Koch et al. 2011)).

Tertiary energy demand calculation method

The tertiary energy demand is also calculated based on Equation 0-1 and as well explained in (McKoen, Koch et al. 2011) The indicator data, i.e. Marketsafe data on number of employees for the limited and non limited companies across the GB is improved using a redistribution of employee method. Moreover, base thermal and electrical benchmarks are corrected with latest weather adjusted benchmarks. The distribution of employee method is explained in section 1.4.1 whereas the weather adjustment procedure is described in section 0.

Results

Total residential and tertiary thermal energy demand in GB is about 458,392 GWh / year. Figure 0-1 describes the geographic patterns of the thermal energy demand density at the MLSOA/IGZ, where it appears that energy density is higher in the big cities.

The total electricity demand in the GB is about 198,747 GWh / year, where residential electricity demand is about 115,908 GWh. Figure 0-2 describes the electricity demand at the MLSOA/IGZ.

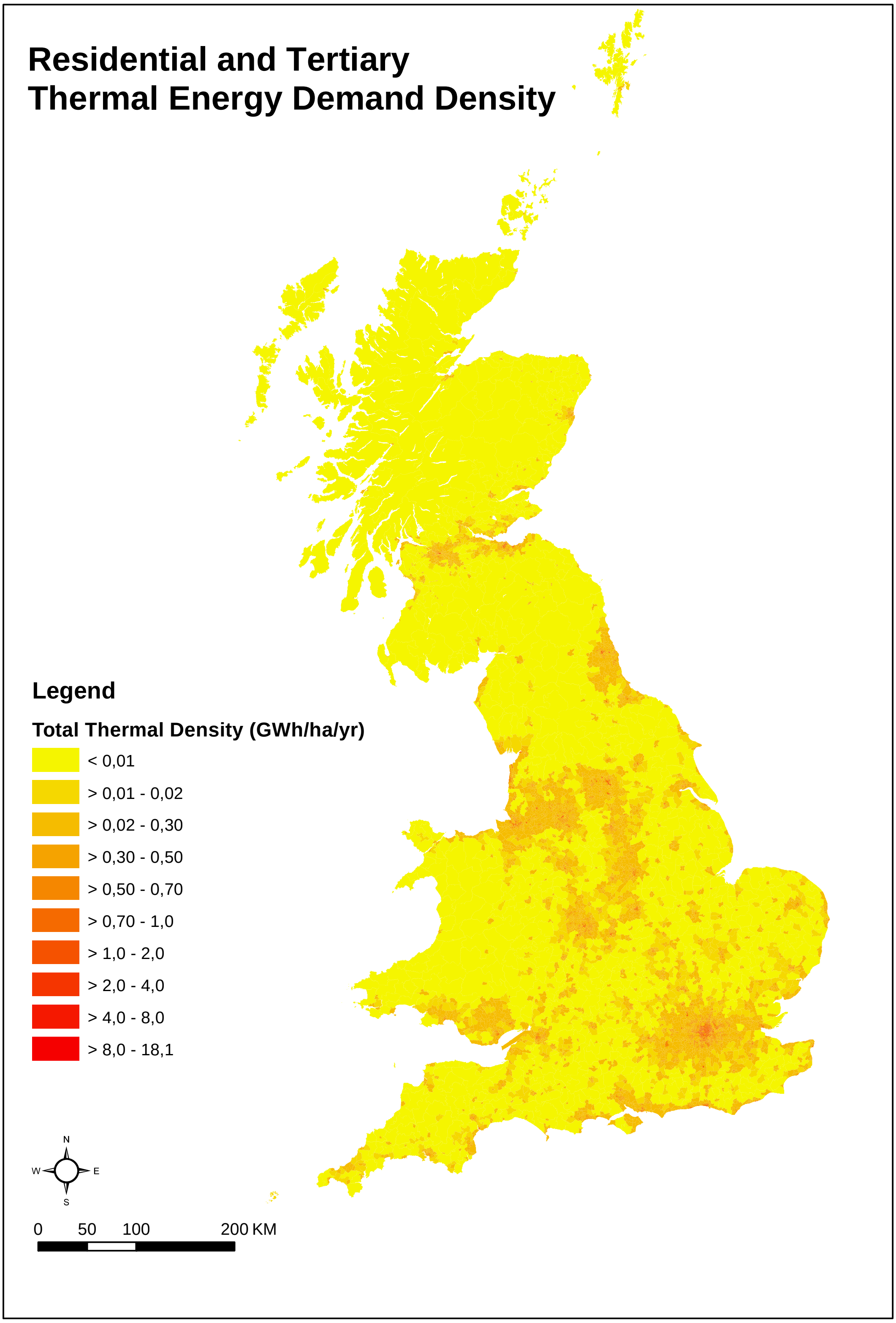


Figure 0-1: Residential and tertiary thermal energy demand density in GB

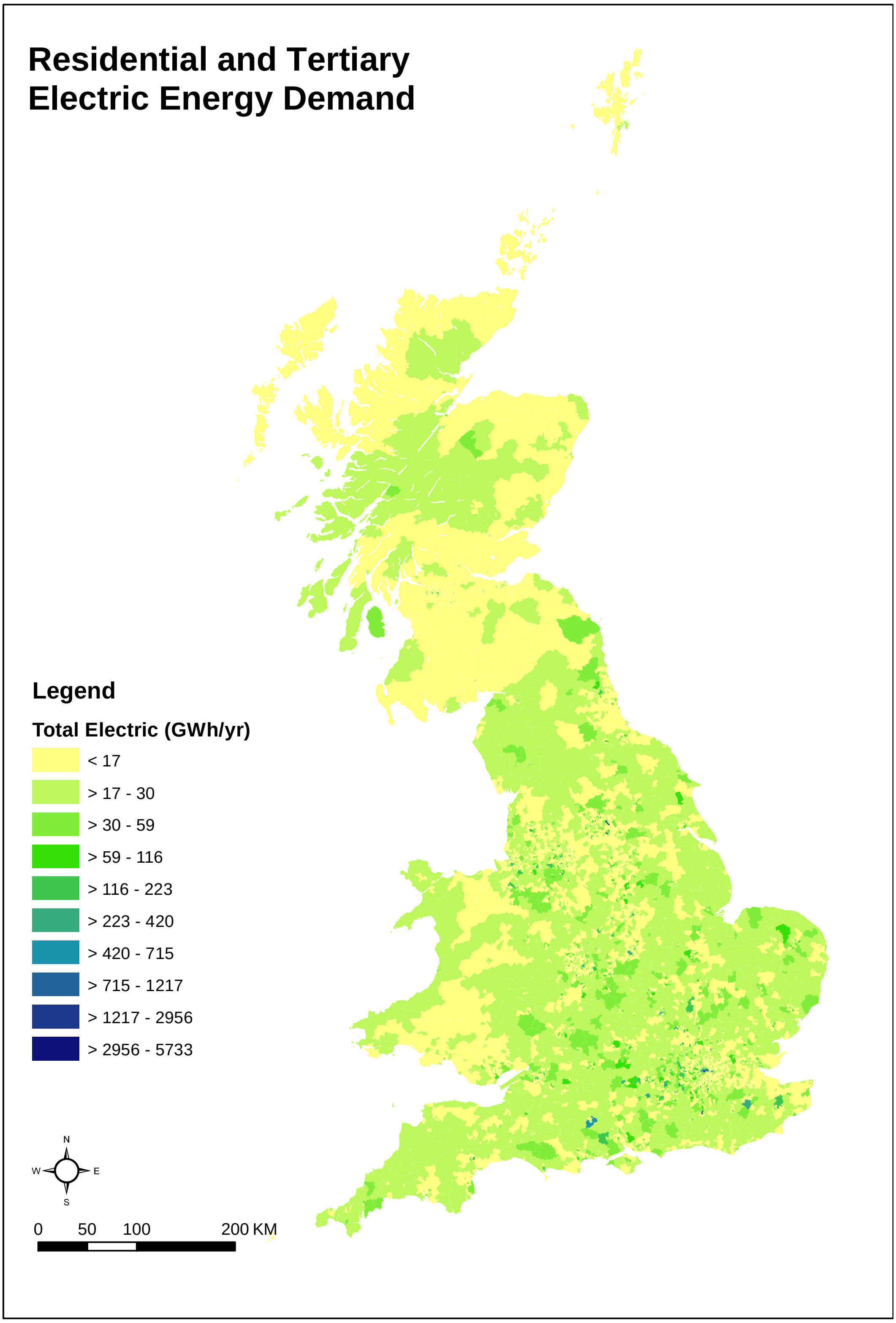


Figure 0-2: Residential and tertiary electric energy demand in GB

Chapter 2 — Clustering

The clustering tool described in this section aims to identify regions in GB where a district heating network (DHN) supplied by a distributed energy system could be economically and environmentally[2] viable. The economic viability in this context is defined by two factors: The heat density[3] and the total heat demand. The environmental effects will be addressed in the GB benefits case of work package 5.1 together with the economic effects (see (Woods 2011), page 2f). The clustering algorithm is based on middle layer super output area (MLSOA) level. There are 8429 MLSOAs in total in GB. Geographical and social information, as well about energy usage has been gathered about them. The clustering tool uses an algorithm that adds MLSOAs together in order to create as many viable distributed energy zones (DE zones) as possible.

Boundary conditions

The input boundary conditions are used for the clustering algorithm:

a) Maximal zonal heating demand, which is 500 GWh.
b) Target minimum zonal heat demand density (mHDD), here 200,000 kWh/hectare.
c) DE Zones have to be continuous, thus two geographically separated MLSOAs cannot form one DE Zone.

Furthermore, it was agreed that local authorities do not influence the creation of district heating networks, i.e. DE zones are not limited within one local authority or region.

Clustering Algorithm: Growing and Merging

The first phase of the clustering algorithm is shown in Figure 0-3 (b) to (f), which can be described as the growing and merging phase. In the beginning, all MLSOAs with a heat demand density above the mHDD become DE zones (see Figure 0-3 (b)). If two zones are neighbours, they get automatically merged (see Figure 0-3 (c)).

During the second step of growing and merging, neighbouring MLSOAs with a HDD down to half of the mHDD are added to the already existing DE zones until either the zonal HDD is dropping below the mHDD or there are no neighbouring MLSOAs of this medium category left. If during this step two DE zones become direct neighbours, they are once more merged. This second step is repeated until no more MLSOAs are added to the DE zones. The effects can be seen in Figure 0-3 (d) – (f).

While the first step starts only with densely populated areas with high energy densities to ensure economic viability, the second step tries to cover as much of the GB heat demand within the DE zones as possible. The high density MLSOAs above the mHDD are used to cross-subsidize the medium dense MLSOA which are not directly viable for district heating.

---

[2] In terms of CO_2 emissions
[3] Defined as annual heat demand in a given area divided by the total land area

Clustering Algorithm: Cutting

During the second phase of the algorithm the DE zones created get reduced to the maximal heat demand of 500 GWh p.a.. Hence, this phase does not affect the DE Zones with a heat demand below this value. DE zones with a demand up to 2000 GWh get cut into their minimal number of zones, i.e. between 500 and 1000 GWh into two zones, above 1000 and up to 1500 GWh into three and between 1500 and 2000 GWh into four DE zones.

All zones with a larger demand than 2000 GWh are split into two first. The step is repeated until no further zones need to be cut. This phase can be seen in Figure 0-3 (g) – (h).

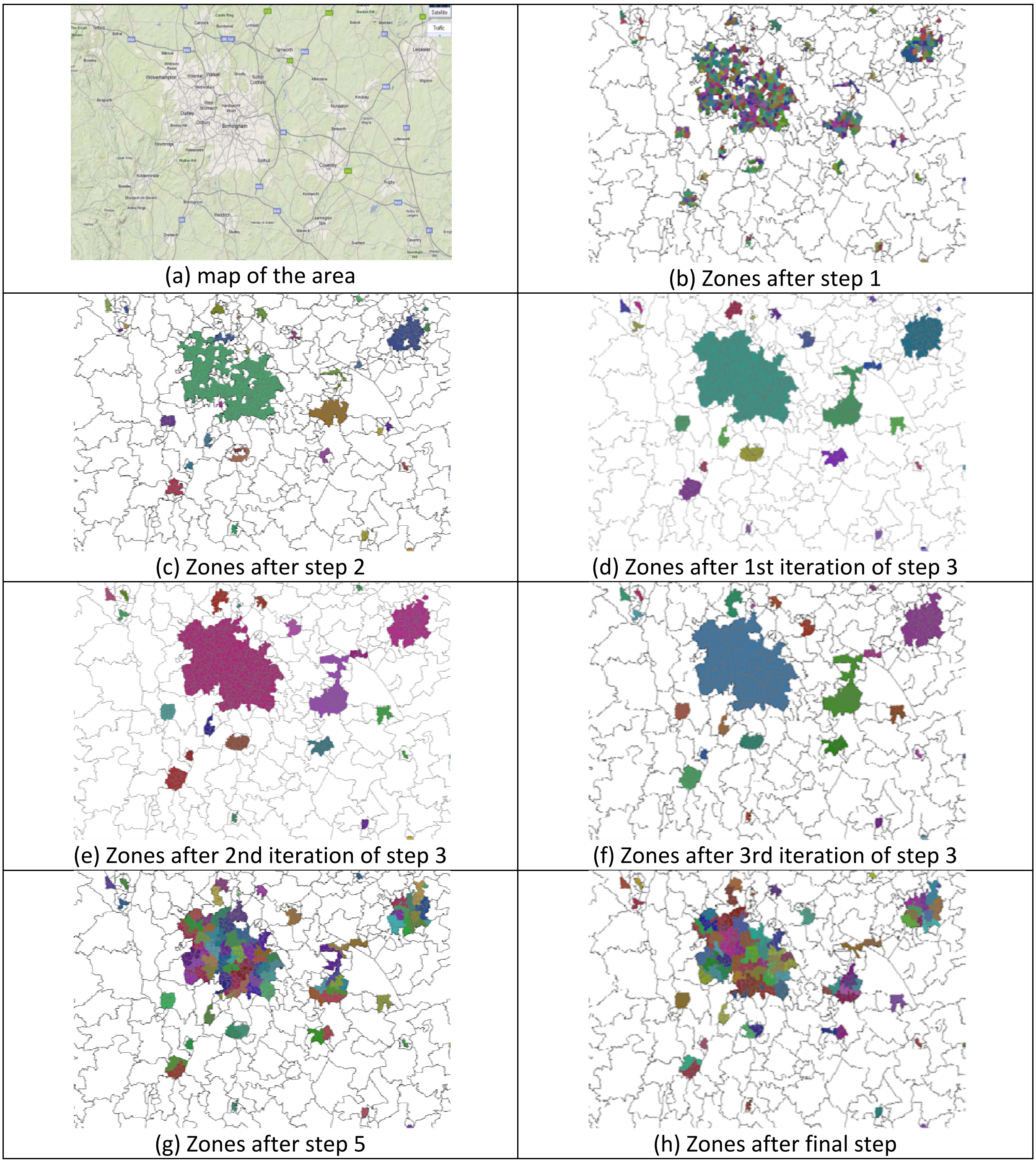


Figure 0-3: Zone evolution of the Birmingham region

In the meantime, it is regularly checked that the DE zones do not drop below the mHDD. Where this happens, single MLSOAs are cut out from the zones until the mHDD is reached again. The effect can be seen in Figure 0-3 (h).

Results and Conclusion

The Clustering algorithm created 946 DE zones in GB, out of which 937 fulfil all three input boundary conditions. 11 of them are still greater than the maximum heat demand of 500 GWh. These are single MLSOA zones that cannot be split further and are therefore kept as DE Zones of their own although their heat demands are too high. They are discussed later in the Appendix J — Outliers.

In total, the DE zones cover more than half of the MLSOAs in GB and nearly half of the demand. The actual figures are shown in Table 0-2. It is remarkable that the DE zones cover only about 4 % of the GB total land area. It needs to be kept in mind that MLSOAs include both rural areas and cities and not all of the MLSOAs' areas are built with homes and tertiary sites. It can be concluded that district heating on this macro scale level is economically viable in cities and larger locations. To give recommendations for smaller settlements and towns the smallest data points need to be smaller than MLSOAs.

Facts about DE zones:

- 946 zones created
- Including 55,3 % of MLSOAs (4660 out of 8429)
- Covering 56.5 % of total calculated residential and tertiary heat demand of GB (258,963 GWh of 458,392 GWh)
- Covering approximately 4 % of the GB area (9,079 km² of 229,012 km²)
- 11 zones marked as outliers (total heat demand greater than 500 GWh per year)

Table 0-2: Summary of created DE zones during the clustering algorithm

Section 3 — Temporal Demand Calculation

To estimate the size of the heat supply centre and to optimise its run-time schedule the temporal energy demands (heat and electric) aggregated for DE zones (see Chapter 2 — Clustering) have been calculated. These load curves represent the total annual demand split up into 8760 hourly values. Finally, the load curves get simplified into 40 characteristic values which comprise the peak and 39 values to describe a typical week in three seasons (i.e. 13 values for each week in each season). The process of determining these time bands contains five sequential steps:

Step 1) Calculation of annual energy demand per house type and per CIBSE TM46 tertiary building type for each DE zone

Step 2) Preparation of annual profiles (heat and electricity)

Step 3) Mapping of demand types to profile types

Step 4) Calculation of zonal temporal energy demand curves (heat and electrical)

Step 5) Calculation of time bands and characteristic hourly energy demand values (seasonal and total peaks and bases)

The first step is analogue to the energy demand calculation (see Chapter 1 — Energy Demand Calculation) except that the tertiary demand is kept in the 29 separate values of the CIBSE TM46 classes.

Preparation and Mapping of Profiles[4]

Heat Profiles

The heat profiles have been prepared according to the method developed by the TU Munich for the BDEW (see ((BDEW 2006). It distinguishes four residential (single family house old and new, multi-dwelling house old and new) and 15 tertiary profile types. For each of these profiles a relative ratio of the annual demand for every day is calculated based on the average outside temperature of that day and the specific slope of a sigmoid function. Then an hourly distribution of the daily values is added which is unique for each category and based on the outside temperature of that day.

Electrical Profiles

The methods to prepare the electrical profiles differ between the tertiary and the residential sectors. The two residential profiles, for the two tariff types ordinary and economy 7, have been directly provided by EDF Energy and consist of 8760 hourly values, which have been normalised, so that their sum is equal to one.

The tertiary electric profiles were based on a method of the VDEW[5] (see (BDEW 1999) and (BDEW 2000)). There are 11 different tertiary and agricultural profiles, which in general differ in the operational hours (day, evening, night or 24 hours) and special usage like cooling or heating.

All of them contain three seasons (winter, summer, transition) and two day types (working day and non-working or weekend day). These days are fitted together in corresponding sequences of weekdays and weekend days of the three seasons to generate one year with 8760 hourly values each. In the last step, these sequences were normalised to result in an integral of 1.

Mapping of the Residential Sector

The next step is to bring the annual demand sectors and the prepared profiles together. Since the categories do not fit together completely, a judgement had to be made to match the existing demand categories to the existing profiles.

[4] In this report Profiles are referred to as normalised temporal demand curves whose integral result into 1. The curves of profiles are showing the shapes of the demand, but need to be multiplied by the annual demand to become complete temporal demand curves.

[5] The VDEW has been merged into the BDEW in 2007 together with three other organisations.

Calculation of Zonal Load Curves

Thermal load curves

To calculate the thermal load curves of a zone, the residential and tertiary parts have to be calculated separately and afterwards summed together. The 35 house types were mapped into the 4 different profile sets and each hourly relative demand value multiplied with the annual demand figure to become the total residential demand of this particular hour. The thermal tertiary load curves have been prepared in the same way with 15 tertiary heat profiles into which the 29 TM46 sectors are mapped. In the end, all values of one hour are summed up to become the zonal heat demand of that particular hour. The integral of all hourly values of one year corresponds to the zonal annual heat demand.

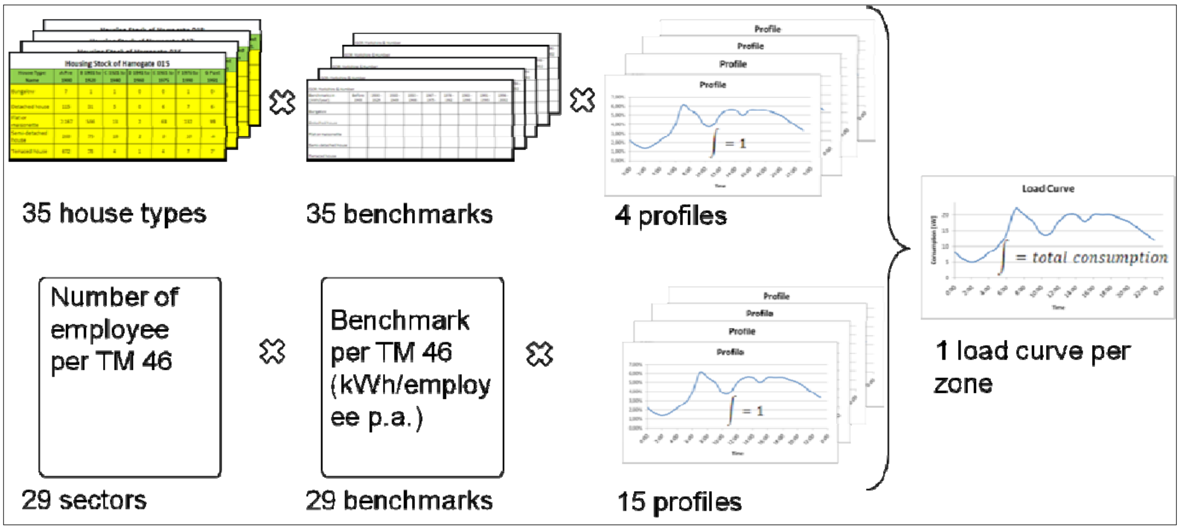


Figure 0-4: Graphical explanation of the heat load curve calculation process for residential (upper) and tertiary (below)

Electrical load curves

On the tertiary side the eleven tertiary and agricultural profiles are matched to the 29 TM46 sectors in the same way as the heat profiles and afterwards the load curves of the sectors have been calculated. For the residential load curves the two profiles from EDF Energy have been multiplied with the domestic electricity demands for both tariffs given by DECC.

The electrical load curves per zone were calculated by the summation of all tertiary electrical load curves and the domestic ones. The integral of all hourly values of one year corresponds to the zonal annual electrical demand.

Time Bands

The two load curves for heat and electrical demand are then simplified into the so called time bands. These are 40 characteristic values consisting of three seasons with two day types of either 6 or 7 time bands and the demand of the annual peak hour. The working days are created out of seven time bands and the weekend days out of six.

The selected hours of each time band are shown in Table 0-3.

Time Bands	Weekday	Time Bands	Weekend
WD 1	0:00 - 04:59	WE 1	0:00 - 04:59
WD 2	5:00 - 6:59	WE 2	5:00 - 6:59
WD 3	7:00 - 9:59	WE 3	7:00 - 9:59
WD 4	10:00 - 16:59	WE 4	10:00 - 16:59
WD 5	17:00 - 18:59	WE 5	17:00 - 21:59
WD 6	19:00 - 21:59		
WD 7	22:00 - 23:59	WE 6	22:00 - 23:59

Table 0-3: Defined time bands for a weekday and a weekend day for a given season

Chapter 4 — Classification

The classification aims to group DE zones with similar attributes together to find commonalities in different parts of GB and reduce the number of DE zones to be analysed in the benefits case of WP 5.1. Up to 25 classes would be created from about 1000 DE zones, thus on average 40 zones will be grouped together to form one class.

Afterwards, out of each class, one zone will be extracted which will be used as representative for its class. These zones are selected to be as close to the centre of their classes as possible and will be called Characteristic Zones.

Cluster analysis

The typical cluster analysis as it is used for the classification process in this case covers the following steps:

1. Selection of objects and variables (parameters)
2. Standardization of variables' values
3. Selection of distance measure
4. Selection of classification method
5. Determination of number of classes
6. Evaluation of classification results
7. Description (interpretation) and validation

Cluster analysis is generally the process to assign a set of objects into groups. These groups are supposed to be different from each other, but the members of each group to be similar. It is often used for data mining and knowledge discovering within large databases.

Methods of cluster analysis

The following two methods are used during the classification process:

The k-means algorithm creates out of a data set k groups with the smallest distance to their class centres. The k centres are randomly chosen in the beginning. When all objects of the data set are assigned to the closest centre, the class' centroids are re-calculated. These two steps are repeated until no more reassignments are registered and the class centres are fixed.

Ward's method is strongly biased towards producing classes with roughly the same number of observations. At each class generation, the within-class sum of squares is minimized over all partitions obtainable by merging two classes from the previous generation.

During the classification process, the ward's method is used to determine the number of classes that should be created at each step of the classification. The k-means algorithm is used afterwards to create the best number of classes.

Both methods are broadly known and accepted in the scientific community and have been included in the software used for classification (see section 4.3 Classification with SAS)

Decision on classification parameters

Before the classification can be executed, the parameters to be used have to be defined. Therefore groups of different parameters have been created by their correlations to each other. These groups can be seen in Figure 0-5.

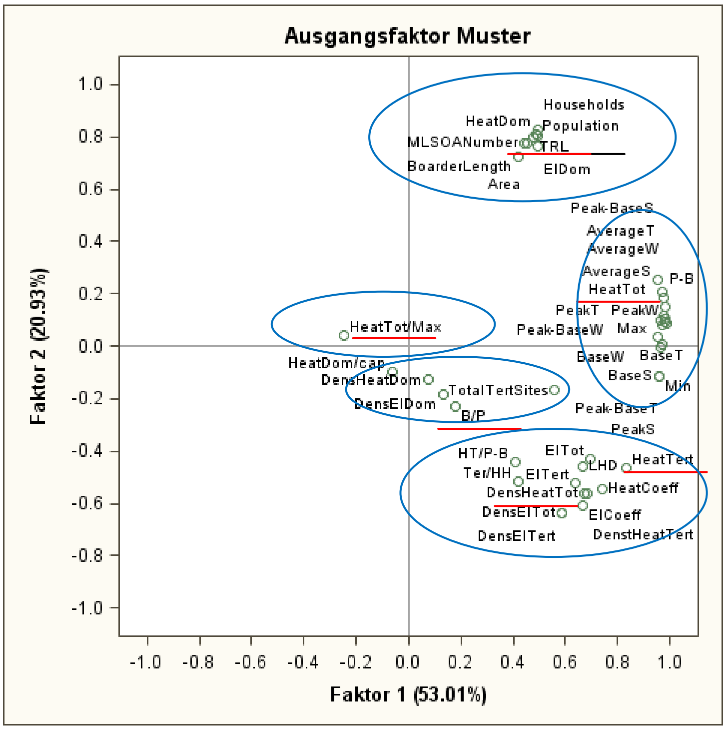


Figure 0-5: Plot of the Factor Analysis with two synthetic parameters. The other paramters are plotted according to their correlations values to these two synthetic ones. The blue circles indicate the five identified groups of parameters out of each one represantative for the further classification has been chosen (indicated by red lines).

Next, from each group of parameters the most important representative parameter was selected. Only less correlated parameters are appropriate for classification because they have the ability to discriminate within the set of objects.

After analysis and discussions within the consortium, the following five parameters were chosen for further classification, which for each group have the greatest impact on the benefits of a DE zone:

- heat demand in the tertiary sector: HeatTert (in kWh)
- total heat demand (the sum of domestic and tertiary heat demand): HeatTot (in kWh)
- The density of total heat demand (total heat demand per hectare) – DensHeatTot (in kWh per hectare)
- The ratio of annual base and peak heat demand – B/P (dimensionless or kWh/kWh)
- The Load Factor, calculated as the total heat demand divided by the product of 8760 and the annual peak heat demand - HeatTot/Max (dimensionless or kWh/kWh)

Out of this group of five parameters, the two parameters total heat demand and density of total heat demand have been decided to be most important for the further analysis of the DE zones.

Intermediate Results

The classification itself has been executed in several sequential steps. This has been necessary to avoid classes with less then five members. These steps and the number of DE zones per class are shown in Figure 0-6.

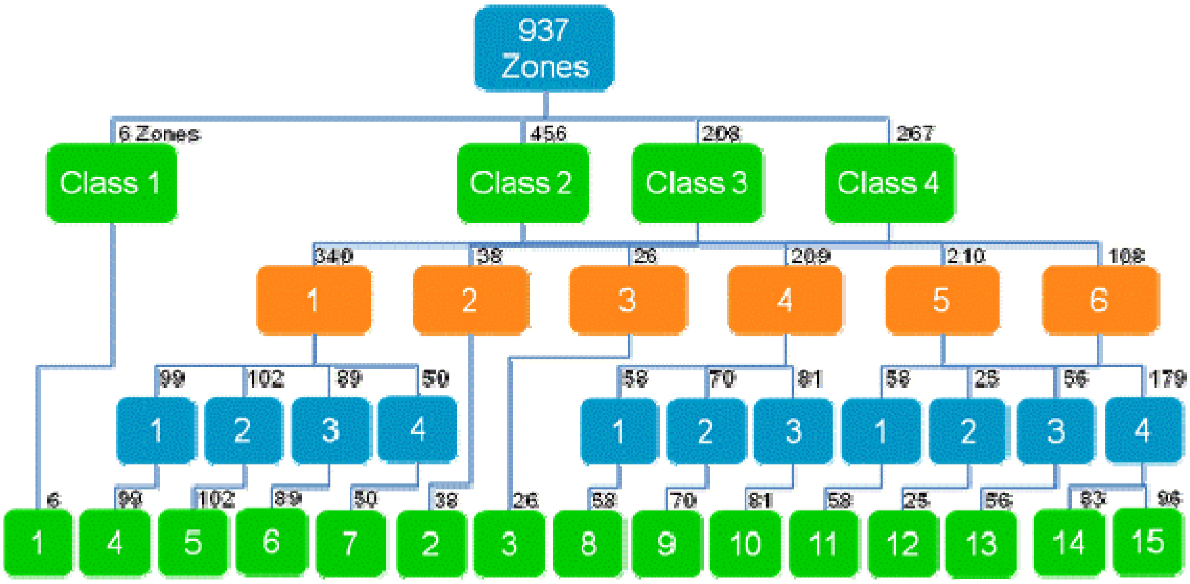


Figure 0-6: Classification steps towards the 15 classes of the intermediate results. The white numbers in the boxes represent the class number and the black figures above the number of DE zones in the classes.

The intermediate results of the classification process with their variation coefficients[6] for the classification parameters are shown in Table 0-4. These results have been analysed and used to start the discussion about homogeneity of classes. It has been decided that some of the classes needed further improvements which was done in the following classification steps.

Class	# zones in class	Variation Coefficients (%)					
		Heat Tertiary	Total Heat	Heat Density	Base / Peak	Load Factor	LHD
1	6	22.51	19.55	24.00	8.83	4.18	28.26
2	38	67.74	23.20	57.74	19.84	4.92	48.21
3	26	30.56	21.33	44.97	13.72	7.26	34.97
4	99	42.57	5.34	22.84	14.42	4.30	13.90
5	102	48.96	7.46	35.51	18.78	5.55	21.46
6	89	50.26	5.25	27.95	18.09	5.47	18.66
7	50	33.62	4.04	30.16	15.57	4.96	19.99
8	58	37.09	12.65	57.48	12.45	4.36	43.01
9	70	37.29	6.39	40.04	13.25	3.49	26.68
10	81	36.43	7.75	47.34	12.56	3.50	26.96
11	58	84.21	15.09	31.79	29.07	6.79	25.32
12	25	38.07	11.22	43.92	21.16	5.74	50.09
13	56	52.92	9.86	25.52	16.44	4.44	24.91
14	83	80.06	28.56	34.83	28.26	7.21	31.83
15	96	59.32	15.12	30.60	24.78	6.22	21.78

Table 0-4: Intermediate results of the classification process showing the number of DE zones per class and the variation coefficients for the five classification parameters.

[6] The variation coefficient defined as the standard deviation of a data set divided by the average of the same set. It can be interpreted as the relative standard deviation, which makes two different data sets comparable to each other.

Improvement of intermediate results

The intermediate results of classification, covering the 15 classes, did not satisfy all expectations. The variation coefficients for the two most important parameters total heat and heat density were in several cases much higher than 20%, which resulted in unsatisfactory homogeneity of zones within classes (see Table 0-5). Thus, it has been decided to re-iterate the classification process for some of the classes which showed potential for improvements. For the following steps the number of parameters has been reduced two total heat and heat density. The effects on the other three parameters are of lesser importance and thus the rules for their variation coefficients have been loosened. For consistency the variation coefficients are still shown for all five parameters in the following Table 0-5 – Table 0-8.

	Class	# Zones	Heat Total	Density Heat Total	Load Factor	Linear Heat Density	
red	1	6	19.55%	24.00%	4.18%	28,26%	red
red	2	38	23.20%	57.74%	4.92%	48,21%	red
red	3	26	21.33%	44.97%	7.26%	34,97%	red
green	4	99	5.34%	22.84%	4.30%	13,90%	green
yellow	5	102	7.46%	35.51%	5.55%	21,46%	yellow
green	6	89	5.25%	27.95%	5.47%	18,66%	green
yellow	7	50	4.04%	30.16%	4.96%	19,99%	yellow
yellow	8	58	12.65%	57.48%	4.36%	43,01%	yellow
red	9	70	6.39%	40.04%	3.49%	26,68%	red
red	10	81	7.75%	47.34%	3.50%	26,96%	red
green	11	58	15.09%	31.79%	6.79%	25,32%	green
green	12	25	11.22%	43.92%	5.74%	50,09%	green
yellow	13	56	9.86%	25.52%	4.44%	24,91%	yellow
yellow	14	83	28.56%	34.83%	7.21%	31,83%	yellow
green	15	96	15.12%	30.60%	6.22%	21,78%	green

Table 0-5: Intermediate results of 15 classes are represented by sample size and different figures of variation coefficients. The colour code at the left and right is an indicator for the class homogeneity (green = good, yellow = medium, red = bad).

Phase 1 – Re-classification

The re-classification of the intermediate results has been carried out on six of the fifteen classes. These classes were identified by the homogeneity analysis. To this set belong the classes 2, 3, 8, 9, 10 and 12 (see Table 0-5). The remaining nine classes have been left unchanged.

The six above mentioned classes have been re-classified using the same cluster analysis procedures, but the number of classification factors has been reduced to the total heat demand and the heat density. As a first result, three new classes have been obtained. In the second step of phase 1, the first two groups have been split further to improve the variation coefficients. In total, 7 groups have been created.

The classes obtained in this stage were put together with the remaining nine classes of the intermediate results, resulting in 16 new classes (see Table 0-6).

Class	# Zones per Class	Variation Coefficient (%)					
		Heat Tertiary	Heat Total	Heat Density	Base / Peak	Load Factor	LHD
1	6	22,51	19,55	**24,00**	8,83	4,18	28,26
2	68	44,95	6,45	**26,45**	24,24	7,28	30,37
3	56	67,35	11,57	**25,08**	26,39	5,94	24,09
4	99	42,57	5,34	**22,84**	14,42	4,30	13,90
5	102	48,96	7,46	**35,51**	18,78	5,55	21,46
6	89	50,26	5,25	**27,95**	18,09	5,47	18,66
7	50	33,62	4,04	**30,16**	15,57	4,96	19,99
8	12	74,53	18,42	16,79	12,26	7,79	31,73
9	14	72,58	10,24	19,09	14,84	4,71	21,67
10	64	43,34	5,26	21,74	21,20	6,76	21,06
11	58	84,21	15,09	**31,79**	29,07	6,79	25,32
12	57	38,04	4,98	19,12	16,51	4,54	20,37
13	56	52,92	9,86	**25,52**	16,44	4,44	24,91
14	83	80,06	**28,56**	**34,83**	28,26	7,21	31,83
15	96	59,32	15,12	**30,60**	24,78	6,22	21,78
16	27	56,48	**23,84**	**24,63**	22,34	10,86	24,59

Table 0-6: Intermediate results of all 16 classes after phase 1 of the re-classification.
Variation coefficients on the two parameters total heat demand and heat density are marked red if above 20%.

Phase 2 – Discriminant analysis[7]

In order to improve the results the discriminant analysis has been executed on the two parameters total heat demand and total heat density.

Discriminant analysis is a statistical method, which allocates elements to already existing classes using created classifier (discriminant function). Based on its value the probability of affinity to the particular class is tested. In practice, the distribution of analysed class attributes is analysed and a probability density function is used as the discriminant function. As a result, possible new localisations for all elements are found. If the value of the classifier of a given element is higher in a new class than in the original one, then the element is moved to the new class. The discriminant analysis can be performed with more than one parameter.

The improved results are shown in Table 0-7.

[7] Discriminant analysis is a general method to separate a data set (objects) into a defined number of classes based on a combination of variables.

Class	# zones in class	Variation Coefficients (%)				
		Heat Density	Total Heat	Load Factor	Base / Peak	Heat Tertiary
1	6	24,00	19,55	4,18	8,83	22,51
2	79	28,89	3,34	8,99	36,09	70,82
3	50	22,93	5,28	6,18	25,61	60,43
4	81	21,23	2,93	7,34	30,68	62,36
5	93	23,33	3,94	7,08	30,62	73,61
6	78	22,20	3,39	6,40	27,60	60,53
7	51	21,37	2,08	6,04	24,09	60,43
8	26	26,20	16,17	9,28	26,18	82,02
9	19	28,56	6,96	4,90	20,38	81,58
10	81	24,75	2,42	6,79	28,15	64,95
11	56	32,28	14,40	6,83	29,31	82,76
12	55	18,89	3,15	6,25	24,70	47,26
13	64	28,58	9,86	6,48	29,34	72,70
14	78	36,21	28,17	7,12	28,06	79,74
15	103	28,72	16,91	6,28	24,69	60,64
16	17	22,44	17,50	11,09	22,79	43,62

Table 0-7: Results of the 16 classes after the discriminant analysis on the parameters total heat and heat density. Values above 20% on the two parameters total heat and heat density are indicated in red.

Final Iteration

Due to the still unsatisfactorily high values of variation coefficient for some of the classes, an additional classification step was executed for classes with first either high number of classes or large share of the calculated GB demand within all zones and secondly a high variation coefficient for either heat density or total heat demand.

Therefore, it was decided that classes 1, 3, 7, 9, 12, 13 and 16 show good results or cannot be improved further and thus have been left without changing. Classes 2, 5, 10 and 14 needed updating because of their high variation coefficients of total heat (class 14) or of heat density (classes 2, 5 and 10), while classes 4, 6, 8, 11 and 15 were considered to be updated, but finally were left unchanged, because no improvements could be achieved.

The step has been performed according to the schema presented in Table 4-22. Class 2 has been split at a density above and below 310,000 kWh/ha. The splitting point in class 5 could be found at 300,000 kWh/ha, while 280,000 kWh/ha has been used for class 10. Class 14 has been split on the parameter total heat. Here the value of 32 GWh has been used.

Results

Final Classes

Finally, 20 classes have been created during the classification process, which have been presented to the consortium. Most of the classes show satisfying results with variation coefficients of 20 % or less, at least for the two main classification parameters total heat and heat density. Further re-classifications have been tested, but they did not improve the results significantly. Thus, Table 0-8 shows the final results of the classification process.

		Variation Coefficients (%)				
Class	# Zones	Total Heat	Heat Density	Tertiary Heat	Base/Peak	Load Factor
1	6	19,55	24,00	22,51	8,83	4,18
2	17	3,82	12,83	64,28	40,82	11,67
3	50	5,28	22,93	60,43	25,61	6,18
4	81	2,93	21,23	62,36	30,68	7,34
5	36	4,15	11,57	73,41	35,26	8,77
6	78	3,39	22,20	60,53	27,60	6,40
7	51	2,08	21,37	60,43	24,09	6,04
8	26	16,17	26,20	82,02	26,18	9,28
9	19	6,96	28,56	81,58	20,38	4,90
10	26	2,09	17,38	70,39	33,35	7,82
11	56	14,40	32,28	82,76	29,31	6,83
12	55	3,15	18,89	47,26	24,70	6,25
13	64	9,86	28,58	72,70	29,34	6,48
14	30	8,76	22,24	60,66	26,33	7,16
15	103	16,91	28,72	60,64	24,69	6,28
16	17	17,50	22,44	43,62	22,79	11,09
17	62	2,58	12,88	71,48	32,56	8,17
18	57	3,82	11,93	68,94	23,88	5,85
19	55	2,58	10,69	61,71	25,59	6,27
20	45	20,58	15,30	94,71	23,65	3,59

Table 0-8: Results of all 20 classes, showing the number of zones in each class and the classes' variation coefficients for the 5 classification parameters

Characteristic Zones

Out of each class, the DE zone closest to the class centroid has been determined to become the Characteristic Zone. These Characteristic Zones are sent together with detailed information about the area and their time bands to the University of Manchester for the homogeneity analysis (see Chapter 5 —). Afterwards these zones will be used as representatives of their classes to analyse the national case and the benefits case. Table 0-9 shows the Characteristic Zones and their detailed information.

	Characteristic Zones (Total Heat & Heat Density)									
Class Nr.	1	2	3	4	5	6	7	8	9	10
Zone Nr.	154	711	239	619	70	509	127	145	29	138
Heat Total (kWh)	3,49E+08	3,10E+08	2,30E+08	4,00E+08	2,70E+08	3,28E+08	4,80E+08	2,81E+08	4,52E+08	3,64E+08
Electric Total (kWh)	2,42E+08	8,97E+07	7,41E+07	2,37E+08	1,93E+08	1,02E+08	1,36E+09	1,12E+08	4,37E+08	1,36E+08
Heat Density (kWh/ha)	4,31E+06	3,18E+05	2,23E+05	2,29E+05	3,28E+05	2,22E+05	2,29E+05	6,01E+05	8,47E+05	3,46E+05
Elec. Density (kWh/ha)	2,99E+06	9,20E+04	7,20E+04	1,36E+05	2,35E+05	6,90E+04	6,46E+05	2,39E+05	8,19E+05	1,29E+05
Zone Population	15089	42350	28259	36680	37325	45254	64948	41887	47184	67152
Zone Area (hectare)	81	976	1027	1749	821	1478	2097	467	533	1051
# of Households	8361	18551	11605	14890	15296	18787	27243	18779	21924	28590
# of Tertiary	3206	892	2052	3848	8388	1842	5657	2999	15295	2981
# of MLSOAs	2	6	4	5	6	6	9	6	7	9
Thermal Peak (kWh)	1,42E+05	1,08E+05	7,71E+04	1,40E+05	1,04E+05	1,13E+05	1,90E+05	1,07E+05	1,60E+05	1,34E+05
Thermal Base (kWh)	6,20E+03	2,60E+03	2,70E+03	4,90E+03	2,80E+03	3,10E+03	3,90E+03	3,10E+03	8,30E+03	4,20E+03
Class Nr.	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Zone Nr.	566	313	686	535	734	148	23	547	701	909
Heat Total (kWh)	1,13E+08	4,44E+08	1,81E+08	3,76E+07	5,67E+07	3,74E+08	2,97E+08	2,68E+08	3,64E+08	2,33E+07
Electric Total (kWh)	2,78E+07	1,83E+08	5,61E+07	1,16E+07	2,04E+07	5,02E+08	9,94E+07	9,01E+07	1,23E+08	7,30E+06
Heat Density (kWh/ha)	2,48E+05	2,23E+05	2,24E+05	2,52E+05	2,22E+05	1,46E+06	2,79E+05	2,78E+05	2,07E+05	2,14E+05
Elec. Density (kWh/ha)	6,10E+04	9,20E+04	7,00E+04	7,80E+04	8,00E+04	1,96E+06	9,40E+04	9,30E+04	7,00E+04	6,70E+04
Zone Population	15128	49510	21070	5510	7724	22075	36020	23745	41394	3222
Zone Area (hectare)	455	1993	806	149	255	256	1063	964	1757	108
# of Households	6461	19912	8768	2312	3150	10586	15063	10543	18045	1576
# of Tertiary	476	3807	323	158	563	12227	3551	2154	1968	70
# of MLSOAs	2	7	3	1	1	3	5	3	5	1
Thermal Peak (kWh)	3,82E+04	1,53E+05	5,98E+04	1,36E+04	1,98E+04	1,45E+05	1,19E+05	8,86E+04	1,21E+05	7,30E+03
Thermal Base (kWh)	1,10E+03	4,90E+03	2,10E+03	4,00E+02	5,00E+02	7,00E+03	2,30E+03	3,40E+03	4,30E+03	3,00E+02

Table 0-9: Values of Characteristic Zones

Additional Zones (quartiles)

In addition to the Characteristic Zones, a set of 4 further zones from each class has been selected. The role of these zones is to verify the homogeneity within their classes. The lesser the difference between these points, the more homogeneous and compact are the classes. This means that Characteristic Zones represent their own classes in a good way.

In order to define the additional points from classes, the quartiles for the two parameters total heat and heat density for all 20 classes have been determined (see Appendix F — List of Quartile Zones).

Chapter 5 — Application of DE design tool to evaluate characterisation of zones

The ETI Macro DE design tool is applied to the characteristic zones defined in Work Package 2 of the Macro DE project. The design tool generates the design for an optimised energy centre and estimates costs and performance of DE solutions (the energy centre and district heating network, (DHN)) to meet the heat requirements of each characteristic zone. The purpose of the characteristic zones is to facilitate understanding of the potential of DE for Great Britain by evaluating only a limited number of characteristic zones, each representing a class of demand zones.

The purpose of this analysis is to assess whether i) the characteristic zones generated in Work Package 2 are distinctively different from each other, and ii) the classes are sufficiently homogeneous. The report summarises the DE design methodology and assumptions. The results of applying the DE design tool are presented for 20 characteristic zones (intended to represent 20 classes of zones) and a further four zones in each class, representing a range (the two middle quartiles) of the thermal demand and thermal demand density in the class.

The performance results indicate that the characteristics of the classes (including annualised capital cost and total annualised cost of the DE centre, the district heating network and the overall DE solution, and CO_2 emissions from fuel consumption by the DE centre) vary across a wide range, supporting the proposal that the classes are indeed distinctive.

For the five zones examined within each class, performance results are presented; these include: annualised capital cost and total annualised cost, CO_2 emissions from fuel combustion and avoided CO_2 emissions associated with electricity generated in the energy centre. To assess the homogeneity of the classes, the maximum difference between the performance of the characteristic zone and that of the other zones in the class is quantified on a ‘per unit of heat delivered’ basis.

It is observed that in seventeen of the twenty classes, all the quartile zones were relatively similar to the characteristic zone, *i.e.* that their total annualised costs (per unit of heat delivered) were within 20% of that of the characteristic zone. In the remaining three classes, the total annualised costs of the quartile zones were within 30% of that of the characteristic zone; the simplified cost model for the district heating network is shown to be significantly responsible for the differences in these cases. The performance results in terms of CO_2 emissions are very similar for all classes. The results confirm that the classes generated in Work Package 2 and the associated characteristic zones are appropriate for assessing the potential for macro DE for GB.

Chapter 6 — Methodology to assess industrial waste heat and its potential for district heating

Introduction

The broad objective of the MacroDE project is to assess the opportunity for providing low carbon distributed energy (DE) solutions to aggregated energy demand zones across the UK. One way to increase DE potential is to make use of heat sources, for example coming from industrial processes that are currently being lost. Consequently, the work described in this section is aimed at assessing the industrial waste heat[8] resources and their locations across GB.

The report includes the description of the methodology developed to select industrial sectors and evaluate waste heat potential as well as the confidence level in the results. The deliverable also includes an excel table providing results of waste heat potential per industrial sector and per site location (postcode), in terms of quality (temperature) and quantity (energy). A map indicating location and size of heat source is also provided.

A methodology in 4 steps

The approach adopted to calculate the industrial waste heat (or industrial excess heat which will be used synonymously) is composed of 4 steps. The procedure used in step 1 is largely described in (McKenna 2009), but step 2, 3 and 4 will be described in this section of this report.

- **Step 1: Determination of energy consumption per industrial site**
  Since, there are no public data of energy consumption per industrial site, a procedure based on the use of site-specific data contained in the EU-ETS National Allocation Plan (DEFRA 2010), which provide annual CO2 emissions per committed industrial site, was adopted.

- **Step 2: Determination of waste heat ratios**
  The work determines for each industrial sector the ratio of waste heat to input fuel energy, in other words, the fraction of the primary energy consumed by a site that is lost from the site as waste heat. It is based on previous analysis made by University of Bath (McKenna 2009), complemented by EDF-R&D feedback from audits. When several different industrial processes exist within a given sector (i.e. blast furnace or basic oxygen furnace for the iron and steel sector) several ratios can be provided. The ratios are expressed with "lower" and "upper" bands, to take into account the large discrepancy between industrial sites.

- **Step 3: Determination of waste heat potential per industrial site**
  Application of waste heat ratio (determined in step 2) to energy consumption (determined in step 1) provides the waste heat potential per industrial sector (UK-SIC codes).

- **Step 4: Visualisation**
  The locations of each industrial site with a potential for usable waste heat (and their sector reference) have been marked on a UK map. A picture of this map is included in this report and added in a higher resolution to this deliverable.

[8] "waste heat" refers here to "excess of heat", i.e. heat that should be valorised out of the industrial site (on-site use or recovery is out of scope for this study).

Quality of heat

Although the quantity of waste heat available is an important parameter, it is not alone an effective measure of waste heat recovery opportunity. Indeed, it is important to specify the waste heat quality, which means the usefulness of the waste heat, as determined by its temperature.

Waste heat streams are often grouped into 3 temperature categories to distinguish how these streams might be used, with different energy recovery technologies: low (< 100°C), medium (100-500°C), and high (> 500°C) (see examples in Table 6-1).

Consequently, we can consider **100°C as reference temperature for heat recovery**.

Quantity of heat: potential coverage

The quantity of waste heat contained in a waste stream is a function of both the temperature and the mass flow rate of the stream:

$$E = m \times \Delta H(T)$$

Where *E* is the waste heat (kW); *m* is the waste stream mass flow rate (kg/s); and $\Delta H(T)$ is the waste stream specific enthalpy (kJ/kg) as a function of temperature, relative to a reference state.

Enthalpy is not an absolute term, but it must be measured against a reference state. In this report, the enthalpy of waste heat streams is calculated at room temperature and atmospheric pressure ambient.

Figure 6-3Figure 0-7 illustrates that the majority of heat recovery is estimated to be within the temperature band from 100°C to 500°C.

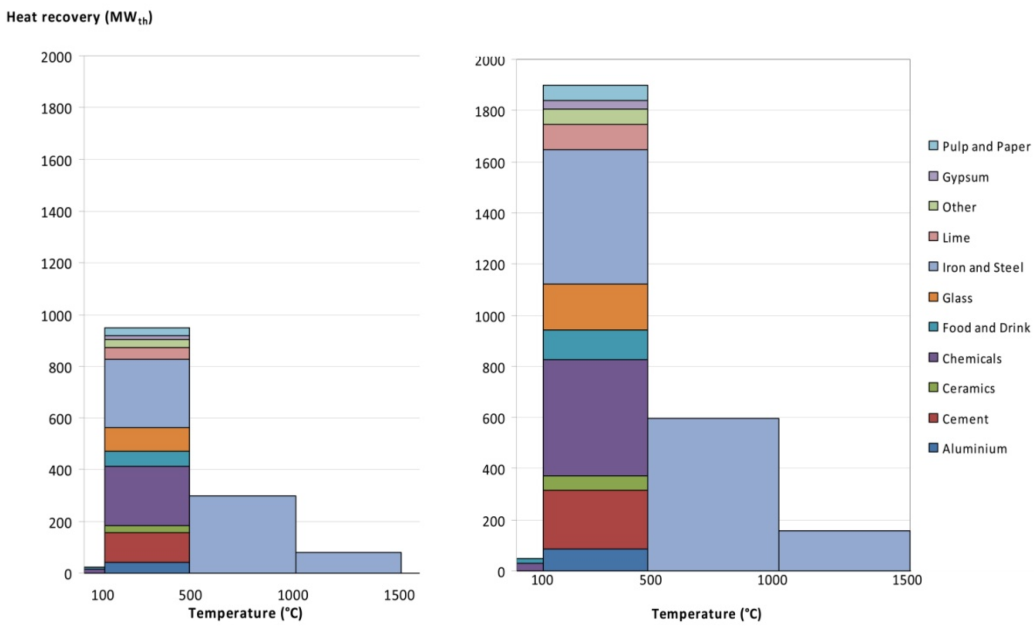


Figure 0-7 : Heat recovery potential by sector against temperature, low scenario (left) and high scenario (right) (see (McKenna 2009)).

Energy consumption per industrial site

In total, more than 260 industrial sites are considered, each site being characterized by:

- Company name
- Postal address (including postal code)
- UK-SIC code (2 digits), which refers to the sector reference
- Annual Energy consumption (electricity and per fuel type : coal, gas, oil, coke) in MWh or PJ (10^{15} J)

The following sectors are included in the report:

- Cement
- Lime
- Glass
- Iron & steel
- Aluminium
- Pulp & Paper
- Chemicals
- Ceramics
- Food & drink (breweries, malteries, sugar cane, sugar beets)

Calculation of waste heat ratios per sector

The same approach was taken for Cement, Glass, Pulp & Paper, Ceramics and Food & drink, while for aluminium, iron and steel, chemicals, and lime specific methodologies were adopted either due to their heterogeneous nature and/or because parts thereof have not been included in the National Allocation Plan (DEFRA 2010). For these last four sectors the method employed involved information about production capacities for individual sites and products, in conjunction with specific energy consumptions for these processes[9].

Summary of data

The waste heat potential of each industrial sector is defined by the exhaust temperature of the waste heat (quality) and the ratio of energy consumption (quantity) that can be made available for further use. All above mentioned industrial sectors have been analysed to obtain these two figures. The detailed description of the analysis can be found in section 6.4. The results are summarised in Table 0-10 below.

[9] Specific energy consumption data was mostly obtained from the relevant sector BREFs (Bat REFerence document – BAT : Best Available Technologies) and EU ETS Benchmarking studies (see DEFRA (2010). EU Emissions Trading Scheme: Approved Phase II, National Allocation Plan 2008 - 2012. London, Department of Environment, Food and Rural Affairs.)

Sector / Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Cement	150 – 300	10	20
Lime	150 – 300	10	20
Glass	500	10	20
Iron & steel Coke oven gas	1000 – 1100	10	25
Iron & steel Coke oven waste gas	200	50	100
Iron & steel Sintering	1350	50	100
Iron & steel Blast furnace	200	15	30
Iron & steel Basic Oxygen Furnace	1600	10	20
Iron & steel Electric Arc Furnace	1200	10	20
Iron & steel Continuous casting	980	25	50
Iron & steel Hot rolling	900	30	60
Aluminium	100	10	20
Pulp and paper	150	10	20
Food and drink Breweries	150	5	10
Food and drink Distilleries	80	5	10
Food and drink Maltings	40	5	10
Food and drink Sugar beet	200	5	10
Food and drink Sugar cane	150	5	10
Ammonia	350	5	10
Ethylene	150	5	10
Chlorine	100 – 150	5	10
Ceramics	150	10	20

Table 0-10: Summary of the waste heat potential of different industrial sectors

Validation of waste heat coverage

When applying the ratio 'waste energy' to 'final energy consumption' to the total energy consumption per sector as shown in Table 0-10, the share of industrial waste heat treated in this study, with regard to total potential industrial waste heat across UK, can be obtained. Results are illustrated in Table 0-11. The study covers about 70% of total waste heat.

	Energy consumption (PJ)		Waste Heat	Waste heat potential (MWth)	
Industry sector	Our study	ECUK	ratio	Our study	ECUK
Aluminium	26	25	0,10	87	83
Cement	44	61	0,20	264	366
Ceramics	19	23	0,10	63	76
Chemicals	203	190	0,10	678	634
Glass	29	55	0,20	195	369
Iron and steel	251	67	0,13	1074	287
Lime	16	14	0,20	96	84
Pulp and paper	18	109	0,10	51	307
Pulp and paper - CHP	60		0,00	0	0
Food and drink	42	143	0,10	133	453
Food and drink - CHP	42		0,00	0	0
Other non minerals	3	450	0,20	20	1268
Others	27		0,08	61	
Gypsum	9		0,10	32	
'otal Energy consumptio	789	1137	Total Waste heat	2754	3930
			Waste heat covered in study (%)		70

Table 0-11 : Comparison DUKES and EU-ETS (our study) on waste heat coverage per industrial sector

Visualisation of waste heat

All potential industrial sites in the UK have geographical co-ordinates and the recoverable waste heat (both low and high scenario) are visualised using Geographic Information System. Afterwards, waste heat arising from the individual sites is aggregated at each MLSOA/IGZ, and thus provides a spatial representation of combined recoverable waste heat. One example of the calculated waste heat potential for GB is shown in Figure 0-8. It shows the solution for the high exhaust fraction combined with the available waste heat and its temperature on the level of MLSOAs.

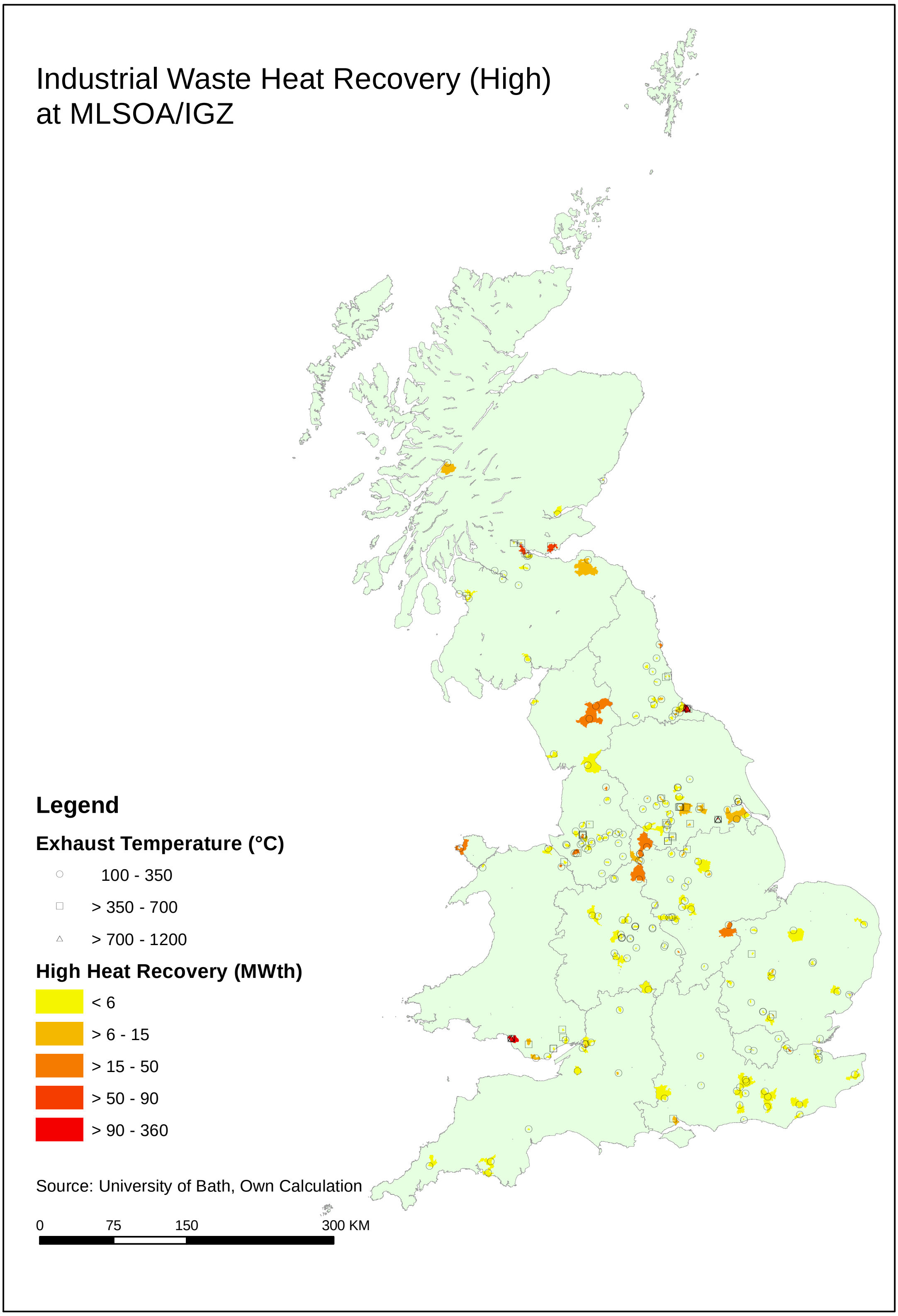


Figure 0-8: Maps presenting high heat recovery scenarios and corresponding waste heat potential across GB

Final Conclusion

Work Package 2 of the ETI Macro DE project investigates the energy demand of all MLSOAs in GB, creates macro-scale DE zones of up to 500 GWh where district heating is viable and clusters them into 20 classes. These classes have been tested for their homogeneity within a class and the heterogeneity between classes. Additionally, the temporal demand for the residential and tertiary sectors is calculated and a study about the potential of waste heat in GB conducted.

The total GB demand has been calculated for the residential and the tertiary sector, both electrical and heat demand. Indicators data, benchmarks and profiles have been collected on different levels, down to postcodes, to ensure the requested granularity. It can be concluded that the knowledge about the tertiary sector is lower than the residential sector. The reasons for this could be the diversity of the tertiary sectors, the lower heat demand, but most likely the fact that historically the tertiary sector has been of less interest. In conclusion, the total calculated GB heat demand of these two sectors is 458,392 GWh per year and the total electrical demand 198,747 GWh per year.

Of this demand nearly 50 % has been assessed as viable for district heating. These 4660 MLSOAs out of 8429 have been clustered into 948 DE zones. Although more than 50 % of the MLSOAs are included within the DE zones, they cover only about 4 % of the area of GB, which leads to the conclusion that macro-scale district heating can only be established in areas of high residential and tertiary energy demand, thus cities or other densely populated areas. A limitation of this study has been the size of MLSOAs, especially in rural areas where a lot green space like forests or fields are included in the MLSOA area. For these MLSOAs a further study, e.g. on the base of LLSOA or even smaller is recommended.

During the classification process the 948 created DE zones have been grouped into 20 classes. Out of each class, one DE zones has been chosen as representative for all DE zones of this class. These 20 Characteristic Zones will be used for further analysis within the ETI Macro DE project. The Characteristic Zones have to be analysed independently of their actual locations, because they are typical for their class in terms of energy demand, but not in terms of area or region.

The 20 classes have been tested on their distinction between each other and the homogeneity within each class. The first task – the distinction – can be shown by the difference of annual heat demands, energy densities or further parameters.

The second part of work package 2 was a study about the availability of waste heat in GB. It was shown that industrial sites can be an additional heat source for DE zones as the heat is already produced and is currently being lost. There is the need for further analysis to assess the temporal availability of the waste heat source and how surplus heat[10] can be damped, but this should be done on a site by site basis.

[10] Here, surplus heat means the exceeding of the heat demand by the waste heat supply.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

AD	Annual Demand
AM	Assignment matrix
CHP	Combined heat and power
CIBSE	Chartered Institution of Building Services Engineers
DB	Database
DECC	Department of Energy and Climate Change
DH	District heating
DHN	District heating network
ERM	Entity Relationship Model
GB	Great Britain
GIS	Geographic Information System
GOR	Government Official Region
GSP	Grid Service Points
HDD	Heat demand density
IGZ	Intermediate Geography Zone
MFH	Multi Family House
mHDD	Minimal heat demand density
MLSOA	Middle Layer Super Output Area
MZHD	Maximal zonal heat demand
NBHM	Neighbourhood matrix
NSPD	National Statistics Postcode Directory
ONS	Office of National Statistics
SFH	Single Family House
SIC	Standard Industrial Classification
TM46	Tertiary energy benchmarks defined by CIBSE (see Table B-1)
WD	Weekday
WE	Weekend Day
WP	Work Package
ZLCV	Zonal Load Curve Value

Introduction

The ETI Macro DE project is a feasibility study about the potential of macro scale district heating systems to reduce CO2 emissions in GB funded by the Energy Technologies Institute. The current report describes the work and results of the work package 2. Its aim in the project was to calculate the energy demand in Great Britain on the level of MLSOAs, assign these areas into district heating (DH) zones using an intelligence where DH is economic viable and group these zones into up to 25 classes out of which one representative zone per class was chosen (WP 2.1). Additionally, a feasibility study about waste heat as a supply for DH was created (WP 2.2).

The first part was split into 5 tasks:

1) Calculation of the energy demand of all 8429 MLSOAs (residential and tertiary, heat and electrical demand (Energy Demand Calculation)
2) Creation of DH zones (Clustering)
3) Calculation of load curves and time bands for all DE Zones (Temporal Demand Calculation)
4) Grouping the DE zones into up to 25 classes (Classification)
5) Verification of results
    a) Distinctiveness between classes
    b) Homogeneity of zones within each class

These tasks form the frame of this report. The results of this work will be used in work package 5 for the benefits case. To streamline further work load, the estimated 1000 DE zones have been grouped into max. 25 classes out of which one DE zone per class has been chosen as Characteristic Zone to represent the whole class in further calculations and conclusion for GB.

Chapter 1 — Energy Demand Calculation

1.1 Description of demand database and GIS data catalogue

This section describes the development process and characteristics of the demand database and the Geographic Information System (GIS) data catalogue which are fundamental for the energy demand calculation as well as Clustering and Classification of zones.

1.1.1 Demand database

The demand database contains all of the data required for the calculation of energy demand. They include indicators, benchmarks and conversion factor data for residential and tertiary sectors as well as administrative boundary data (e.g. postcode, MLSOA, GOR). This dataset was collected from many different sources at different resolutions and formats. They have been carefully analysed and structured to prepare a demand database.

A consistent demand database is needed because:

- The large extent (GB) and high resolution (MLSOA) of demand calculation requires consistent structure of the data.
- Working with millions of records (e.g. Marketsafe data includes several tables having millions of data) is not possible with the standard spreadsheet based programs (e.g. MS Excel) or databases (e.g. MS Access).
- Interdependences/relationships among different types of data (e.g. geographic reference) can be better identified through Entity Relationship Models (ERM), which are fundamental to any relational databases.
- The demand database is designed to ensure the flexible nature of input and output data, which means input data can easily be aggregated to the next upper level or the output data can be generated at any preferred level (e.g. MLSOA/IGZ or GOR). Demand outputs can also be generated according to end use (electricity or gas) or to specific sector (e.g. detached house built after 1990 or retail sectors) in any geographic level. In this way, preparation of load curves, clustering and classification can be carried out at the later stage of the project.
- Finally, the database output can be integrated into the GIS data catalogue for visualisation of the results and preparation of maps.

Therefore, the database is needed to ensure the handling of the different and extensive data sets.

Preparation of Demand Database

The raw datasets are mainly collected in .txt, .csv, .xls, .sas format. Their formats have been checked and corrected according to database specification. Afterwards they are imported into the MySQL server to prepare a relationship database. Several steps are required to prepare the demand database (See Figure 1-1).

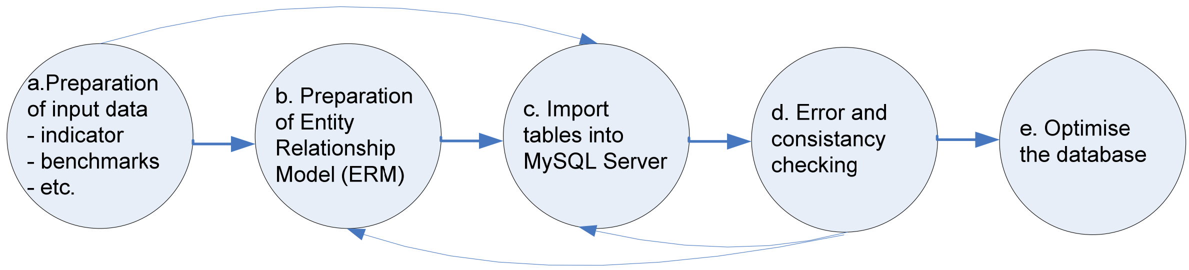


Figure 1-1: Steps required creating demand database

A detailed description of each step follows in this sections.

a. Preparation of input data

Table 1-1 explains the lists of data and their sources that are required to perform the annual energy demand calculation. Many of the datasets have already been described in the trial demand analysis report (see (McKoen, Koch et al. 2011)) whereas some additional data (especially the EDF Energy weather data, Marketsafe data) have been collected to improve the demand calculation method.

Sector	List of tables	Source
Geographic data	Country boundary	Office of National Statistic (ONS)
	GOR boundary	ONS
	GSP regions	EDF Energy
	MLSOA IGZ boundary	ONS
	Postcode boundary	NSPD
Residential sector	Housing stock	EDF Energy
	Age class	EDF Energy
	Housing type	EDF Energy
	Floor area	EDF Energy
	Residential Thermal Benchmark	EDF Energy
	Residential Electric Demand	DECC
Tertiary sector	Company information	Marketsafe data (Creditsafe Business Solution Limited)
	Employee numbers	
	SIC code	

	Post code	
	Premise type	
	Standard Industrial Classification SIC2003	ONS
	TM46	Chartered Institution of Building Services Engineers (CIBSE)
	Conversion	Own Assumption
	Tertiary Thermal Benchmark	CIBSE
	Tertiary Electric Benchmark	CIBSE

Table 1-1: List of required datasets for the demand database

EDF Energy weather data

EDF Energy has provided weather data for about 1000 typical annual scenarios across all the Grid Service Points (GSP) in GB. The scenarios include half hourly temperature and wind speed values for a year (01 May – 30 April). This data is processed by SAS software to determine the median scenario. This will be used to 'weather adjust' the residential and tertiary thermal benchmarks and prepare the profile for the load curve calculation.

Marketsafe data

The Marketsafe data consists of limited and non limited company information in the GB. The company data on the number of employees, post code, SIC, premises types etc. are delivered in several delimited text files which are the basis for calculating tertiary energy demand in the GB. A list of all the limited and non limited bulk data as well as their description are given in (McKoen, Koch et al. 2011).

All these datasets are made compatible to import into the database.

b. Preparation of Entity-Relationship Model (ERM)

Entity-Relationship Model (ERM) is a database modelling method, used to produce a conceptual schema of a relational database. In this project, an ERM is built to explain the abstract and conceptual representation as well as relationship among all the tables. Figure 1-2 describes the ERM diagram for the demand database.

c. Import of tables into MySQL server

The input tables are normalised to reduce the redundancy of data. The goal of database normalization is to decompose relations with anomalies in order to produce smaller, well-structured relations. Afterwards, the bulk datasets are systematically imported into the databases. This import process also includes consistency check of the data formats, naming convention as well as completeness of data retrieval.

d. Error and consistency checking of data

In this step the consistency and completeness of the data have been checked. Inconsistency on data records (especially with Marketsafe data) have been corrected by validating with other sets of data. For example, the SIC information in Marketsafe data do not always match with the official SIC2003 records of Office for National Statistics (see (ONS 2009)). Moreover, the postcode information within the Marketsafe data is sometimes incomplete or missing. They are corrected by mapping with the official postcode data, e.g. National Statistics Postcode Directory (NSPD) datasets.

Some of the database tables (especially the Marketsafe data) consist of millions of records with several columns that are the main input indicators for demand calculation methodology. For this purpose, Structured Query Language (SQL) queries are written and applied within the database schema. In this process, many intermediate tables are generated. The creation of intermediate tables (rather than generating the final output table) ensures the 'consistency check' in the calculation and helps to ensure correct final results (see section 1.4 and Table A-1:).

e. Optimise the database

In the final step, the demand database schema is optimised to perform the operation efficiently. For this purpose, database engineering techniques (e.g. indexing, definition of Primary Key (PK), Foreign Key (FK), etc.) are applied to each table. It ensures optimal design of the schema and that the SQL queries can be used to perform the database calculation, in terms of time and output format.

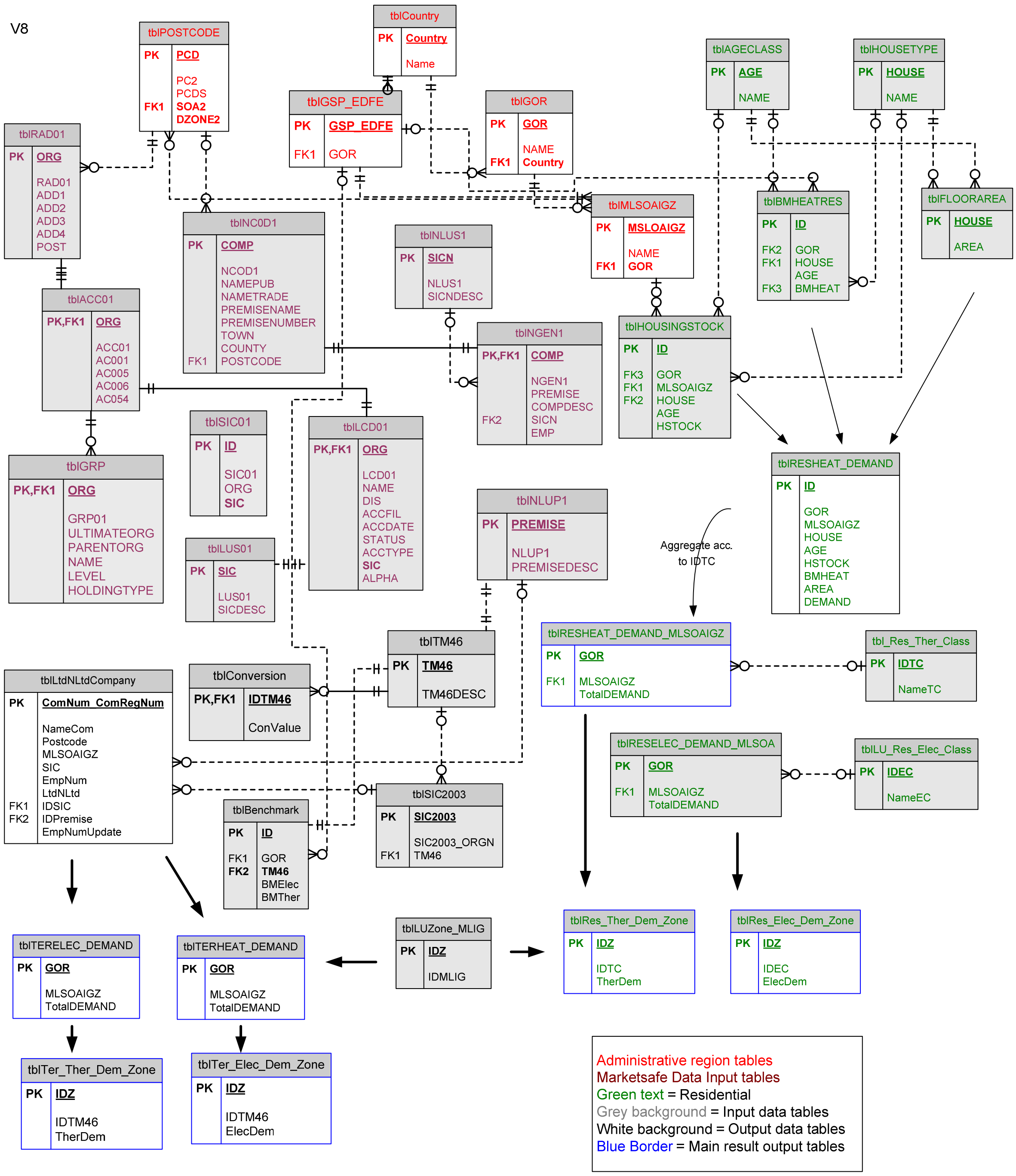


Figure 1-2: Snapshot of the Entity Relation Model (ERM) of the Demand Database

Database platform

The demand database is built upon MySQL Community Server 5.5.13, which is open source database software provided under the GPL License. A physical machine having one 4 GHZ Processor, 2TB of HDD, 4GB of RAM has been dedicated to the demand database.

On the other hand, MatLab, SAS and MS Excel have been used to analyse and process some of the input datasets as well. These have been imported to the MySQL as well.

1.1.2 GIS data catalogue

A GIS data catalogue is prepared to allow visualisation, analysis and mapping of results. For this purpose, GB geographic boundary data (collected from ONS) are merged with the DECC energy consumption data of 2007 to prepare a MLSOA/IGZ level geo-database. The energy demand calculation output from the demand database can also be connected with the GIS data catalogue for further geographic analysis and visualisation of results. For example, to carry out the task of clustering and classification of zones, neighbourhood relationship matrix and shared border length among the neighbouring MLSOA/IGZ are calculated within the GIS.

ArcGIS 10 has been used to perform data preparation, GIS analysis and visualisation of results.

1.1.3 Conclusion

Demand database and GIS data catalogue are fundamental to energy demand calculation in the GB. Therefore, a robust database has been prepared considering the state of the art database techniques, consistency checking of input data as well as completeness of records.

1.2 Improvement of energy demand calculation

The annual thermal and electricity energy demand calculation for the residential and tertiary sectors in the GB are carried out based on the methodology explained in the trial report (see (McKoen, Koch et al. 2011)).

$$Annual\ Energy\ Demand = \sum_{Sector} \sum_{Sites} (indicator\ \times factor\ \times benchmarks)$$

Equation 1-1: Annual energy demand calculation

However, in this report, energy demand calculation methods are refined with the availability of additional granular datasets. Compared to the trial report, following improvements are made:

- Incorporation of GB wide residential and tertiary indicator data to calculate the energy demand
- Improvement of thermal benchmarks by applying ‘weather adjustment’ with EDF Energy weather scenario data (see section 2.2)
- Modification of demand estimation method, especially with the redistribution of number of employees in the tertiary sector (see section 2.4)

Table 1-2 gives an overview of all the thermal and electrical energy demand components for residential and tertiary sector. Each modified demand component is emphasized in bold fonts.

Demand components	Energy use	Residential sector	Tertiary sector	
			Limited	Non limited
Indicator		**GB wide data**	**GB wide data**	**GB wide data**
Benchmark	Electric	Not changed	Not changed	Not changed
	Thermal	**Weather correction with EDFE weather scenario**	**Weather correction with EDFE weather scenario**	**Weather correction with EDFE weather scenario**
Factor		Not changed	Not changed	Not changed
Demand method	Electric	Not changed	**Redistribution of employees rather than restoration of employees**	Not changed
	Thermal	Not changed	**Redistribution of employees rather than restoration of employees**	Not changed

Table 1-2: Improvement of demand estimation method

1.2.1 Description of indicator data

The number of residential buildings within each MLSOA is used as an indicator to calculate residential thermal energy demand. The detailed housing stock data in the GB was collected from the EDF Energy's customer research team. It is categorised according to five housing types and seven age classes (see (McKoen, Koch et al. 2011)).

The number of employees working at tertiary sites in GB is considered as the indicator data for tertiary energy demand. This data has been bought from the Creditsafe Business Solution Limited which offers credit rating related services and data on 4.4 million GB companies in all tertiary sectors. It also includes information on company address, SIC, link to parent company, etc. This data is categorised into two main business activities- limited company and non limited company. A detailed description of the dataset is given in (Creditsafe 2010).

1.2.2 Weather adjustment of benchmark data

Weather adjustment of benchmarks was performed using a formula specified in CIBSE's TM46 guideline (see (CIBSE 2008)).

In the trial demand analysis (see (McKoen, Koch et al. 2011)), weather adjustment to the 35 residential thermal energy benchmarks (derived from the EDF Energy customer energy savings tool) was based on a recent year. In this report, the weather compensation was carried out based on representative regional weather scenario data provided by EDF Energy's meteorology team.

The advantages of refining the benchmark with EDF Energy data are:

- EDF Energy weather scenario data is more detailed - both in spatial (for 17 Grid Supply Points (GSP) regions) and temporal scale (half hourly daily temperature and wind speed value)
- The median scenario can be defined from over a thousand scenario
- This dataset is generated by the EDF Energy's meteorology team and EDF Energy uses this data set for business decision/strategies
- The more detailed a record of temperature data helps to calculate more accurate Heating Degree Day (HDD)

Description of the weather adjustment procedure

The weather data consists of 1010 weather scenarios for each of the 17 Grid Supply Points (GSP) in GB. The detailed procedures to calculate the degree days and then to apply weather corrections to the base benchmarks is elaborately explained in Figure 1-3.

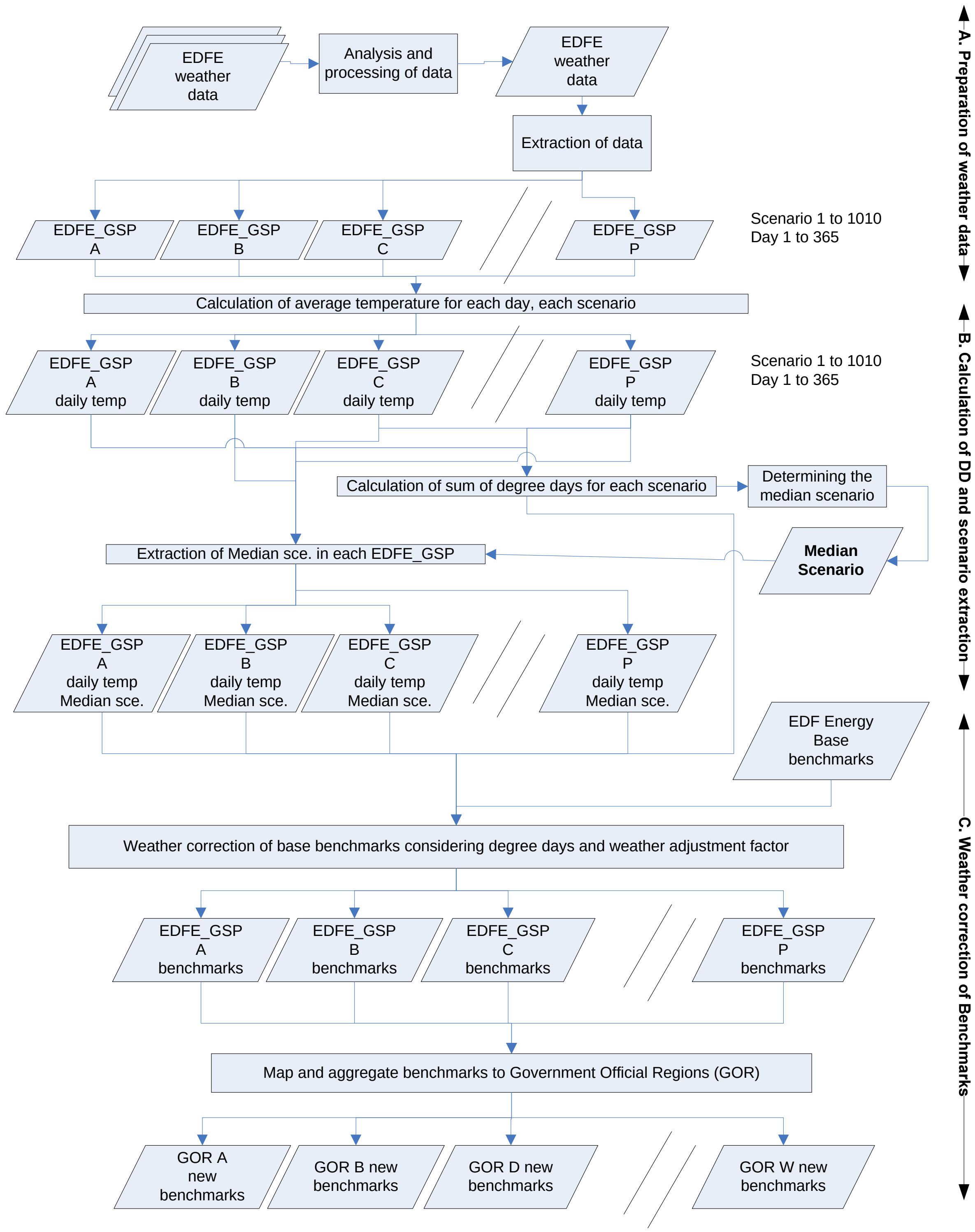


Figure 1-3: Detailed methodology of weather adjustment of the base benchmarks

Three main steps are involved to weather adjust the benchmark data:

a. Preparation of weather data
b. Calculation of degree days and extraction of median scenario
c. Weather correction of benchmarks

a. Preparation of weather data:

The weather data are delivered in 10 SAS files and are worked out using "SAS Enterprise Guide" software. Each file contains 101 weather scenarios and distributed across 17 EDF Energy GSP regions, over a year. These files contain half hourly daily temperature and wind speed values from 01 May to 30 April.

All the files are appended into a single file to combine all the 1010 scenarios. This helps to calculate the degree days for each GSP region.

b. Calculation of degree days and extraction of median scenario

In this step, a median scenario which represents a typical year of GB weather has been identified from the scenarios. It is the closest scenario to the medians of the GSP. Therefore, the following process has been done for every scenario for all 17 GSP:

- Sort scenarios by GSP
- Calculated average daily temperature for each scenario
- Calculate the sum of degree days for each scenario
- Select median scenario

Method to calculate Degree Day

Heating degree days (HDD) are defined relative to a base temperature and are selected with the most appropriate base temperature for a building. In this study, a base temperature of 15.5°C is selected to calculate the HDD, similar to (CIBSE 2008) and recommended by EDF experts. If the daily temperature is smaller than the base temperature (15.5˚) then we calculate the difference between the base temperature and the daily average temperature. The number of degree days can be obtained from the sum of these values over one year. Table 1-3 shows an example output of DD in one GSP for a particular scenario.

Calendar date	AVG of temperature	DD
01.05.	13,04	2,46
02.05.	14,02	1,48
03.05.	13,73	1,77
04.05.	16,02	0
05.05.	13,50	2,00
06.05.	11,36	4,14

Table 1-3: Example of degree day in one GSP within one scenario

Selection of median scenario

Selection of median scenario involves following steps:

- Sum of degree days for all scenarios is calculated.
- Sum of all squared residual (difference between each scenario's sum of degree days and the median sum of degree days for that scenario) is calculated. We calculate a squared residual because in doing the sum, negative and positive values could counter one another and not provide us with the desired scenario closest to the median values.
- The scenario with the minimum sum of all squared residuals is identified.
- Calculate average daily temperature for chosen scenario.

The sums of degree days of the chosen scenario calculated in each of the 17 GSPs are used for weather correction of the residential and tertiary benchmarks.

c. Weather correction of benchmarks

The benchmarks differ according to the geographic location. Therefore, weather adjustments have been applied to the initially developed residential and thermal benchmarks[11], considering the sum of degree days in each GSP region and a weather adjustment factor from the (CIBSE 2008).

Following formula is applied:

$$Weather\ adjusted\ Benchmark_n = Base\ Benchmark\ \times (1 + weather\ adjusting\ factor \times \left(\frac{Sum\ of\ degree\ day\ in\ GSP_n}{Sum\ of\ degree\ days\ in\ base\ benchmark}\right))$$

Equation 1-2: Calculation of weather adjusted benchmarks

[11] Residential and tertiary benchmarks are described in McKoen, K., A. Koch, et al. (2011). Development of a methodology to calculate energy demand. Karlsruhe, Paris, EIFER. **2.0.1**.

Finally, the weather compensated benchmarks in all the GSP regions are mapped and aggregated to the Government official regions (GOR).

1.2.3 Residential energy demand calculation method

The residential energy demand is calculated following the equation 1. However, in this report, the residential thermal energy demand calculation method is improved, considering GB wide EDF Energy's housing stock data and weather adjusted benchmarks (see section 0). Residential electricity demand is extracted from DECC energy consumption data for standard and economy tariff (see (DECC 2008)). The detail description of the basic method is explained in Macro DE trial report (see (McKoen, Koch et al. 2011)).

1.2.4 Tertiary energy demand calculation method

The tertiary energy demand is also calculated based on equation 1, and as explained in (McKoen, Koch et al. 2011). The indicator data, i.e. Marketsafe data on number of employees for the limited and non limited companies across the GB is improved using a redistribution of employee method. Moreover, base thermal and electrical benchmarks are corrected with latest weather adjusted benchmarks. In the following section, the distribution of employee method is explained whereas the weather adjustment procedure is described in the earlier section.

Redistribution of employee numbers

In the trial demand analysis it was investigated that many GB companies (particularly the limited ones) do not register the correct number of employee at every site. Therefore, employee numbers restoration method was applied and tested in a trial MLSOA HG015. In this study, we have applied this method for the whole GB, but the number of employees and thus the calculated energy demand have been too high. An investigation into the Marketsafe data reveals that number of employees is incorrectly restored in many companies and sites. Most of the companies register employee numbers at their headquarters. Therefore, the earlier adopted restoration method (which was successful in HG015) mistakenly allocated large employee numbers. In Harrogate, none of the existing companies have big headquarters, therefore, the restoration method worked. But this method is not suitable for GB wide data.

Hence, a method of redistributing the number of employee has been proposed to allocate the employees away from the headquarters to their subsidiary companies or branches. The main objective of this method is to keep the total number of employees of every company the same. The companies having no branches do not need to redistribute the employee numbers. This method is applied only to the limited companies in GB. The non limited companies do not have any branch and the employee numbers are allocated only to the respective company. For this reason, the employee numbers of the non limited companies do not require redistribution. The validation of this method is described in the next section.

The Marketsafe data is available in many several separate tables/files containing company information on accounting, credit ratings, addresses, etc. One of the tables in limited company bulk data (GRP01 table) contains information on parent organisation and ultimate organisation of a particular company and the hierarchical level where the particular company is located. These companies are identified and an algorithm is developed to distribute the employee numbers from the parent companies to the subsidiaries based on their levels (e.g. from 1 to 15) within the hierarchy of the enterprise.

The distribution procedure is explained in Figure 1-4. The re-distribution procedure is applied to the companies that have subsidiary companies. Therefore, the first step is to identify those companies with subsidiaries and employee data at the headquarters. Then from level 1 to 15, those companies and branches with 0 employees are identified as they require re-distribution. This is done by selecting the subsidiary companies without employees and allocating a share of the parent company's employees considering the average employee of both SICs and total number of employees in the parent company. The following equation is used:

$$\begin{aligned} &Redistributed\ employee\ of\ subsidiary\ company \\ &\quad = \#\ of\ employees\ of\ parent\ company \\ &\quad \times \left(\frac{average\ \#\ of\ employees\ for\ SIC}{avg\ \#\ of\ emp\ of\ parent\ company + \sum_{subsidiaries} avg\ \#\ of\ emp}\right) \end{aligned}$$

Equation 1-3: Calculation of employee numbers during the redistribution

This step is iterated for all the levels and the employee numbers are updated within each step.

The redistribution method of employee numbers is not required for the non limited companies. Therefore, the energy demand calculation method for the non limited companies is only updated with the newly adopted weather benchmark.

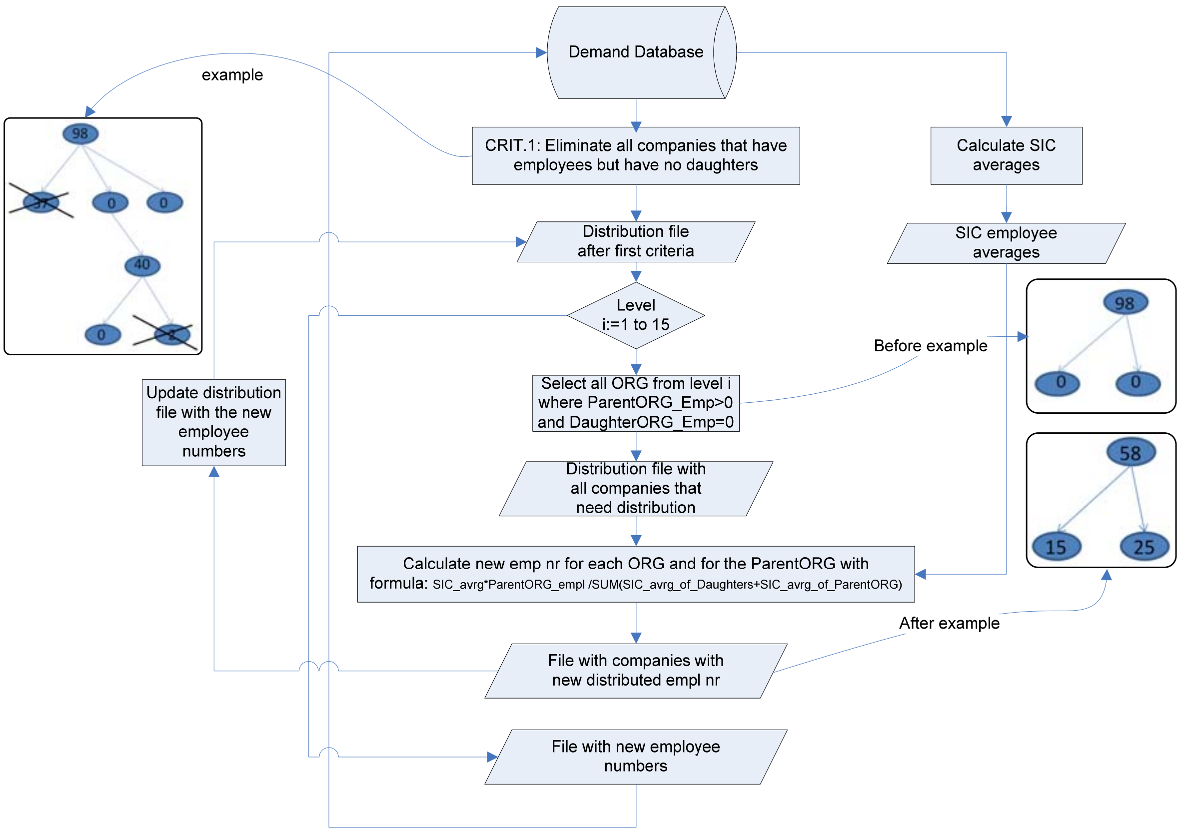


Figure 1-4: Methods of distribution of employee numbers from the limited parent company to its daughters companies

1.2.5 Conclusion

In this chapter improvement of energy demand calculation for the residential and tertiary sectors have been discussed. It is done by incorporating detailed GB wide indicator data, spatially different weather adjusted thermal benchmarks as well as improving limited company employee information through a redistribution process.

1.3 Calculation of energy demand and discussion of results

The calculations of electrical and thermal energy demand in residential and tertiary sectors are carried out following the methodology explained in section 1.2.3 and 1.2.4.

In the GB, there are about 25 million residential buildings which contribute to about 377,614 GWh/yr thermal energy demand. Figure 1-5 shows the distribution of thermal energy demand for different types and age classes of residential buildings.

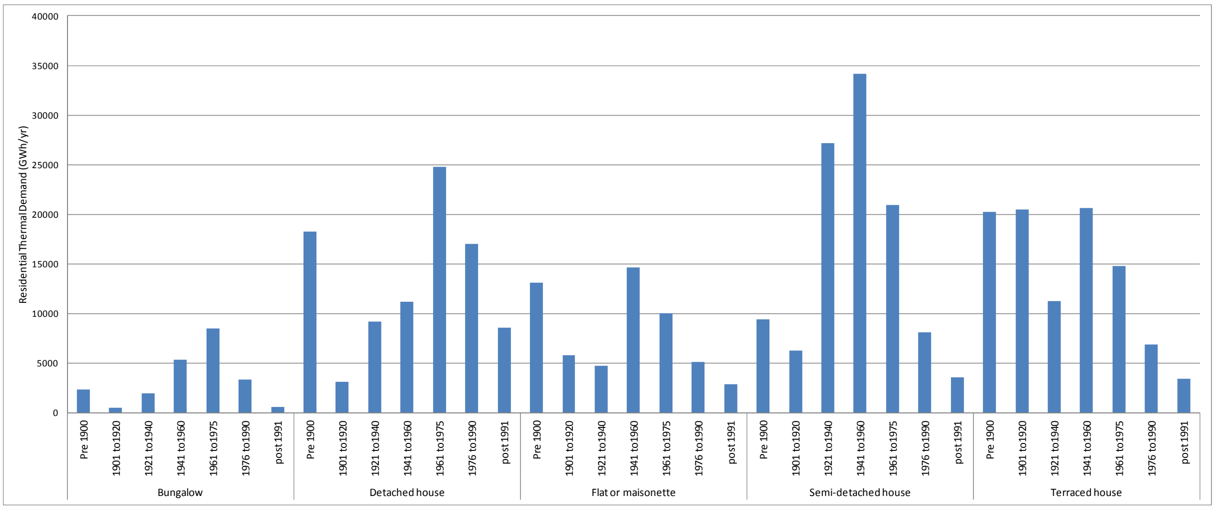


Figure 1-5 Distribution of residential thermal energy demand in GB

The majority of the demand (29%) are in the semi-detached houses, followed by terraced (26%) and detached house (24%) (Figure 1-6). Considering the construction ages of the residential building, most thermal energy demand occurs in the old buildings, e.g. buildings built before 1976 contribute to 84% of the total residential heat demand in the GB (Figure 1-7).

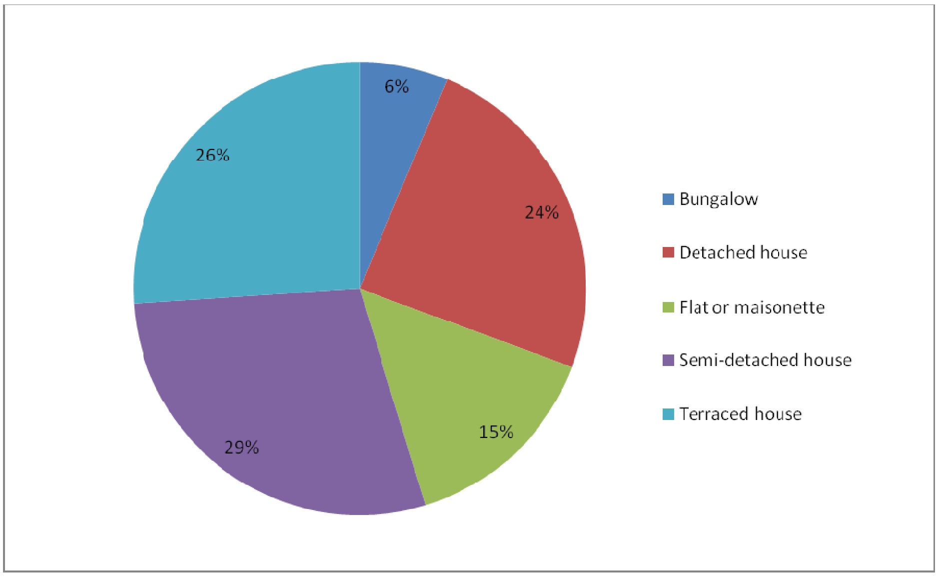


Figure 1-6: Share of residential thermal energy demand according to the house type.

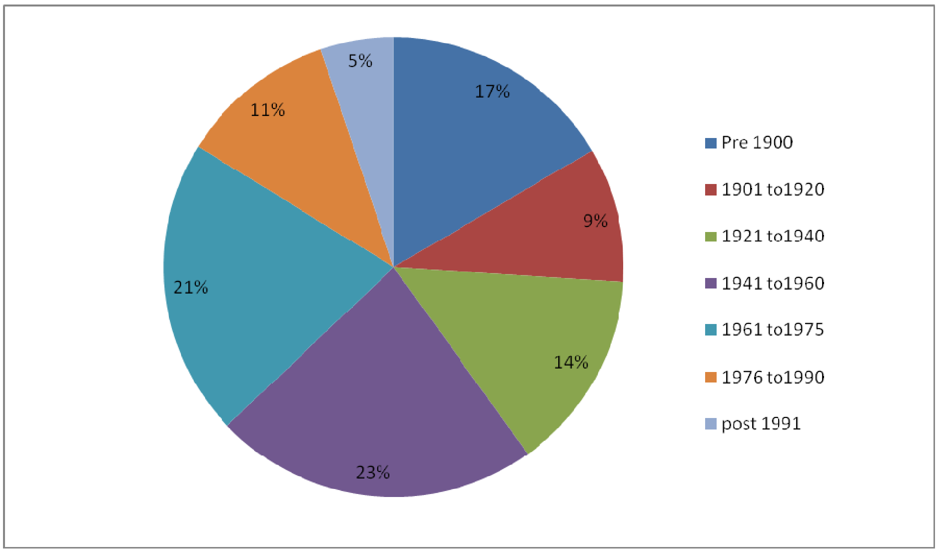


Figure 1-7: Share of residential thermal energy demand according to different age groups in GB.

The geographic distribution pattern of the residential thermal energy demand shows a uniform spread across the MLSOAs in England and Wales. Scotland has significantly low energy demand, compared to the other part of the GB.

Figure C-1: displays the demand at MLSOA/IGZ level.

The total yearly tertiary thermal energy demand in the GB is about 80777 GWh, where schools and seasonal buildings contribute maximum (about 39%) thermal energy demand, followed by hospital (13%) and general offices (10%). Figure 1-8 shows the distribution of total tertiary thermal energy demand for 29 TM46 classes (See Table B-1). Moreover,

Figure C-2 displays the geographical distribution of tertiary thermal demand for each MLSOA in GB.

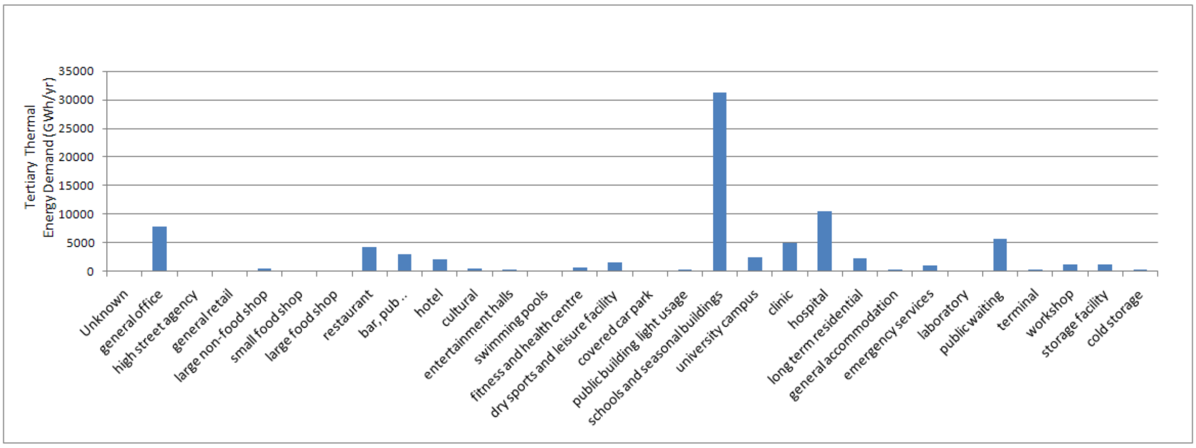


Figure 1-8: Distribution of total tertiary thermal energy demand at each TM46 sectors in GB

The total yearly tertiary electric energy demand calculated in this project is about 82838 GWh. The high street agency contribute to maximum electric demand (68%), followed by schools and seasonal buildings (10%) and general offices (7%). Figure 1-9 shows the comparative picture of the electricity demand for all the 29 TM46 types.

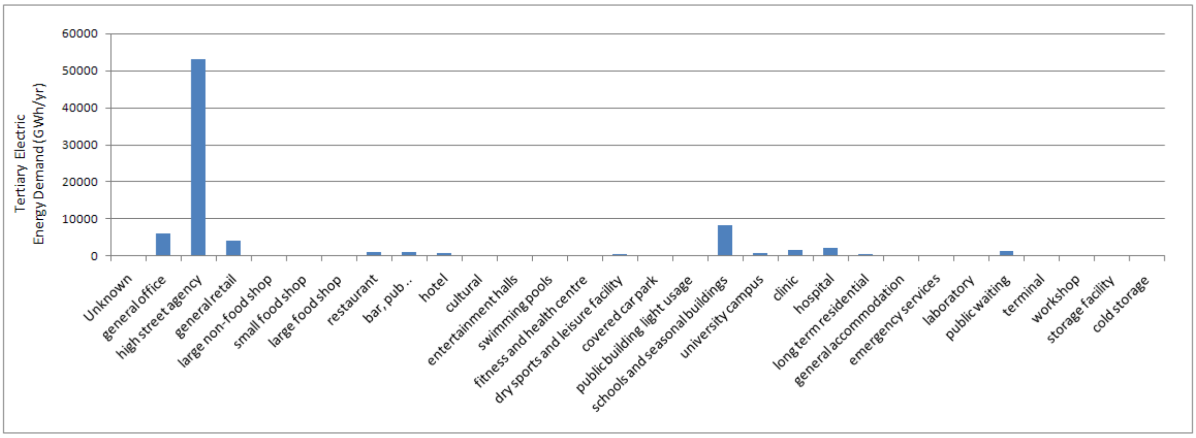


Figure 1-9: Distribution of total tertiary electric energy demand at each TM46 sectors in GB

Total residential and tertiary thermal energy demand in GB is about 458,392 GWh / year. Figure 1-10 describes the geographic patterns of the thermal energy demand density at the MLSOA/IGZ, where it appears that energy density is higher in the big cities.

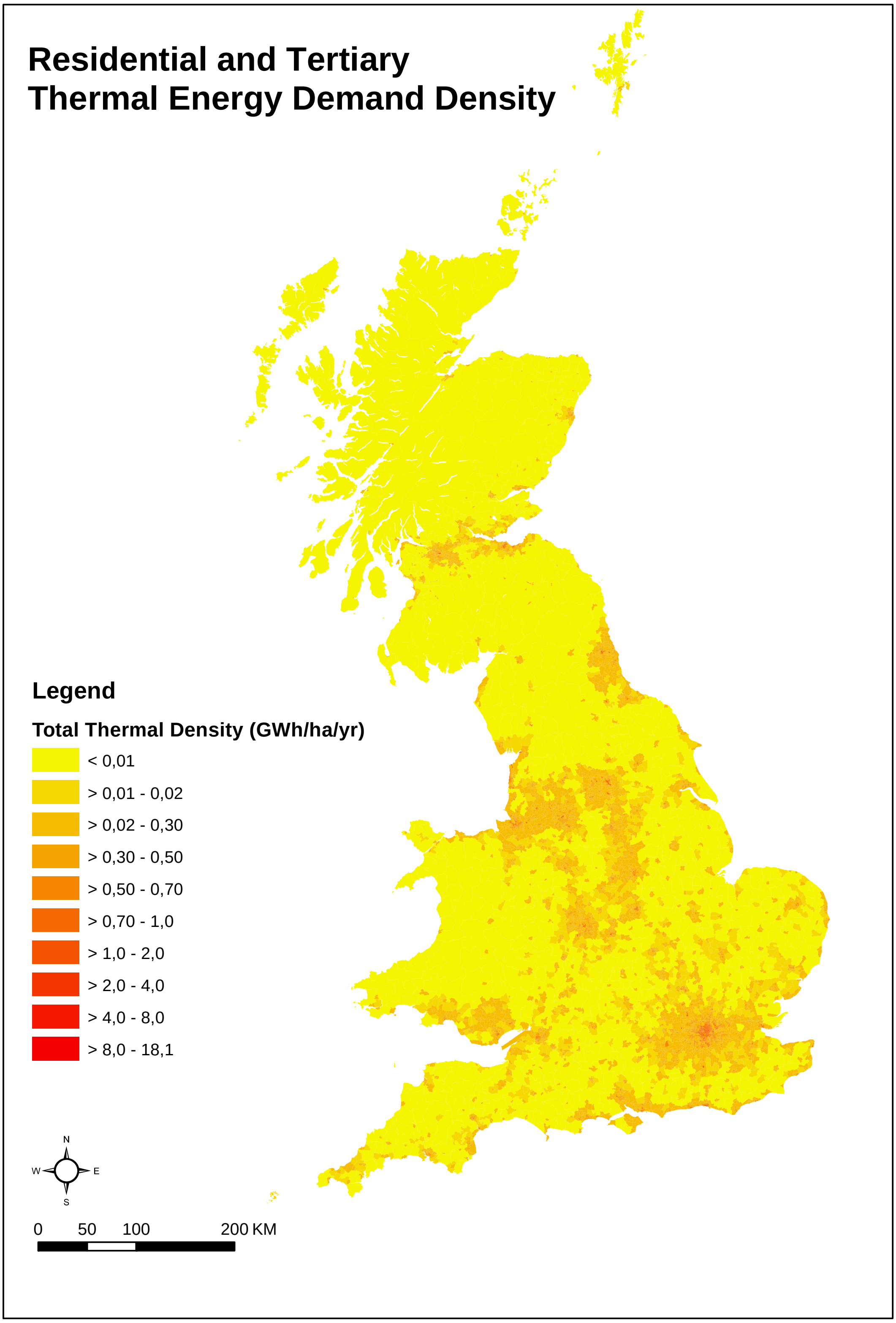


Figure 1-10: Residential and tertiary thermal energy demand density in GB

The total electricity demand in the GB is about 198,747 GWh / year, where residential electricity demand is about 115,908 GWh (calculated from (DECC 2008), as explained in earlier sections). Figure 1-11 describes the electricity demand at the MLSOA/IGZ.

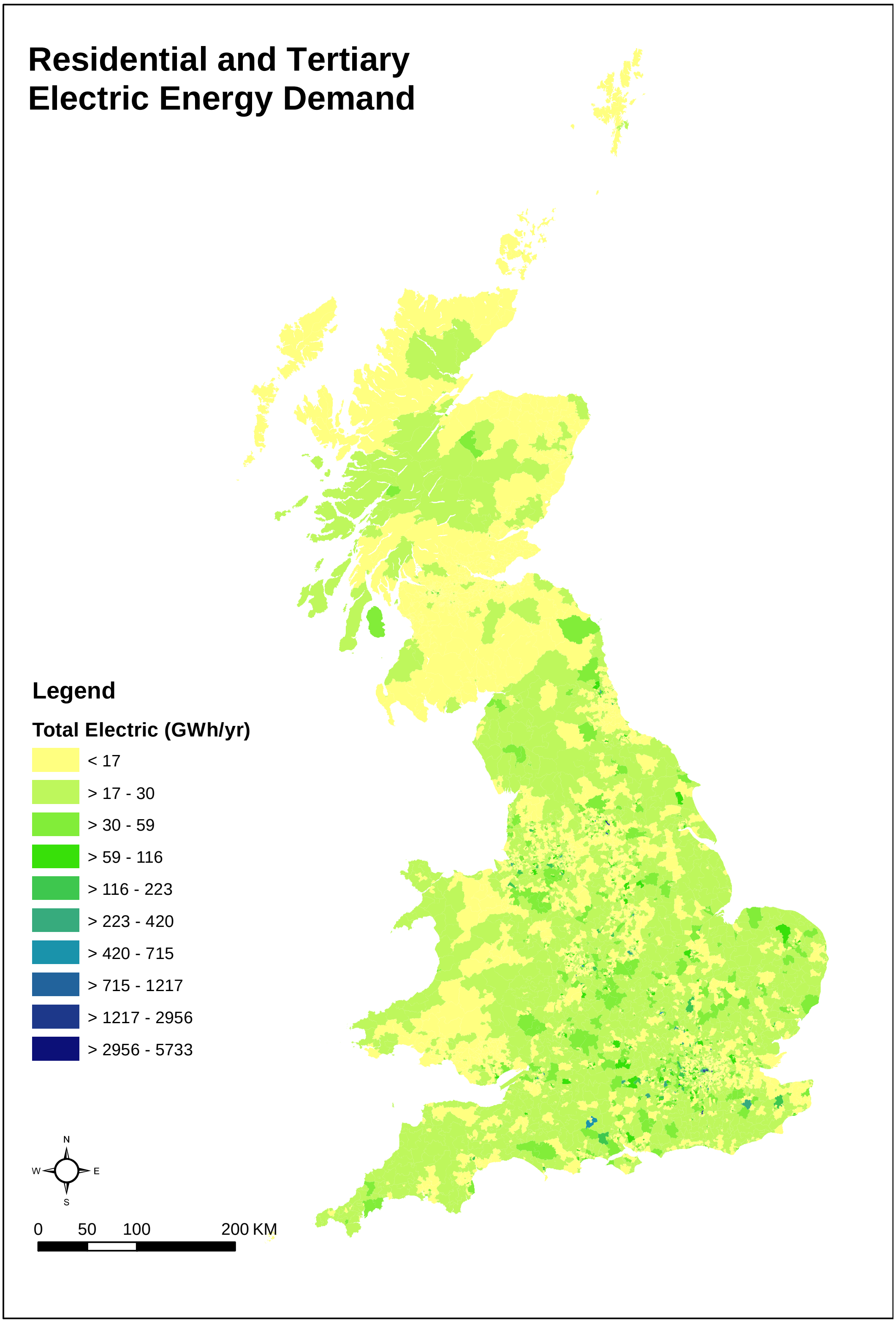


Figure 1-11: Residential and tertiary electric energy demand in GB

Moreover, Figure C-3 and Figure C-4 show the aggregated thermal and electric energy demand density for the residential and tertiary sectors across the MLSOA/IGZs in GB.

1.3.1 Conclusion

Finally, the demand per MLSOA has been calculated for the residential and the tertiary sectors. The results (electricity and thermal) are presented in maps to visualise the geographic pattern of demand in GB. Total yearly electricity and thermal energy demand in the GB is 198,747 GWh and 458,392 GWh respectively.

1.4 Validation of energy demand calculation

Detailed statistical analyses of energy demand components and calculation method as well as validation of results have been explained in the trial report (see (McKoen, Koch et al. 2011)). It is not within the scope of the project to repeat such validations across the whole of the GB.

In this report, however, an investigation into the newly purchased GB wide registered company datasets (e.g. employee numbers, company sites) has been carried out.

1.4.1 Distribution of employees in Marketsafe data

Number of employees working in the limited and non limited companies is the key indicator for energy demand calculation in tertiary sector in the GB. This employee information has been collected from the Marketsafe data where a total of 36 million employees have been recorded in 2010. However, the statistical office in the UK, records about 31m employees (see (ONS 2009)), which also includes Northern Ireland.

For many tertiary sites, the Marketsafe data does not contain employee information (Murshed et al, 2010). In many cases, employee numbers in the subsidiary companies is registered at the headquarters. In this report, the missing employee numbers have been restored by distributing employees from the parent companies to the subsidiaries (see section 1.2.4). Table 1-4 shows such an example which explains that the total number of employees in limited companies within each MLSOAs in Harrogate is kept almost same.

Limited Company			
MLSOA name	Employee numbers		Difference in %
	INITIAL Marketsafe data	Employees after distribution	
Harrogate 001	620	620	0,00%
Harrogate 002	84	98,75	17,56%
Harrogate 003	1413	1544,87	9,33%
Harrogate 004	1842	1553,07	-15,69%
Harrogate 005	110	110	0,00%
Harrogate 006	10	10	0,00%
Harrogate 007	306	306	0,00%
Harrogate 008	287	287	0,00%
Harrogate 009	193	193	0,00%
Harrogate 010	3	3	0,00%
Harrogate 011	144	144	0,00%
Harrogate 012	1410	1390,76	-1,36%
Harrogate 013	1403	1403	0,00%
Harrogate 014	2283	2283	0,00%
Harrogate 015	3964	3984,10	0,51%
Harrogate 016	1273	1288,28	1,20%
Harrogate 017	7858	7651,93	-2,62%
Harrogate 018	7328	7328	0,00%
Harrogate 019	1831	2184,94	19,33%
Harrogate 020	6098	5919,34	-2,93%
Harrogate 021	29	29	0,00%
TOTAL	38489	38332,06	-0,41%

Table 1-4 Comparison of employees in original Marketsafe data and after distribution process with the Harrogate MLSOAs

1.4.2 Distribution of company sites in the GB

In the report "Comparison of Marketsafe data with Geoplan data" it was addressed that many company sites in Marketsafe data do not contain complete SIC, employee numbers and postcode information (see (Murshed, McKoen et al. 2010)). An investigation into the database of Companies House[12] revealed that the data had been originally submitted by the companies in that manner. It is not possible to correct individual company records, but we have identified how many such companies have such missing information (see Appendix 1) and tried to retrieve the company records with supplementary information. Following paragraphs explained the method.

The company information in Marketsafe data is delivered at postcode level which is aggregated to MLSOA level. For this purpose NSPD data which contains postcode references to the MLSOAs is used.

[12] http://www.companieshouse.gov.uk/, where all companies submit company information

Some postcode addresses in Marketsafe data are either incomplete, missing or mismatch with the official NSPD postcode records. Those companies could not be analysed. Some companies also contain partial SIC values or miss SIC information. The demand methodology cannot be applied in the later type of companies.

Moreover, the Marketsafe data is delivered in several separate files which need to be joined using SQL queries and using the unique company identification number (ID) in order to apply the demand methodology. It has also been revealed that not all of the company IDs are available in all the tables which cause missing of some company information. However, the companies with missing employee numbers could not be taken into consideration.

In the limited company information, 2,806,228 distinct companies are recorded of which 95.93% companies are lacking employee information. However, after correcting some records and joining different files, 2,530,617 distinct companies are identified with appropriate information on company registration numbers, postcodes, SICs and employee numbers.

In non limited company information, 1,041,745 distinct companies are identified with correct company registration number, postcodes, SIC and employee numbers from a total of 1,126,503 non limited companies.

Therefore, a total 3,572,362 limited and non limited companies are considered for final demand estimation and 360,369 companies are excluded because of incomplete or missing information on Postcodes, SIC, or Company ID. It is not possible to identify the type and location of these 'missing' companies and therefore, is not possible to quantify the corresponding energy demand.

The detail information on the different files and their descriptions are given in Table A-1:.

1.4.3 Conclusion

In this chapter, the Marketsafe data has been further investigated to identify the number of company sites as well as employee numbers and checked against the statistical number in GB. An overview on the number of missing company sites and employee numbers has been given.

Chapter 2 — Clustering

2.1 Introduction

The clustering tool described in this section aims to identify regions in GB where a district heating network (DHN) supplied by a distributed energy system could be economically and environmentally[13] viable. The economic viability in this context is defined by two factors: The heat density[14] and the total heat demand. The environmental effects will be addressed in the GB benefits case of work package 5.1 together with the economic effects (see (Woods 2011), page 2f). The clustering algorithm is based on middle layer super output area (MLSOA) level. There are 8429 MLSOAs in total in GB. Geographical, social and energetic information has been gathered about them. The clustering tool uses an algorithm that adds MLSOAs together in order to create as many viable distributed energy zones (DE zones) as possible.

2.2 Input Parameters

The following input parameters are needed for the clustering algorithm:

1. Target parameters

a. Maximal zonal heating demand, which is 500 GWh.
   One of the boundary conditions of this project is the upper limit for the electrical power generation installed in the DE zones. This value has been set to 50 MW_e[15] and converted into a general maximum heat demand of a zone of 500 GWh (see Table 2-1). The conversion is based on three different rules of thumb that have been presented in a memo (see (Meidl and McKoen 2010)):

   1. The generation ratio between electricity and heat of a CHP unit is about 1:1.
   2. The size of the CHP unit will typically be designed to run about 5000 hours per year at full capacity.
   3. The CHP unit will satisfy at least 50 % of the zonal heating demand - the additional need for heat will be satisfied by other units such as gas boilers.

   The lower boundary of 100 kW_e was irrelevant for creating DE zones because any single MLSOA already exceeds this value.

b. Target minimum zonal heat demand density (mHDD), here 200,000 kWh/hectare. This figure was agreed on within the consortium as a lower boundary above which district heating is likely to be economically viable. In the literature this value is often higher, e.g. in the Nahwärmekonzept (see (WMBW 2007)), in which a minimum density of 250 ,000 to 300,000 kWh/hectare is recommended.

[13] In terms of CO_2 emissions
[14] Defined as annual heat demand in a given area divided by the total land area
[15] This value has been fixed in the contract and is derived from the Electricity Act which requires power stations above 50MWe to be approved by the Secretary of State whereas stations below this output are subject only to local planning approvals under the Town and Country Planning Act.

It was decided to reduce this value to take non-built up areas of MLSOA into account and to create more zones around the threshold of economical viability. This issue will be analysed in more detail in the benefits case in WP5 (see (Woods 2011)).

Assumption/Statement	Electricity	Heat
Range of CHP units available in this project	100 kW_e to 50 MW_e	-
Heat to electricity ratio of a CHP unit assumed to be 1:1	100 kW_e to 50 MW_e	100 kW_{th} to 500 MW_{th}
A CHP unit should at least run 5000 hours per year at full capacity	500,000 kWh_e to 250,000 MWh_e → 500 MWh_e to 250 GWh_e	500,000 kWh_{th} to 250,000 MWh_{th} → 500 MWh_{th} to 250 GWh_{th}
CHP unit satisfies 50 % of zonal demand	1,000,000 kWh_e to 500,000 MWh_e → 1,000 MWh_e to 500 GWh_e	1,000,000 kWh_{th} to 500,000 MWh_{th} → 1,000 MWh_{th} to 500 GWh_{th}

Table 2-1 Conversion of output power of macro-scale CHP units to zonal annual electricity and heat demand (see (Meidl and McKoen 2010))

In addition to these two requirements, every DE zone's MLSOAs have to be contiguous (in contact with each other) at the end of the runtime. Thus, throughout the MLSOA clustering, it needs to be checked that zones have not been split into parts.

2. MLSOA information

a. A list of all the neighbouring MLSOAs of each MLSOA (realized as a Boolean[16] 8429 x 8429 matrix)
b. Total Heat Demand of each MLSOA (in kWh)
c. Total Area of each MLSOA (in hectare)
d. Energy Density of each MLSOA (in kWh/hectare), calculated out of b. and c.
e. Latitude and longitude of the centres of each MLSOA (GPS coordinates)
f. Matrix of all lengths of common borders between MLSOAs
g. List of lengths of the total outlines of every MLSOA

3. The role of local authorities

It has been widely discussed if the boundaries of local authorities should be respected in the zone creation process. The consortiums has agreed that this is not the case, because the goal of the project is to estimate the impact of district heating by distributed energies will have on the CO_2 emission reductions. Therefore, limitations like regional or local authority boundaries where different local governments are in charge of the investment decision have to be overcome to achieve the highest possible reduction in CO_2 emissions.

[16] A Boolean variable is a data type having two values (e.g.: true and false)

2.3 Clustering Algorithm

The clustering algorithm consists of 9 sequential steps. A flow chart of these steps can be seen in Figure 2-1. They are mainly based on logical operations, but comprise heuristics and fuzzy logic. Some of the steps are iterated recursively.

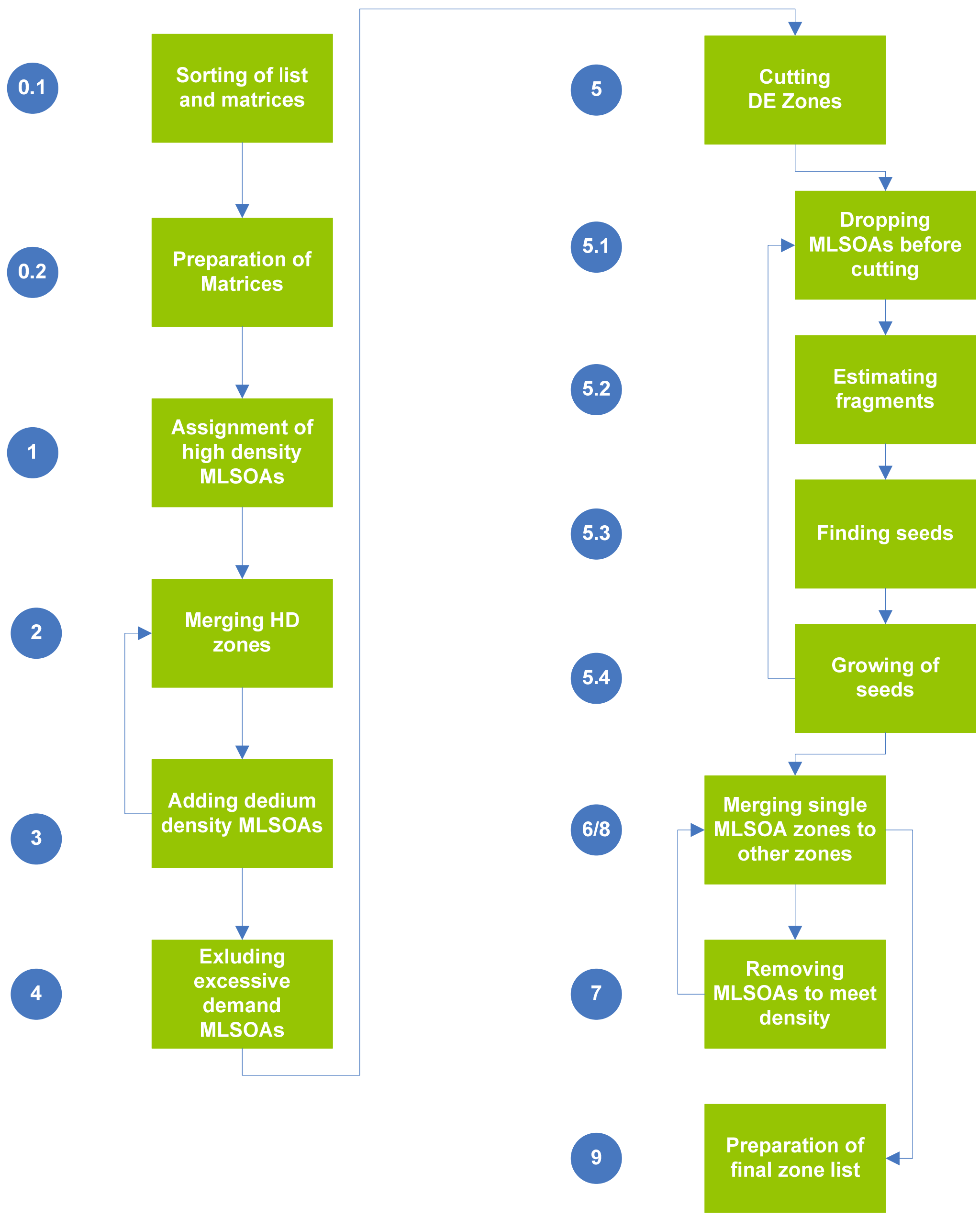


Figure 2-1: Flow chart of the Clustering algorithm

2.3.1 Step 0.1 - Sorting of the lists and matrices

All the input information is represented as lists (vectors) and matrices that are alphanumerically sorted by the MLSOA code in order to have a consistent identification.

2.3.2 Step 0.2 - Preparation of the matrices:

All the steps make use of two Boolean variables that are organized in matrices. The first one - the Assignment Matrix (AM) - holds the information about the assignments of MLSOAs to DE zones. It is used as a decision variable. The second step - the Neighbourhood Matrix (NBHM) – contains information about the neighbourhoods of all the MLSOAs. This matrix serves as an input variable. Both matrices are 8429 rows times 8429 columns large ([number of MLSOA] x [number of MLSOA]).

The Assignment Matrix can be altered during program execution. At the start of the program it is filled with zeros, as no MLSOA is assigned to a zone.
The NBHM holds geographical input parameters that were created by information obtained from a Geographic information system (GIS). This matrix is not changed during execution. The rows and columns of the matrices represent the MLSOAs (alphanumerically sorted by their MLSOA code). The value [true[17]] in the cell of column N and row M of the NBHM means that MLSOA N is a neighbour of MLSOA M and vice versa (the cell of column M and row N also contains [true]). The NBHM is therefore symmetric.

In the AM, the value [true] in the cell of column N and row M means that MLSOA M is assigned to a zone represented by MLSOA N, but not vice versa. The cell of column M and row N contains [false], with the exception of M = N. This process promotes MLSOA N to the kernel of the DE zone (the kernel is not necessarily the geographic centre). In most cases, the MLSOA with the highest thermal demand is defined as the zone kernel. This is not necessarily the case during all the steps of the algorithm, but is represented in that way at the termination of the program.

In an alternative version of the algorithm, the neighbourhood of two MLSOAs will be ignored if their border length is below a pre-defined length. (The NBHM will only contain [true] if the MLSOAs share a border of a certain length). The length is given as a relative parameter that is defined as a length ratio of the common border of the two neighbouring MLSOAs and the larger of the two total MLSOA perimeters.

2.3.3 Step 1 - Assignment of high density MLSOAs

For the remainder of the process, the MLSOAs are categorized into 3 different types:

- High heat demand density MLSOAs
- Medium heat demand density MLSOAs
- Low heat demand density MLSOAs

High heat demand density (HDD) MLSOAs are those MLSOAs that exceed the minimal zonal HDD (mHDD[18]). Medium HDD MLSOAs have a HDD of less than the minimal zonal HDD, but more than half the mHDD. Low HDD MLSOAs have an HDD of less than half the mHDD. These low HDD MLSOAs will not participate in the clustering process.

0.5 mHDD < Medium HDD < mHDD (1)

[17] The value „true“ will be used synonymously to „1“; accordingly the value „false“ corresponds to „0“
[18] Here 200.000 kWh/hectare

In the first step, all high HDD MLSOAs can be directly assigned to a DE zone kernel as they fulfil all criteria to form a DE zone on their own. This section of the algorithm simply assigns each MLSOA, which exceeds the minimal zonal HDD, to itself.

```
for MLSOA=1:8429
        if HeatDensity(MLSOA)<MinimalHeatDensity
                Assignment (MLSOA,MLSOA)=[True]
        end
end
```

2.3.4 Step 2 - Merging HD Zones

Step 2 is analysing the (single MLSOA) zones that were created by step 1. In this step, if two zones neighbour each other they will be combined into one zone.
This is implemented by two nested loops, one running through the rows of Assignment matrix, the other one through the rows of the zone Neighbourhood matrix. Whenever the sum of the scalar product of the two selected columns is bigger than 0, it is identified that two zones neighbour each other. Thus, they will be merged into one zone. (For later iterations, all MLSOAs of both zones will be assigned to the zone that is first in the alphabetical order of them).

2.3.5 Step 3 - Adding Medium Density MLSOAs

All of the MLSOAs that are assigned to zones so far have a HDD higher than the mHDD. Thus all DE zones have a HDD exceeding the mHDD.

Step 3 adds MLSOAs that neighbour the existing zones and have a HDD below the mHDD, but more than half of it (medium density MLSOAs).

A medium HDD MLSOA can only be added to a zone if the resulting total HDD of the zone stays above mHDD.
Because MLSOAs have been added, the program checks for newly neighbouring zones (step 2). If this is the case, the two zones are merged. These steps are executed repetitively until no MLSOA can be added to any DE Zone anymore.

2.3.6 Step 4 - Excluding MLSOAs with excessive heat demand

Without the limitation of a maximal zonal heat demand (MZHD[19]), the final clusters would have been obtained by the end of step 3. The maximal zonal heat demand limitation requires a post processing of the zones to reduce their heat demand. Step 4 separates out all the single MLSOA zones that exceed the MZHD. These MLSOAs show good potential for district heating as they usually have a very high heat demand density, but exceed the limit of the maximal power unit. They will be dealt with separately as outliers (see Appendix J — Outliers).

[19] Here 500 GWh, see section 2.2 Input Parameters

2.3.7 Step 5 - Cutting of DE Zones

After excluding the single MLSOA zones with excessive demand, there will still be several zones of more than one MLSOA left that exceed the maximal zonal demand. Therefore, these zones need to be cut into smaller fragments. The cutting algorithm does not literally "cut" the zone; it splits the zone into smaller clusters of MLSOAs. It does this by defining seeds in the corners of the zones and then growing new zones out of these seeds. The algorithm is separated into several sub-steps:

Step 5.1 - dropping MLSOAs before cutting

First, all zones that exceed the MZHD will be identified. All other zones will be ignored for the further steps. (If no zone exceeds the MZHD, the algorithm will continue with step 6). The algorithm will start by trying to drop the MLSOA with the lowest heat demand density of the zone. If the demand of the zone thereby falls below the maximal zonal demand, no cutting will take place. If the zone demand still exceeds the maximum demand, the dropping will be undone and instead a cutting of the zone will take place.

Step 5.2 estimating fragments

Before an actual cutting takes place, the estimated number of fragments is determined by dividing the zonal demand by the maximal demand (see Equation 2-1).

$$\#\ of\ fragments = ceil\left(\frac{Zonal\ Heat\ Demand}{MZHD}\right)$$

Equation 2-1: Calculation of the number of fragments a DE zone needs to be cut into

If the estimated number of fragments is two[20] or higher than four[21], the algorithm will try to cut the zone into two similar zones (similar in terms of heat demand). If the estimated number of fragments is three or four, the algorithm will try to create three, respectively four similar zones.
The algorithm is reiterated until all zones have been cut down to the maximal zonal heat demand.

Step 5.3 - finding zone seeds

The MLSOA with the largest distance to its geographical zone centre is identified (The geographical zone centres are defined as the average latitude and longitude of all MLSOA centres within one zone). These MLSOAs are announced as first seeds.
Next, the MLSOA of the same zone with the highest distance to the first seed is identified and announced as the second seed. If the zone needs to be cut in three or four pieces, iteratively, the MLSOA that is the most distant from the average of the before announced seeds is announced as third / fourth seed.

[20] In our case, this would mean a total annual thermal demand of 500 GWh
[21] This would mean a total annual thermal demand of more than 2000 GWh

Step 5.4 - growing from seeds

After defining the seeds, the new fragments will start growing from the seeds. The thereby created new zones will be referred to as fragment zones.

The MLSOAs that are part of these new zones will be stored in separate list starting with the seed. All MLSOAs, which will be added later in this step, will follow in this list.

First, all MLSOAs of the original zone (except the seeds) will be put into a pool of available MLSOAs.

Next the following procedure is followed:

One of the fragment zones is chosen and the MLSOA with the highest heat demand which fulfils the following criteria will be identified and added to the fragmental zone:

1. Is part of the pool of available MLSOAs
2. Is neighbouring the fragmental zone.

Afterwards, this MLSOA will be removed from the pool of available MLSOAs.

If the heat demand of the updated fragmental zone is higher than the heat demand of the other fragments, the next fragment is chosen and the procedure repeated, otherwise another MLSOA will be added to the fragment.

If the HDD of a fragmental zone drops below the minimal zonal HDD by adding another MLSOA, or if no MLSOA can be found that fulfils the above mentioned criteria (1. and 2.), the fragment zone will be closed and not further modified.

If all fragmental zones are closed, the "cutting" of the zone is finished and the AM will be updated.

After step 5.4 is completed, the algorithm jumps back to step 5.1.

A visualization of the cutting process of London is displayed in Figure 2-2.

The detailed process is described in the pseudo code below:

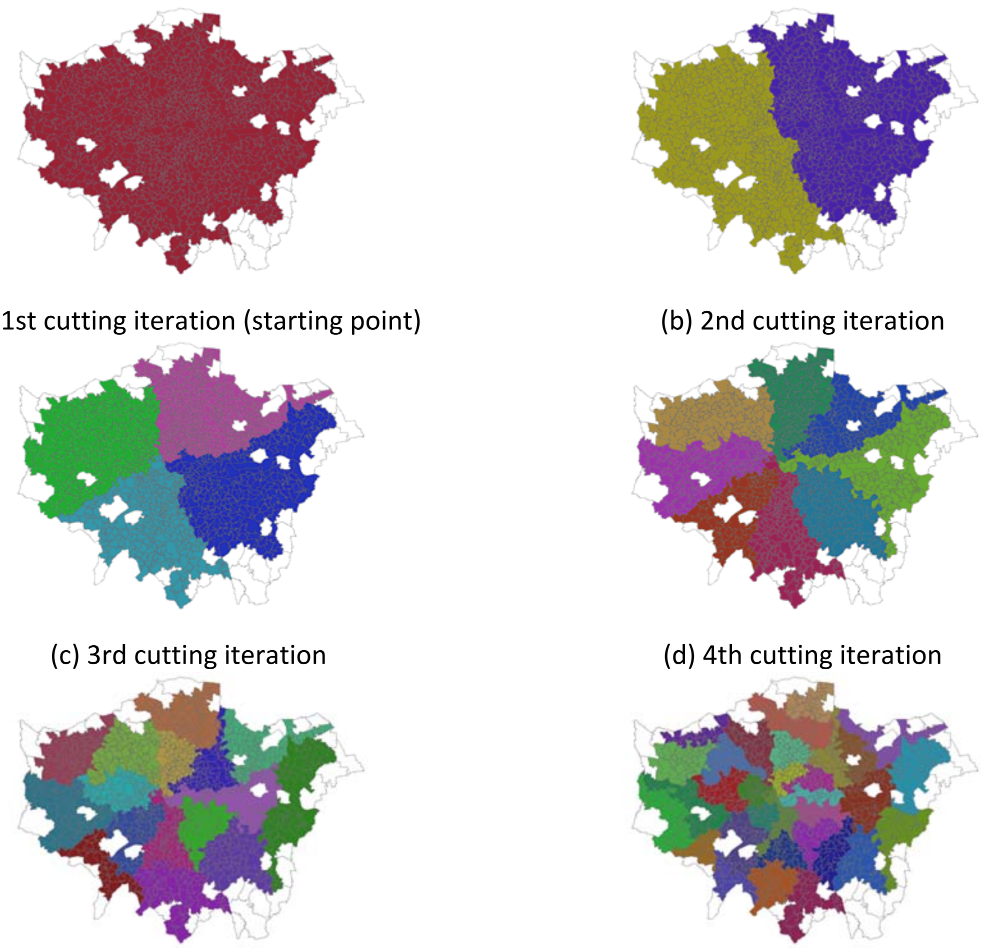
```
1 Choose one fragment zone to start with
2 define i=1
3 find neighbour of Fragment list(i) with the highest Heat demand density, that is part of the old zone
4 if successful==true
5       increase i by one
6       if i is higher than the number of elements in the list
7               jump to step 9
8        else
9               jump back to step 2
10       else
11              Add this neighbour to the new zone
12      Check if the new zone has a lower total heat than all the other fragment zones
13      if true
14              Repeat from step 1
15      else
16              Check if all MLSOAs of the old zone have been added to the fragment zones
17              If true
18                      Stop the growing and continue with the algorithm
19              else
20      Take the next fragment zone and start with 1.
```

2.3.8 Step 6 - merging single MLSOA-Zones to other Zones

Due to the irregular shapes of the MLSOAs, the cutting algorithm might create zones with just one or two MLSOAs. Such zones are not desired. Therefore, step 6 tries to merge zones consisting of one or two MLSOAs with neighbouring Zones.

This process is described in pseudo code below.

1. Find next DE Zone with one or two MLSOAs
2. Find neighbouring zone
3. If there is no neighbour left, continue with step 1
4. Check if the total heat demand of both zones combined is lower than the maximal heat demand
5. If true, merge them
6. If false, find the next neighbour and continue with step 3.



(a) 1st cutting iteration (starting point)
(b) 2nd cutting iteration
(c) 3rd cutting iteration
(d) 4th cutting iteration
(e) 5th cutting iteration
(f) 6th cutting iteration (final result)

Figure 2-2: Cutting iterations at the example of the GOR London

2.3.9 Step 7 - removing MLSOAs to meet Density

The zones met the minimal heat demand density before the cutting.
After cutting, the heat demand density might have decreased due to the creation of fragments. This problem is resolved in the following post process.

1. Find next zone that has HDD lower than the mHDD
2. Find the MLSOA with the lowest HDD in this zone
3. Remove this MLSOA
4. Check if the zonal HDD is lower than the mHDD
5. If true, continue with step 2
6. If false, continue with step 1

2.3.10 Step 8 - merging single MLSOA-Zones to other Zones

After step 7, again some zones might have been created that only have one or two MLSOAs. This is why step 6 is executed once more.

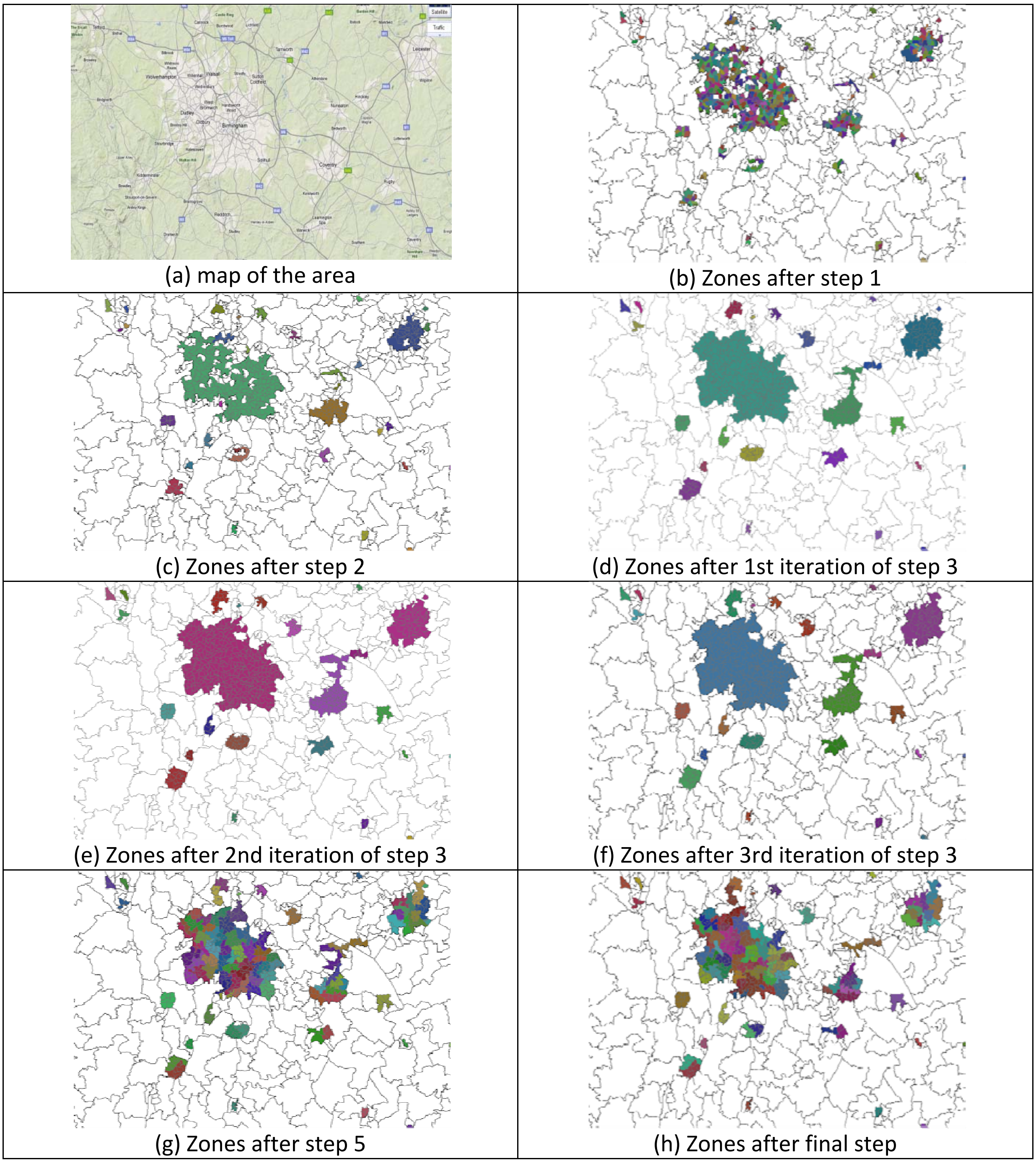
2.3.11 Step 9 - Creating Zones with one single MLSOA exceeding max demand

The MLSOAs that exceed the maximal zonal demand (the ones that have been removed in step 4), will be assigned as single MLSOA zones.

2.4 Output

The output of the algorithm is the final assignment matrix. With the help of this matrix, all further information can be extracted.

An example of the complete clustering algorithm for the region around Birmingham is displayed below in Figure 2-3. Figure 2-3(b) shows the zones of high HDD MLSOAs before the first merging step, which can be found in Figure 2-3(c). The Figure 2-3(d), (e) and (f) iterate the growing and merging algorithm using medium HDD MLSOAs to grow the zones and include more demand within the DE zones. Figure 2-3(g) displays the DE zones after cutting, so that all of them fulfil the maximal demand limit of 500 GWh. During the final step in Figure 2-3(h) MLSOAs of DE zones with an energy density lower than 200,000 kWh/hectare have been excluded, before the single MLSOA zones are added to neighbouring DE zones.



(a) map of the area
(b) Zones after step 1
(c) Zones after step 2
(d) Zones after 1st iteration of step 3
(e) Zones after 2nd iteration of step 3
(f) Zones after 3rd iteration of step 3
(g) Zones after step 5
(h) Zones after final step

Figure 2-3: Zone evolution of the Birmingham region

2.5 Results

During the clustering process 946 zones were created which all fulfilled the three criteria mentioned in the Section 2.2 Input Parameters. A short description of these zones and their attributes can be found in Table 2-2. These zones cover nearly half of the total calculated residential and tertiary heat demand of GB, although they only contain 4 % of the total area. As expected, the areas most suitable for macro scale district heating can be found in city centres or other densely populated areas.

Facts about zones: • 946 zones created • Including about 55,3 % of MLSOAs (4660 out of 8429) • Covering 56.5 % of total calculated residential and tertiary heat demand of GB (258,963 GWh of 458,392 GWh) • Covering about 4 % of the GB area (9,079 km² of 229,012 km²) • 11 zones marked as outliers (total heat demand greater than 500 GWh per year)

Table 2-2: Summary of created DE zones during the clustering algorithm

The two following DE zone histograms show the clustering algorithm results for the two parameters identified as most important: total zonal heat demand (Figure 2-4) and zonal heat density (Figure 2-5).

The first histogram is organised in 50 GWh steps until the maximal zonal heat demand of 500 GWh, where the width of the bins is set to 250 GWh. The majority of the DE zones can be found in the middle heat demand range between 250 and 400 GWh. What is significant is the large number of small DE zones with a total heat demand of 100 GWh or less. These zones can be usually be found in smaller city centres or more rural areas, where only one MLSOA is able to fulfil all three criteria for macro-scale DE zones.

The second histogram in Figure 2-5 shows the distribution of the DE zones according to their energy density. Between 200 and 400 MWh/hectare the width of the bins is set to 25,000 kWh/hectare. Above 400 MWh/hectare the bins get wider, either 50 MWh/hectare for the next two classes or 250 MWh/hectare until 1 GWh per hectare or 500 MWh/hectare for the class above this value. The slope of the histogram nearly follows an exponential distribution with the largest bin between 201 and 225 MWh/hectare, then decreasing rapidly. The small peaks in the two bins 401 – 450 MWh/hectare and 501 – 750 MWh/hectare correspond to the widening of the classes.

The scatter plot in Figure 2-6 shows the large majority of DE zones have a total heat demand below 500 GWh and an energy density below 1 GWh/hectare. The group of six blue diamonds with a higher energy density (y-axis) are very dense city centres with a high ratio of tertiary demand. This set will be further discussed in Chapter 4 — Classification. The eleven red squares to the right of the threshold value of 500 GWh form the group of outliers which will be discussed in Appendix J — Outliers. These single MLSOA high demand zones cannot be split although they exceed the maximum energy demand considered within this project.

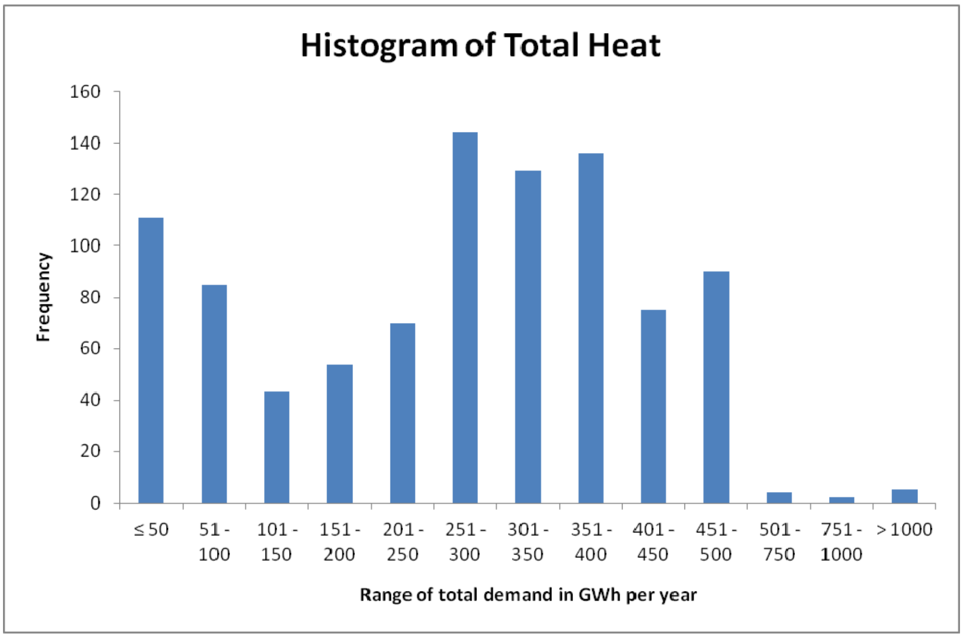


Figure 2-4: Histogram of distribution of total heat demand of zones

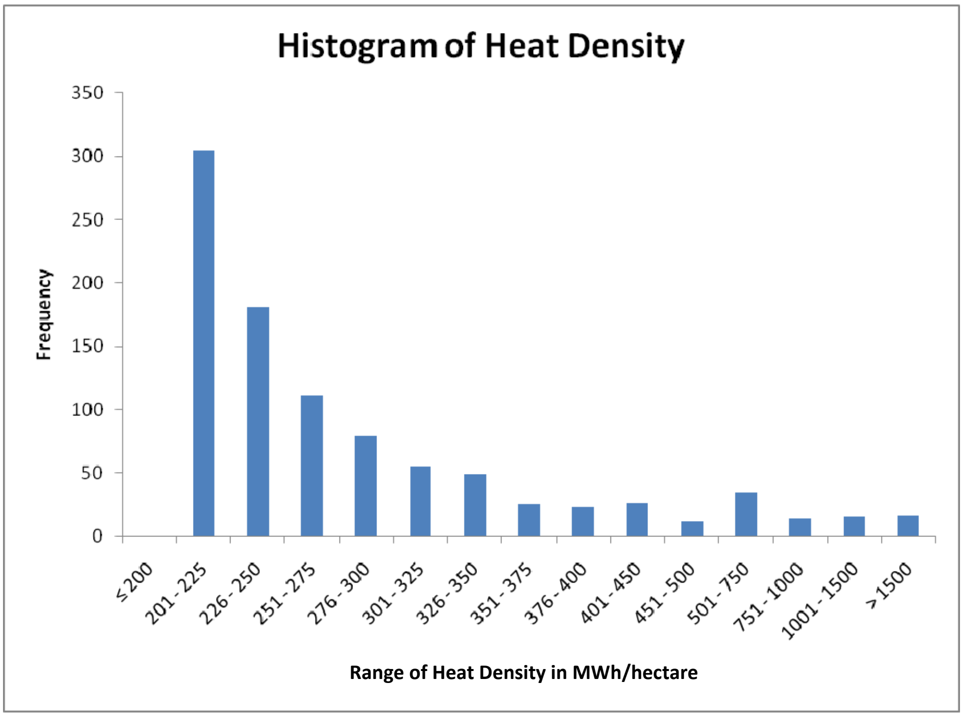


Figure 2-5: Histogram of distribution of heat density of zones
x-axis shows MWh/hectare!

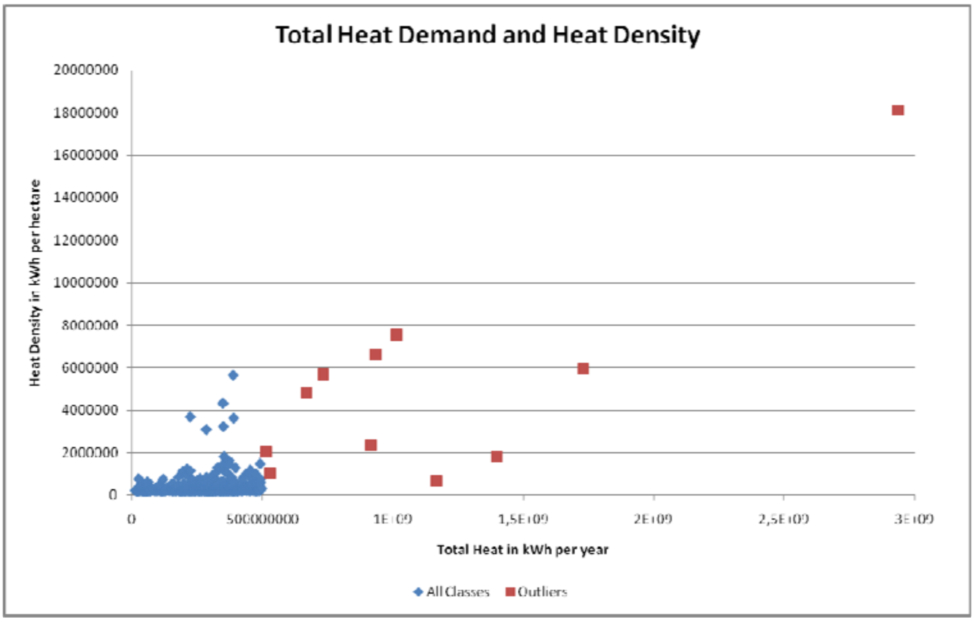


Figure 2-6: Scatter plot of total heat demand and heat density per zone

2.6 Conclusion and Recommendation

District heating supplied by a macro scale distributed energy centre is viable in many different regions in Great Britain. Nearly 50 % of the calculated residential and tertiary heat demand can be satisfied with district heating of this scale. It is obvious that most of the created zones can be found in or around the centres of large or medium sized cities as only 4 % of the total GB area are included in these zones. In these areas the total energy demand and thus the energy density of the area are high enough to run a district heating scheme profitably. For most small towns or villages the residential built-up area can still be beneficial for district heating, but from an MLSOA's perspective which has been used in this project the energy density drops below critical value of 200,000 kWh/hectare. For these areas it can be recommended to do the analysis on a lower scale, e.g. LLSOAs or digital maps directly using GIS, because the MLSOAs tend to include green areas like parks or forests. On the other hand there is the opportunity for the biggest cities in GB which had to be split up into more than one DE zone to combine several district heating networks, either directly into one network or indirectly as several separated ones with connection pipes to each other as a kind of emergency support. The connections can be used to reduce the number of emergency boilers for all networks or during low demand phases in the summer.

In general, it can be concluded that the 948 DE zones that have been created in the project are a first identifier that and where district heating can first reduce the emission of CO_2 and secondly be operated economic viably.

Chapter 3 — Temporal Demand Calculation

3.1 Introduction and Motivation

To estimate the size of the heat supply centre and its optimised run-time schedule the temporal energy demands (heat and electric) aggregated for DE zones (see Chapter 2 — Clustering) have been calculated. These load curves represent the total annual demand split up into 8760 hourly values. Finally, the load curves get simplified into 40 characteristic values which comprise the peak and 39 values to describe a typical week in three seasons (i.e. 13 values for each week in each season). The process of determining these time bands contains five sequential steps:

Step 1) Calculation of annual energy demand per house type and per CIBSE TM46 tertiary building type for each DE zone

Step 2) Preparation of annual profiles[22] (heat and electricity)

Step 3) Mapping of demand types to profile types

Step 4) Calculation of zonal temporal energy demand curves (heat and electrical)

Step 5) Calculation of time bands and characteristic hourly energy demand values (seasonal and total peaks and bases)

The first four steps of this method have been described in detail in the ETI Macro DE Trial Report (see (McKoen, Koch et al. 2011)). These steps have, however, been updated in a number of ways.

3.2 Annual demand calculation

The annual heat and electricity demand were calculated as described in Chapter 1 — Energy Demand Calculation. At this first stage, the energy contributions from different tertiary sectors and residential house types were kept distinct.

First considering the heat demand for every zone, 29 different tertiary contributions and 35 different residential contributions were calculated. The residential data for the number of dwellings per zone and the annual demand per dwelling type was provided by EDF Energy. The tertiary benchmarks were based on a CIBSE publication, while the tertiary indicator data was calculated from the number of employees per SIC in the Marketsafe data. A list of the 29 different CIBSE TM46 building categories is shown in Table B-1 in Appendix B — List of CIBSE Benchmarks.

The calculation of the tertiary electrical demand follows the same methodology as used for the tertiary heat demand. The same CIBSE TM46 categories were used together with data on the number of employees from the Marketsafe dataset. On the other hand, the residential electrical demand was

[22] In this report Profiles are referred to as normalised temporal demand curves whose integral result into 1. The curves of profiles are showing the shapes of the demand, but need to be multiplied by the annual demand to become complete temporal demand curves.

not calculated, but taken directly from data provided by DECC on the electric consumption per MLSOA for both ordinary and economy 7 tariffs (see (DECC 2008)).

The calculated annual heat and electric demands per zone are kept separate for the 29 tertiary sectors and the 35 house types. These consumption types are then mapped to the corresponding profiles (see section 3.4).

3.3 Profile Preparation

The profile sets used in this step were provided by EDF Energy and the German Association of Energy and Water Industries (BDEW). A detailed description of them can be found in the ETI Macro DE Report 2.0 (McKoen, Koch et al. 2011).

3.3.1 Heat Profiles

The heat profiles have been prepared according to the method developed by the TU Munich for the BDEW (see ((BDEW 2006). It distinguishes four residential (single family house old and new, multi-dwelling house old and new) and 15 tertiary profile types, generated according to the same method. For each of these profiles a relative ratio of the annual demand for every day is calculated based on the average outside temperature of that day and the specific slope of a sigmoid function (see Figure 3-1). Subsequently, these values are divided by the sum of all 365 values (see Equation 3-1). The average outside temperatures that are used in this calculation are the same as described in Chapter 1 — Energy Demand Calculation and therefore depend on EDF-Energy's Grid Supply Points (GSP). The averages are taken per calendar day from the hourly values.

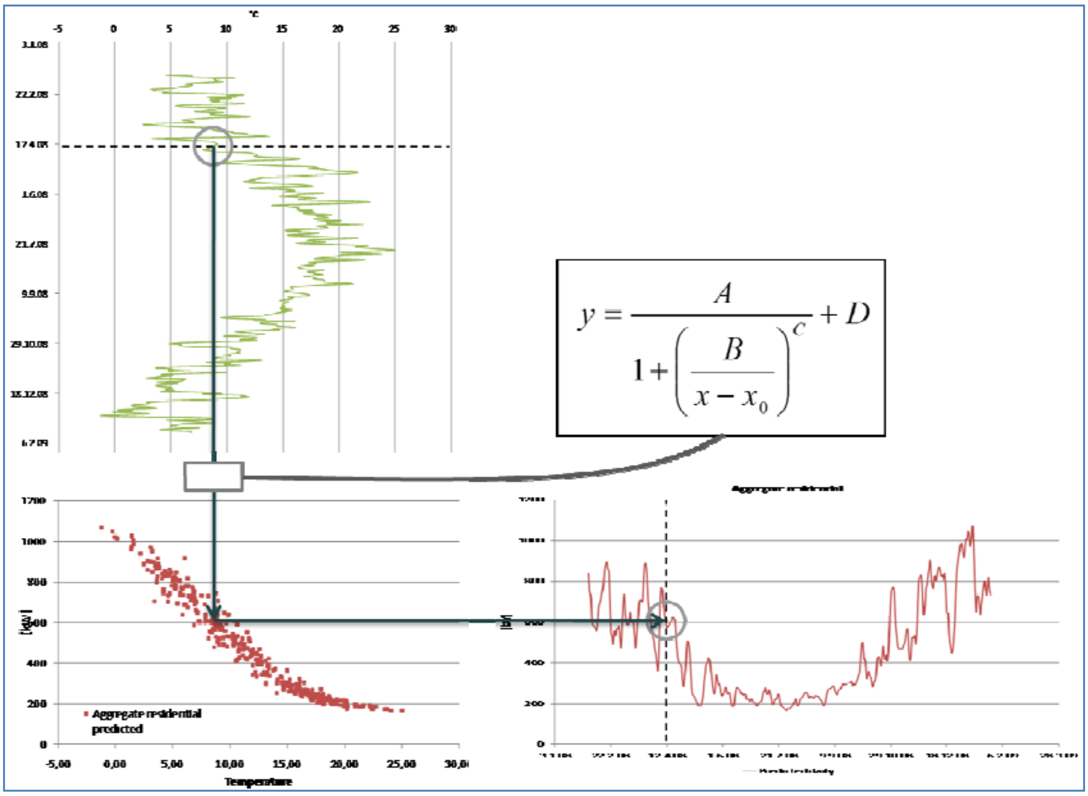


Figure 3-1: Calculation method of the daily heat demand depending on the outside temperature (green curve at the top)

$$daily\ ratio_d = \frac{daily\ value_d(outside\ temperature)}{\sum_{i=1}^{365} daily\ value_i}, d = 1 \dots 365$$

Equation 3-1: Calculation of daily ratios of total heat demand

Hence, 365 different daily values were created, which are normalised to a value of 1. Afterwards a statistical hourly distribution was applied to distribute each daily value into 24 hourly values. These statistical distributions differ from each other by five degree steps depending on the outside temperature. In the end, there are 8760 hourly values, which are also normalised to a value of 1.

3.3.2 Electrical Profiles

The methods to prepare the electrical profiles differ between the tertiary and the residential sectors. The two residential profiles, for the two tariff types ordinary and economy 7, have been directly provided by EDF Energy and consist of 8760 hourly values, which have been normalised, so that their sum is once more equal to one.

The tertiary electric profiles were based on a method of the VDEW[23] (see (BDEW 1999) and (BDEW 2000)). There are 11 different tertiary and agricultural profiles, which in general differ in the operational hours (day, evening, night or 24 hours) and special usage like cooling or heating.

[23] The VDEW has been merged into the BDEW in 2007 together with other organisations.

All of them contain three seasons (winter, summer, transition) and two day types (working day and non-working or weekend day). For each day 15-minutes values were given, which are calculated into hourly averages. Finally, these days are fitted together in corresponding sequences of weekdays and weekend days of the three seasons to generate one year with 8760 hourly values each. In the last step, these sequences were normalised in the same way as previously the daily values in section 3.3.1, only using 8760 hourly values this time (see Equation 3-2).

$$hourly\ ratio_h = \frac{hourly\ value_h}{\sum_{i=1}^{8760} daily\ value_i}, h = 1 \dots 8760$$

Equation 3-2: Calculation of hourly ratios of tertiary electrical demand for the electrical profiles

3.4 Mapping of Profiles

The next step is to bring the annual demand sectors and the prepared profiles together. Since the categories do not fit together completely, a judgement had to be made to match the existing demand categories to the existing profiles.

3.4.1 Mapping of the Residential Sector

The electrical part of the residential sector has been very easy to handle. There were two profiles provided by EDF Energy for the ordinary and the economy 7 tariff and two zonal annual demand figures for the same two categories.

The thermal mapping has been more difficult, because the 35 different house types needed to be grouped into the 4 different profiles for the residential heat demand as the results show in detail in Table 3-1. In general, all building types except the flats or maisonettes are dedicated to single dwelling houses (SFH) which cover all residential for one or two families. The flats or maisonettes are linked to the multi-dwelling house profile (MFH). The cut-off year between old and new buildings has been set to 1976, because larger renovation cycles in the 1980s and the building act of 1984 have been introduced.

House Type Name	Pre 1900	1901 to1920	1921 to1940	1941 to1960	1961 to1975	1976 to1990	post 1991
Bungalow	SFH old	SFH old	SFH old	SFH old	SFH old	SFH new	SFH new
Detached house	SFH old	SFH old	SFH old	SFH old	SFH old	SFH new	SFH new
Semi-detached house	SFH old	SFH old	SFH old	SFH old	SFH old	SFH new	SFH new
Terraced house	SFH old	SFH old	SFH old	SFH old	SFH old	SFH new	SFH new
Flat or maisonette	MFH old	MFH old	MFH old	MFH old	MFH old	MFH new	MFH new

Table 3-1: Mapping of house types to residential heat profiles

3.4.2 Mapping of the Tertiary Sector

The tertiary TM46 categories have been mapped to the heat and electrical profiles of the BDEW. Some can be used for more than one TM46 categories. For the heat profiles an existing mapping for more than one thousand different tertiary, industrial and public sectors in Germany to the 15 profiles could be adapted. There have still been some inconclusive combinations for which a mix of two or three profiles has been used.

The mapping of the electrical profile has been executed in a different way, because the electrical profile primarily depends on the opening hours of the companies and the demand for heating and cooling devices, e.g. deep freezers or cold storage rooms which are running 24 hours a day. There was also a significant difference in the demand pattern if offices or shops needed to be illuminated the whole day and night.

3.5 Calculation of Zonal Load Curves

3.5.1 Thermal load curves

To calculate the thermal load curves of a zone, the residential and tertiary parts have to be calculated separately and afterwards summed together. The 35 house types were mapped into the 4 different profile sets and each hourly relative demand value multiplied with the annual demand figure to become the total residential demand of this particular hour.

The thermal tertiary load curves have been prepared in the same way with 15 tertiary heat profiles into which the 29 TM46 sectors are mapped.

The summation of the residential and the tertiary load curves created the all zonal heat load curve. The calculation is shown in Equation 3-3.

$$ZLCV_h^{Heat} = \sum_{\substack{House\ Type \\ n=1}}^{5} \sum_{\substack{Age\ Class \\ m=1}}^{7} AD_{nm}^{ri} \times Profile_{nm}^{rh} + \sum_{\substack{Tertiary\ Categories \\ k=1}}^{29} AD_k^i \times Profile_k^{rh},$$

$$with \sum_{\substack{hour \\ h=1}}^{8760} ZLCV_h^{Heat} = total\ zonal\ heat\ demand$$

Equation 3-3: Calculation of thermal zonal load curve

With

$ZLCV_h$	Zonal load curve value of hour h
AD-	Annual demand of zone i in GOR r for different residential or tertiary categories
h	Hour
n, m, k	Counters
r	GOR (region)
i	zone number

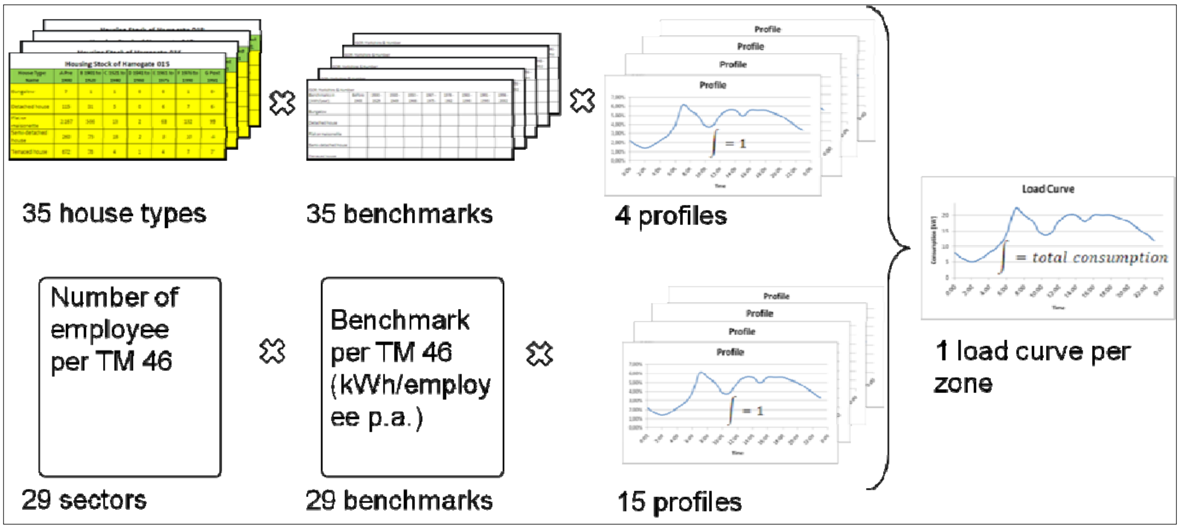


Figure 3-2: Graphical explanation of the heat load curve calculation process for residential (upper) and tertiary (below)

3.5.2 Electrical load curves

On the tertiary side the eleven tertiary and agricultural profiles are matched to the 29 TM46 sectors in the same way as the heat profiles and the load curves of the sectors have been calculated. For the residential load curves the two profiles from EDF Energy have been multiplied with the domestic electricity for both tariffs given by DECC.

The electrical load curves per zone were calculated by the summation of all tertiary electrical load curves and the domestic one.

The calculation of the electrical load curve is shown in Equation 3-4.

$$ZLCV_h^{Elec} = \sum_{\substack{Tariff\ Type \\ n=1}}^{2} AD_n^{DECC,i} \times Profile_n^{rh} + \sum_{\substack{Tertiary\ Categories \\ k=1}}^{29} AD_k^i \times Profile_k^h\ ,$$

$$with \sum_{\substack{hour \\ h=1}}^{8760} ZLCV_h^{Elec} = total\ zonal\ electrical\ demand$$

Equation 3-4: Calculation of electrical zonal load curve

With

$ZLCV_h$	Zonal load curve value of hour h
AD	Annual demand of zone i in GOR r for different residential or tertiary categories
h	Hour
n, m, k	Counters
r	GOR (region)
i	zone number

3.6 Calculation of Time Bands

To use the load curves within the optimisation tool, the data needs to be split up into time bands defined by the University of Manchester. The 8760 hourly values are simplified into 39 characteristic values. These are differentiated by three seasons (summer, winter, and transition) and two day types with either seven different time bands on a weekday and six different time bands on a weekend day. The demand per time band is given as hourly values, thus the units are kWh/h.

Each value was calculated as the average of all hourly values that were sorted into this time band of the particular day type and season. The complete set of time bands can be found in Table 3-2. This example is valid for all three seasons.

<table>
<tr><th>Time Bands</th><th>Weekday</th><th>Time Bands</th><th>Weekend</th></tr>
<tr><td>WD 1</td><td>0:00 - 04:59</td><td>WE 1</td><td>0:00 - 04:59</td></tr>
<tr><td>WD 2</td><td>5:00 - 6:59</td><td>WE 2</td><td>5:00 - 6:59</td></tr>
<tr><td>WD 3</td><td>7:00 - 9:59</td><td>WE 3</td><td>7:00 - 9:59</td></tr>
<tr><td>WD 4</td><td>10:00 - 16:59</td><td>WE 4</td><td>10:00 - 16:59</td></tr>
<tr><td>WD 5</td><td>17:00 - 18:59</td><td rowspan="2">WE 5</td><td rowspan="2">17:00 - 21:59</td></tr>
<tr><td>WD 6</td><td>19:00 - 21:59</td></tr>
<tr><td>WD 7</td><td>22:00 - 23:59</td><td>WE 6</td><td>22:00 - 23:59</td></tr>
</table>

Table 3-2: Defined time bands for a weekday and a weekend day for a given season

To estimate the maximum demand that can occur in a DE zone, the maximum value is taken from the load curve as the peak value. This figure is representing the highest demand throughout the year for one hour. Thus, in a typical year that is used in the calculation herein, this demand could not be exceeded.

Chapter 4 — Classification

4.1 Aim of classification

The classification aims to group DE zones with similar attributes together to find commonalities in different part of GB and reduce the number of DE zones to be analysed in the benefits case of WP 5.1. Up to 25 classes would be created from about 1000 DE zones, thus on up to 40 zones will be grouped together to form a class.

Out of each class one zone will be extracted which will be used as representative for its class. These zones are selected to be as close to the centre of their classes as possible and will be called Characteristic Zones.

4.2 Methodology –general description

The main goal of classification is to find similarity or dissimilarity between objects and next to form the group (class, clusters[24]) from the similar elements. Each class must contain elements which are similar to each other within the given class and at the same time, elements (objects) which belong to different classes must be as different as possible.

The typical cluster analysis as it is used for the classification process in this case covers the following steps:

1. Selection of objects and variables (parameters)
2. Standardization of variables values
3. Selection of distance measure
4. Selection of classification method
5. Determination of number of classes
6. Evaluation of classification results
7. Description (interpretation) and validation

Cluster analysis is generally the process to assign a set of objects into groups. These groups are supposed to be different from each other, but the members of each group to be similar. It is often used for data mining and knowledge discovering within large databases.

4.2.1 Objects and parameters selection

The selection of appropriate classification variables is one of the most important and difficult tasks in classification. The reliability of the final results and the correctness of the taken decisions depends on the quality of the set of variables (parameters) that are selected for classification. Only the parameters that allow a discrimination to be made within the set of objects should be taken into account in the classification process.

The selection of the classification parameters occurs in two steps: (1) selection of initial list of parameters and (2) reduction of the initial list of parameters using different well-established

[24] although this problem is often referred to as *Cluster Analysis*, we shall refer to it as a *Classification* problem to avoid confusion with the clustering of MLSOAs into DE zones described in Chapter 2 — Clustering.

algorithms. The first step is based on experts' experiences and knowledge. In the second step, the grouping of parameters is carried out by factor analysis. The measure of this is the Pearson linear correlation coefficient. In addition the Principal Factor Analysis method was used to determine the appropriate set of parameters. The goal of this procedure is to choose a set of parameters that are not correlated to each other and represent the groups of parameters. The Principal Factor Analysis helps to define how many parameters are needed in the classification and allows the initial number of parameters to be reduced. There are several tests for defining the number of principal factors that exist. In this study the Eigen values criterion and Scree plot was used.

4.2.2 Standardisation

The methods of statistical multidimensional analysis used in this classification approach need to transform all the classification parameters to a comparable scale. This process is known as standardisation, and it converts the original measures to unit-less variables. After transformation, the resulting values have a mean equal to zero and a standard deviation equal to one.

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}$$

Where:

z_{ij}	The standard value of the *j*-th variable for the *i*-th object
x_{ij}	The value of the *j*-th variable for the *i*-th object to be standardized
$\bar{x}_j$	The mean of the *j*-th variable
S_j	The standard deviation of the *j*-th variable

4.2.3 Distance measure

The next step is to compute distances between the objects in order to quantify their degree of dissimilarity. A distance for each pair of object must be calculated.

The distance between objects is usually calculated as the Euclidean distance:

$$d(x_i, x_k) = d_{ik} = \sqrt{\sum_{j=1}^{p} (x_{ij} - x_{kj})^2}$$

Equation 4-1: Euclidean distance

Where:

d_{ik} is the geometric distance between the points with coordinates x_{ij} and x_{kj}.

The most general formula of distance is Minkowski distance:

$$d(x_i, x_k) = d_{ik} = \left(\sum_{j=1}^{m} \left| x_{1j} - x_{2j} \right|^p \right)^{1/p}$$

Equation 4-2: Minkowski distance

Where:

p is any real number larger than or equal to 1. For the Euclidean distance p=2.

x_{1j} and x_{2j} are coordinates of points.

4.2.4 Classification methods

Several classification algorithms exist - the two best known are: (1) partitioning and (2) hierarchical methods.

A partitioning method constructs a predefined number of classes. It classifies the objects into k groups, which together satisfy the requirements of a partition. Each group must contain at least one object and each object must belong to exactly one group. This type of classification method uses an optimizing algorithm with iteration. The algorithm starts with randomly selected objects as a starting, so-called “seed”, groups and in following iterations assigns the remaining objects to the groups, trying to improve the localization of each element in each group.

The representative in SAS[25] of this classification algorithm is the k-means method, which is one of the best known and the most described in the literature of all methods in this category.

Hierarchical algorithms do not construct a single partition with k groups but they deal with all values of k in the same run. There are two kinds of hierarchical algorithms: (1) the agglomerative (bottom up) and (2) the divisive (top down). Agglomerative methods start with all objects in separate groups and in each step two groups are merged, until only one is left. Divisive methods start with all objects as one group and in each step split it up, until there are n groups.

Different types of classification methods calculate the distance between classes in different ways. Usually the name of the method comes from the name of the distance calculating method it uses. Some of the possible methods of classification are described below:

- Single linkage, Nearest neighbour
- Complete linkage, Furthest neighbour

[25] SAS (Statistical Analysis System) is an integrated Business Intelligence system of different analytical tools. One of them is statistical analysis software, which provides comprehensive statistical tools for a wide range of statistical analyses, including descriptive statistics, multivariate analysis like Factor Analysis, Cluster Analysis, Discriminant Analysis and many others.
SAS uses a SAS Enterprise Guide - a Microsoft Windows client application as the default interface for business analysts, statisticians, and programmers.
Besides SAS several others statistical packages exists like commercial Statistica, SPSS or “R”.

- Median classification
- Centroid classification
- Average linkage within groups
- Ward's Minimum-Variance method

In the Single linkage method, the distance between two classes is the minimum distance between an observation in one class and an observation in the other class.

In the Complete linkage method, the distance between two classes is the maximum distance between an observation in one class and an observation in the other class. The Complete linkage method is strongly biased toward producing classes with roughly equal diameters.

In the centroid method, the distance between two classes is defined as the Euclidean distance between their centroids or means.

In average linkage method, the distance between two classes is the average distance between pairs of observations, one in each class. This method tends to join classes with small variances and it is slightly biased toward producing classes with the same variance.

Ward's method is strongly biased toward producing classes with roughly the same number of observations. At each class generation, the within-class sum of squares is minimized over all partitions obtainable by merging two classes from the previous generation.

4.2.5 Determination of number of classes

The total number of classes was arbitrarily defined at the beginning of the process of classification. However, hierarchical methods offer several tools that help to establish the possible number of classes. One of them is the dendrogram – a special kind of graph that indicates the possible cutting point of the initial set of object agglomerations using the relative distance between the groups of objects (classes).

In the literature, one can find the different indices (coefficients) for evaluating the best number of classes. In SAS Enterprise Guide 4.2 software, developed by SAS Institute, the CCC (Cubic Classification Criterion) coefficient, Pseudo-T-Quadrate and Pseudo F test are used.

The CCC criterion is formulated in this way:

$$CCC = \ln\left(\frac{1-\mathrm{E}\left(R^2\right)}{1-R^2}\right)\cdot K$$

Where:

$$K = \frac{\sqrt{\frac{np}{2}}}{\left(0{,}001+\mathrm{E}\left(R^2\right)\right)^{1,2}}$$

Where:

- K is a coefficient, empirically obtained using a Monte Carlo simulation.
- p is the largest integer that is less than the number of classes

The CCC criterion can be interpreted as follow:

- CCC's positive values mean that the set of investigated objects can be divided into groups;
- the CCC's variation with the number of classes indicates the optimal numbers of possible classes that the set can be divided into;
- the best number of classes is the one for which CCC has a global or local maximum value between 2 or 3;
- very low values of CCC mean that outliers (not typical objects) exist;
- the existence of many local maxima indicates that the set of objects have a hierarchical structure.

4.2.6 Evaluation

In order to describe the set of objects systematically, several statistical parameters coming from descriptive statistics are used to evaluate the results of the classification:

- Mean
- Standard deviation
- Variance
- Coefficient of variance[26]
- Dispersion (min-max analysis)
- Range
- Kurtosis
- Skewness
- Quartiles

Variance, range and standard deviation are typical measures of variability and dispersion. Mean is a measure of central tendency.

Standard deviation is a measure of how dispersed the data is around the mean. The Coefficient of variance is defined as the ratio of the standard deviation to the mean, usually expressed as a percentage. The advantage of using the coefficient of variance is the possibility of comparing the results from the different data sets (observations).

Kurtosis and Skewness are measures of the distribution's shape and are typically used to compare the observed distribution with a normal distribution.

[26] The coefficient of variance or variation coefficient defined as the standard deviation of a data set divided by the average of the same set. It can be interpreted as the relative standard deviation, which makes two different data sets comparable to each other.

Skewness describes the asymmetry of a data distribution and defines the degree to which the data is distributed to the right or to the left of the mean. For a normal distribution, the Skewness is equal to zero because values are evenly distributed around the mean. A negative value of Skewness means that majority of observations are concentrated around the highest values and the distribution is left side asymmetric. A positive value of skewness means that majority of observations are concentrated around the lowest values and the distribution is right side asymmetric.

Kurtosis is a measure of the peakiness of the probability distribution of a data set around the mean. For a normal distribution, the kurtosis is equal to zero (so-called mesokurtic distribution). If the kurtosis is positive (leptokurtic distribution) then observation values are more concentrated compared to a normal distribution. A negative kurtosis value (platykurtic distribution) means that observation values are less concentrated than in a normal distribution. A high kurtosis distribution has a sharper peak and longer, fatter tails, while a low kurtosis distribution has a more rounded peak and shorter thinner tails.

A quartile is one of three points that divide a data set into four equal groups. The 1st quartile is a point on the scale which lies above 25% of the objects' population. The 2nd quartile is a point on the scale which represents the 50% population mark and is equal to the median. The 3rd quartile is the point which cuts off the lowest 75% of observation from the population of objects.

Apart from these descriptive statistics, several graphical tools are available from SAS to visualise the results such as box plots, histograms and scatter plots.

In addition to descriptive statistics, a discriminant analysis can be run in order to check whether each object is a member of the correct class. The main purpose of discriminant analysis is to re-allocate objects to earlier-defined classes.

4.2.7 Interpretation and validation

The basis for the description (and interpretation) of the different classes are the variables that were used in the process of classification.

The goal of class validation is to indicate the characteristic attributes of particular classes, thus emphasizing the differences between the classes. The validation is carried out with variables that were not used in the process of classification to prove the assumption that the classes are homogeneous internally, but heterogeneous to other classes. In this study, this step was performed on a test sample (Characteristic Zones), one representative out of each class. These zones have the shortest distance to their class centroids.

4.3 Classification with SAS

4.3.1 Objects and parameters selection

The statistical tools from SAS Enterprise Guide 4.2 were used to perform the classifications and analysis in this project.

The sets of objects that were classified were the DE Zones generated from the clustering of MLSOAs in Great Britain. From around 8500 MLSOAs, almost 1000 Zones were obtained with the clustering algorithm described in Chapter 2 — Clustering of this report. The next step was to group and simplify the set of zones, by classification, in up to 25 classes. From each class one representative was chosen - the so-called Characteristic Zones.

As the ultimate aim of the macro DE project is to establish the viability of DE zones it was decided that the best test of homogeneity within a class would be to apply the DE design tool developed by the University of Manchester to determine whether the cost of heat production, CO_2 savings and technology selection were sufficiently similar for all zones in a class. The Characteristic Zones, plus 4 additional zones from each class, were therefore to checked for homogeneity within each class with respect to their DE design process.

Every zone is described by 42 parameters, which are relating to (1) heat and electricity demand in the domestic and tertiary sectors, (2) geographical and social aspects and (3) energy demand described by information coming from load curves. In addition, (4) some further coefficients have been calculated based on previously available parameters, e.g. energy density or load factor. Table E-1 in the Appendix E — List of Zone Parameters contains a full description of the parameters and the default units of measure applied.

The methods described for selecting classification parameters in this project were chosen from the Correlation and Principal Factor Analysis tools from Multidimensional Analysis Toolkit of SAS Enterprise Guide 4.2. Its correlation tool allows one to identify all strongly correlated parameters and to do a principal factor analysis on them. This analysis classifies parameters into groups and finds the number of parameters that should be used in the classification procedure. Parameter classification is especially useful when there is a large quantity of available parameters.

In SAS it is possible to represent the correlation between parameters in two ways. One way is a two dimensional matrix of parameters with calculated correlation coefficients for each pair of parameters. The second possible way is a scatter plot for each pair of parameters. It is usually advisable to analyse both the correlation matrix and the graphs. The matrix gives values of correlation coefficients, which are in some cases quite difficult to recognize from the graph. Scatter plots on other hand show the concentration of objects within the sample, which helps to score the discriminatory power of the classification parameters.

Figure 4-1 shows the set of scatter plots automatically generated by SAS for 5 different parameters that characterise DE zones.

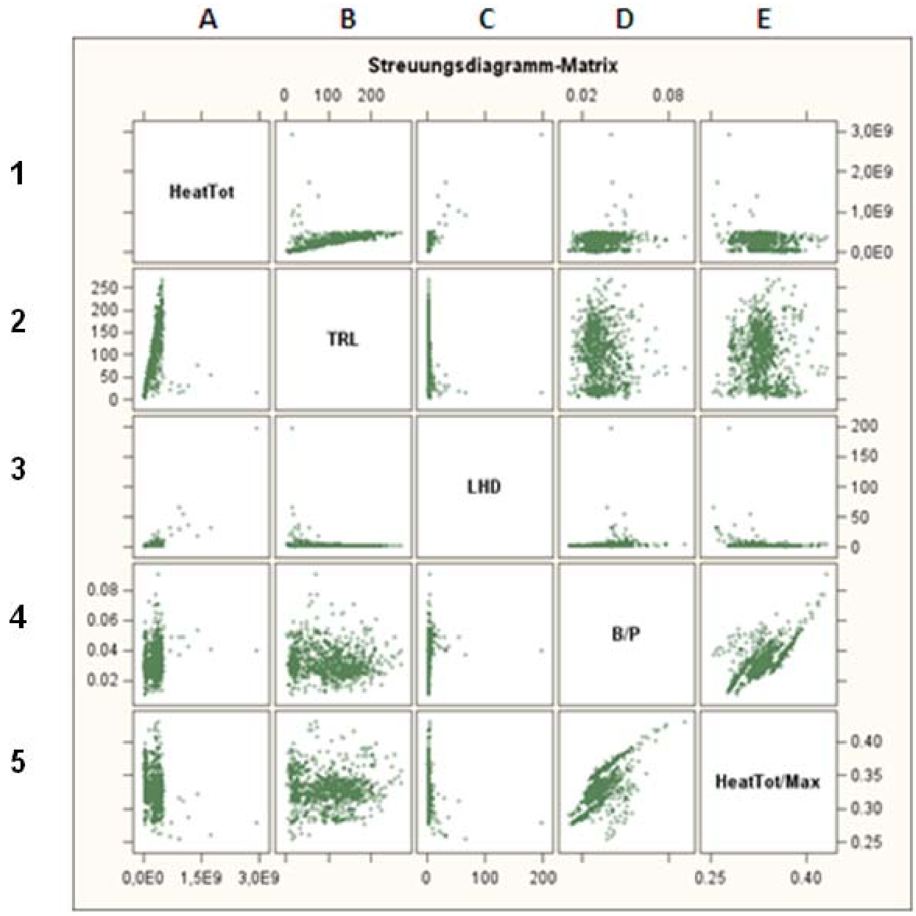


Figure 4-1: Scatter plots of the five parameters total heat demand (HeatTot), total road length (TRL), linear heat density (LHD), ratio of base to peak (B/P) and load factor (HeatTot/Max) indication the correlations

In this example, one can observe that Heat Total is not correlated with B/P (lowest demand hour divided by the highest demand hour) (see A4) nor with Heat Total/Max (Load factor) (see A5) - furthermore Heat Total shows large spreads on both axes. Additionally one can observe the presence of outliers within the population of objects (see C1 – C5).

On the other hand, the Load Factor and B/P are both widely dispersed but nevertheless highly correlated to each other.

Heat Total and LHD (Linear Heat Density) is an example of a pair of parameters that form a well-concentrated group of points. The correlation degree for this pair of parameters is not easy to read from the graph. In this case, the value of their correlation coefficient should be taken from the parameter matrix (see Table 4-1).

NAME	HeatTot	TRL	LHD	B/P	HeatTot/Max
HeatTot	1	0.60	0.60	0.13	-0.23
TRL	0.60	1	-0.14	-0.09	-0.02
LHD	0.60	-0.14	1	0.14	-0.19
B/P	0.13	-0.09	0.14	1	0.66
HeatTot/Max	-0.23	-0.02	-0.19	0.66	1

Table 4-1: Correlation values of the five parameters total heat demand (HeatTot), total road length (TRL), linear heat density (LHD), ratio of base to peak (B/P) and load factor (HeatTot/Max)

Due to the large number of possible parameters, the Factor Analysis tool has been used for determining the right number of classification parameters. Usually, the main idea of this analysis is to replace two or more correlated parameters with one new parameter (principal factor) that provides most of the information of the initial parameters. This factor approximates the regression line between the initial ones.

Therefore, the method can be used in two ways: (1) for reducing the amount of variables (parameters) or (2) for discovering the structure in relationships between variables. In the current study, the second approach was used

A very convenient test is the scree plot. This graph plots the number of factors (new variables) with variance value (Eigen values) for each of them. One should find the point where the smooth decrease of Eigen values appears to level off to the right of the plot. In the example below, the scree test indicates that five is a sufficient amount of coefficients for classification (see Figure 4-2). These five factors cover more than 90% of the information needed to describe the set of objects.

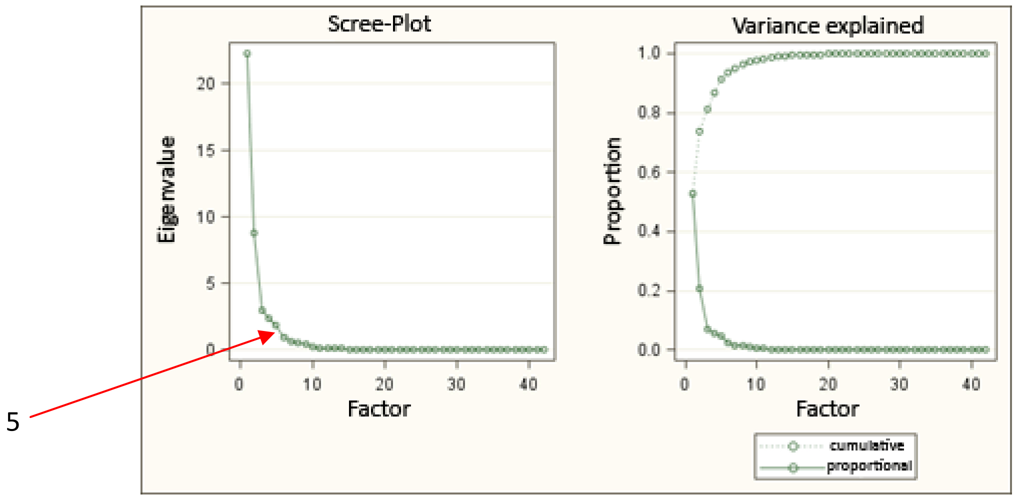


Figure 4-2: Scree plot of the set of DE zones

Finally, the parameters have been mapped on a scatter plot showing their correlations to the two most important factors (see Figure 4-3).

One can see that the first factor covers 53.01% of total variance and the 2nd factor 20.93% . This graph helps to identify which parameters are related to each other by displaying the distances between them.

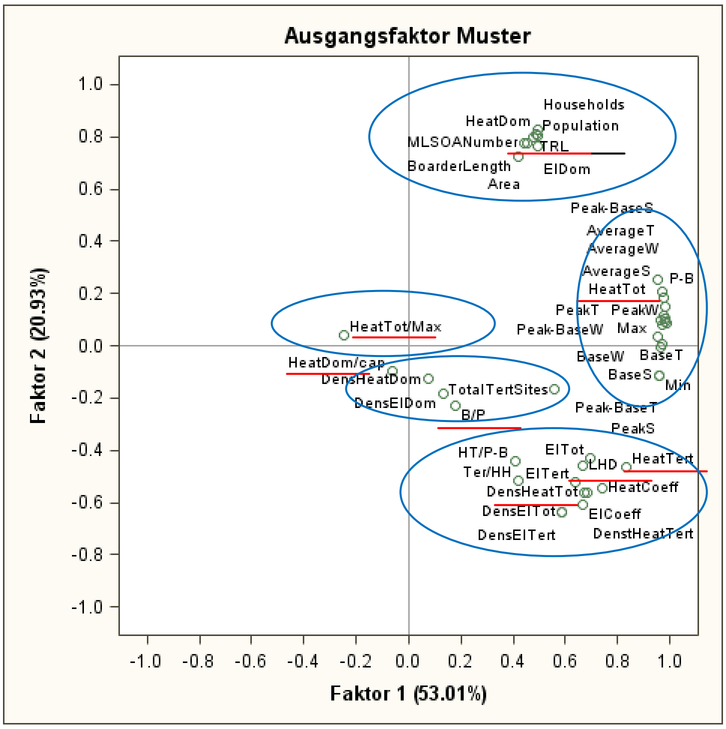


Figure 4-3: Plot of the Factor Analysis with two synthetic parameters. The other paramters are plotted according to their correlations values to these two synthetic ones. The blue circles indicate the five identified groups of parameters out of each one represantative for the further classification has been chosen (indicated by red lines).

The examples above were prepared for all 946 DE Zones.

Next, from each group of parameters the most important representative parameter was selected. Only less correlated parameters are appropriate for classification because they have the ability to discriminate within the set of objects.

After analysis and discussions within the consortium, the following five parameters were chosen for further classification, which for each group have the greatest impact on the benefits of a DE zone:

- heat demand in the tertiary sector: HeatTert (in kWh)
- total heat demand (the sum of domestic and tertiary heat demand): HeatTot (in kWh)
- The density of total heat demand (total heat demand per hectare) – DensHeatTot (in kWh per hectare)
- The ratio of annual base and peak heat demand – B/P (dimensionless or kWh/kWh)
- The Load Factor, calculated as the total heat demand divided by the product of 8760 and the annual peak heat demand - HeatTot/Max (dimensionless or kWh/kWh)

Out of this group of five parameters, the two parameters total heat demand and density of total heat demand have been decided to be most important for the further analysis of the DE zones

4.3.2 Classification with Ward and k-means methods usage

All Zones

For the next steps, the five chosen parameters had to be standardized. As mentioned earlier, standardisation allows the values to be expressed on a dimensionless scale with mean of zero and a standard deviation of one. Thus, it is possible to compare values with different scales and units.

The standardised series of observations were then used for analysis. To determine the number of classes, Ward's Minimum Variance Method was used from whose dendrograms, and CCC[27] graphs or Pseudo-T-Square[28] tests (PST2) the possible number of classes can be inferred.

The example below shows that the CCC has negative values (see Table 4-2). This suggests the presence of outliers within the set of objects.

Development of Classes		
Number of Classes	**CCC**	**Pseudo-T-Square Test**
10	9.27	14.2
9	6.96	85.2
8	3.77	86.8
7	0.43	11.9
6	-4.5	360
5	-7.5	111
4	-8.9	327
3	-16	26.6
2	-12	452
1	0.00	343

Table 4-2: Results of CCC and Pseudo-T-Square tests for different numbers of classes of this set. Two examples have been marked in red and green circles.

[27] CCC: Cubic Clustering Criterion,-see page 6

[28] Pseudo-T-Square test – is a variant of Hotelling's T-squared test and a generalization of the Student's distribution. Usually is employed for multidimensional hypothesis testing.

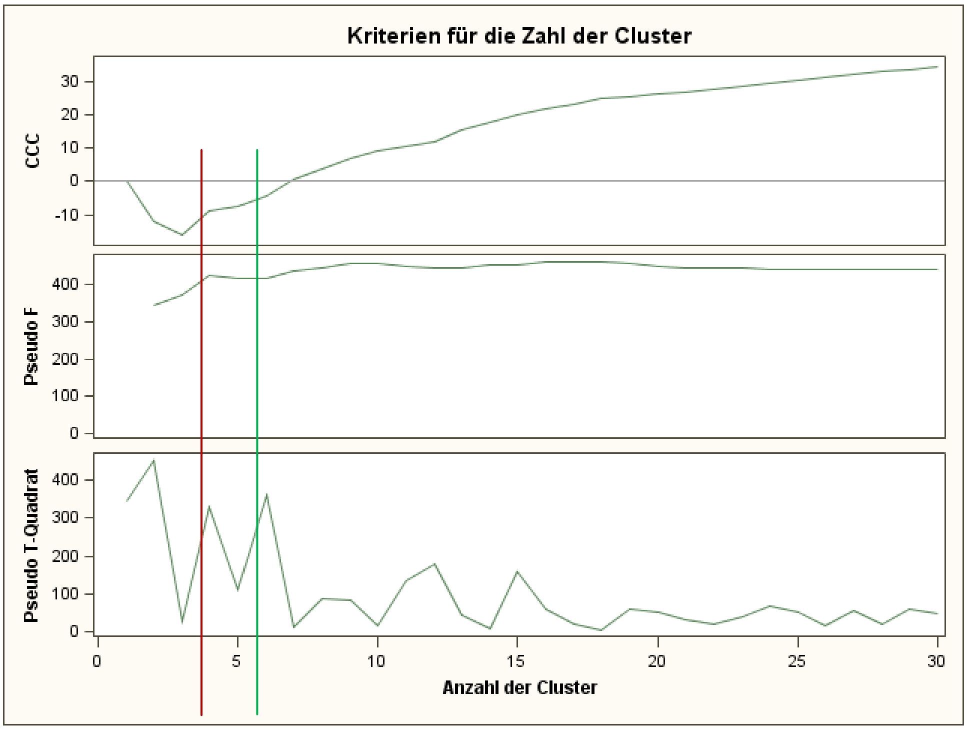


Figure 4-4: Plots of CCC, Pseudo F and Pseudo T-Square tests indicating the number of possible classes. Two examples are marked by the red and green vertical lines.

The possible solutions can be read from Figure 4-4 or from Table 4-2, which contains the values of different statistics. The first values after peaks in the PST2 curve indicate recommended solutions. In the case of the CCC criterion, the peaks themselves indicate the number of classes. Table 4-2 shows significant peaks of the PST2 test for two possible numbers of classes, marked by red and green circles. It is recommended to use more than one criteria, as well as tables and figures, to determine the best number of classes due to possible difficulties of reading the numbers from the graphs only, as is presented in Figure 4-4.

The Outliers problem

In order to perform the classification correctly, isolated elements must be analysed separately. For this purpose, the following steps of analysis were carried out:

1. Using Ward's method, read off a minimum number of classes,
2. Filter out the group of isolated elements and analyse it separately
3. Analyse the main group of elements
4. If a further analysis of the main group still produces the outliers, repeat the procedure.

The following three groups of Zones were found during pre-classification (see Figure 4-5). Group numbers 1 and 3 represent the sets of outliers and the largest one (nb.2) forms a self-consistent group of DE zones with less than 500 GWh of heat demand per year. The scatter plot in Figure 4-5 shows the layout of Zones on the dimensions of Heat total and Load factor.

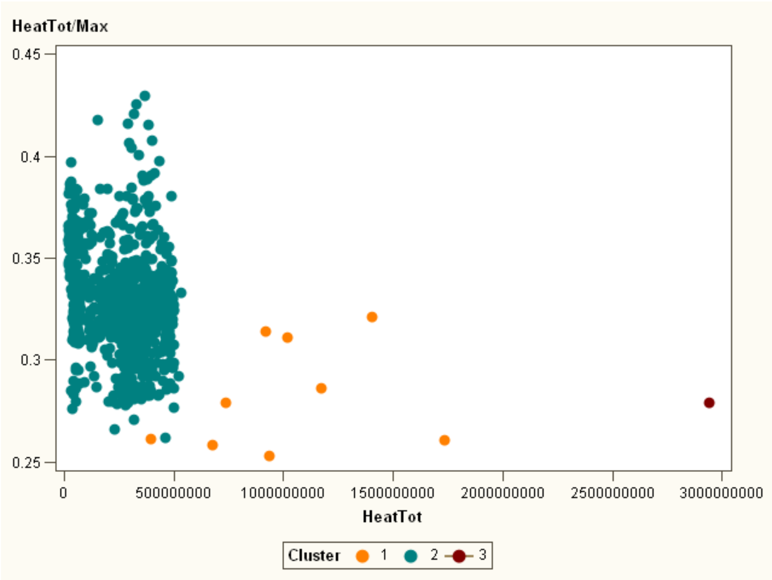


Figure 4-5: Main set of DE zones and outliers on the scales heat total (x-axis) and load factor (y-axis)

Zones below 500GWh

Figure 4-5 shows that the process of eliminating isolated zones has found outliers above 500 GWh of total heat demand, with the exception of one object.

In order to meet the project's requirements, the DE zones with a heat demand larger than 500 GWh were removed from the classification. Eleven zones met this criterion, leaving 937 zones with a heat demand below 500 GWh to be classified. This result is in line with the outlier elimination process.

For this new set the correlation matrix and the principal factor analysis were computed. The graph below (see Figure 4-6) shows the correlation of both the total road length (TRL) and linear heat density (LHD) with the five classification parameters for the remaining DE zones (with total heat demand of less than 500 GWh):

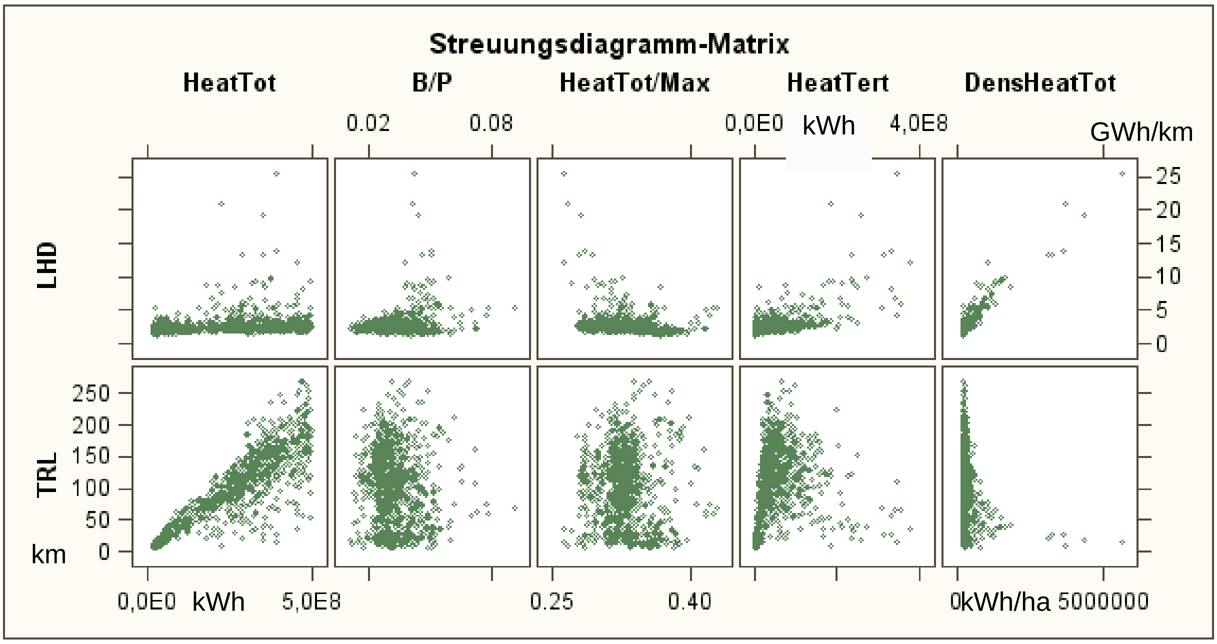


Figure 4-6: Correlation of the five classification parameters total heat demand (kWh), base hour divided by peak hour (dimensionless), load factor (dimensionless), tertiary heat demand (kWh) and heat density (kWh/ha) to the two parameters linear heat density (LHD) (GWh/km) and total road length (TRL) (km)

Correlation between two parameters can be observed in different degrees from -1 (negative correlation/interaction) to 1 (positive correlation/interaction). A correlation value of zero suggests no interaction between two variables.

The graph of correlation above (see Figure 4-6) indicates that total road length (TRL) is highly correlated with total heat demand. However, linear heat density (LHD) is not at all correlated with total heat demand. Furthermore, TRL shows a quite good correlation with tertiary heat demand, in contrast to all other parameters with which TRL forms outliers, especially total heat density.

The set of scatter plots below (see Figure 4-7) shows the dependence between total Heat and five different parameters. Additionally, the table next to the graph (see Table 4-3) includes the value of correlation coefficients corresponding to them. The results indicate the low correlation between total heat demand and total heat density, base hour divided by peak hour and load factor. Additionally, they show good dispersion of objects for total heat demand, base hour divided by peak hour and load factor. In the second graph of Figure 4-7 (Heat Total vs. Density Heat Total) one can observe the narrow range of the sample together with an existing group of outliers.

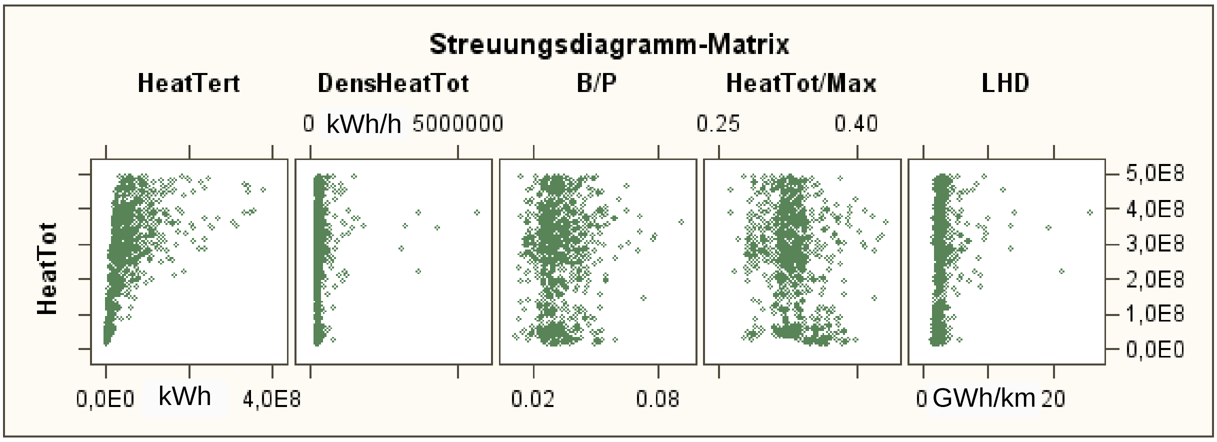


Figure 4-7: Scatter plots showing the correlation of total heat demand with the other 4 classification parameters tertiary heat demand (kWh), total heat density (kWh/ha), base hour divided by peak hour (dimensionless), load factor (HeatTot/Max) (dimensionless) and additionally linear heat density (GWh/km).

NAME	HeatTert	DensHeatTot	B/P	HeatTot/Max	LHD	TRL
HeatTot	0.54	0.12	0.08	-0.18	0.16	0.87

Table 4-3: Correlation coefficients of total heat demand with the six parameters tertiary heat demand, heat density, base hour divided by peak hour, load factor (HeatTot/Max), linear heat density (LHD) and total road length (TRL)

The Principal Factor analysis (see Figure 4-8) once again suggests that five parameters are sufficient for effective classification (see red line on Figure 4-8).

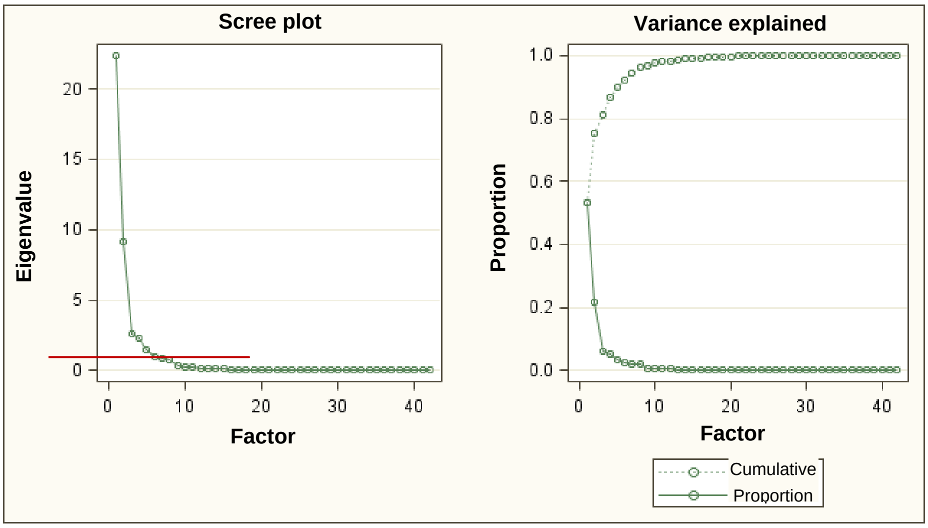


Figure 4-8: Principal factor analysis with a scree plot showing the Eigenvalues of the parameters (left) and the increase of knowledge about the data set with each additional parameter (right)

The results of the principal factor analysis are presented as a two dimensional scatter plot using the correlation of two synthetic parameters as axes. This way it is possible to compare the position of a particular parameter with all other parameters (see Figure 4-9):

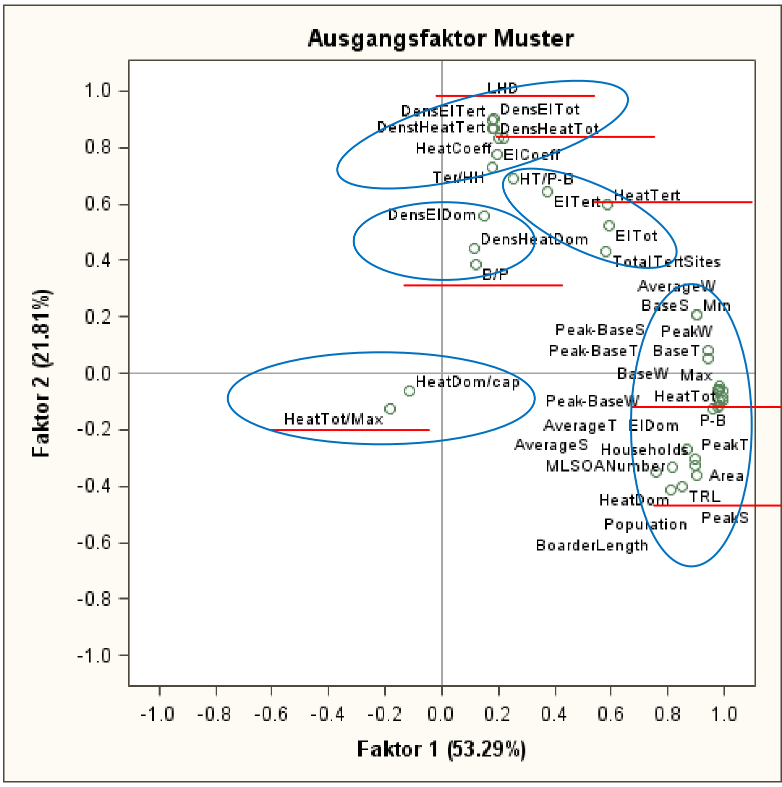


Figure 4-9: Scatter plot of the principal factor analysis using two synthetic parameters

From the above plot (see Figure 4-9), one can read that linear heat density (LHD) and total heat density (Density Heat Total) belong to the same group of parameters (see two red lines at the top of Figure 4-9). The blue circles indicate the five groups of parameters. This process has been described earlier. The values of the correlation coefficient can be seen in the table below (see Table 4-4). Different colours represent different level of correlation, according to the legend underneath the table.

NAME	Heat Tertiary (kWh)	Heat Total (kWh)	Density Heat Total (kWh/ha)	TRL (km)	LHD (GWh/km)	B/P	Load Factor
HeatDom (kWh)	0.19	0.93	-0.10	0.92	-0.08	-0.14	-0.19
HeatTert (kWh)	**1**	0.54	0.54	0.23	0.61	0.51	-0.04
HeatTot (kWh)	0.54	**1**	0.12	0.87	0.16	0.08	-0.18
ElDom (kWh)	0.21	0.87	-0.02	0.83	-0.01	-0.03	-0.19
ElTert (kWh)	0.61	0.30	0.41	0.09	0.44	0.24	-0.15
ElTot (kWh)	0.63	0.52	0.38	0.31	0.40	0.21	-0.19
DensHeatDom	0.12	0.09	0.63	-0.20	0.60	0.14	-0.21
DenstHeatTert	0.59	0.11	0.96	-0.15	0.88	0.29	-0.17
DensHeatTot	0.54	0.12	**1**	-0.19	0.93	0.28	-0.20
DensElDom	0.24	0.10	0.72	-0.19	0.66	0.24	-0.19
DensElTert	0.52	0.10	0.85	-0.13	0.75	0.23	-0.18
DensElTot	0.52	0.10	0.87	-0.14	0.77	0.24	-0.19
Population	0.21	0.91	-0.07	0.87	-0.05	-0.09	-0.21
Area (ha)	0.23	0.83	-0.25	0.94	-0.21	-0.13	-0.09
Households	0.22	0.91	-0.05	0.86	-0.03	-0.01	-0.13
MLSOANumber	0.16	0.82	-0.08	0.84	-0.08	0.05	-0.01
BoarderLength	0.22	0.76	-0.18	0.82	-0.15	-0.10	-0.13
TotalTertSites	0.55	0.53	0.40	0.29	0.38	0.15	-0.29
TRL (km)	0.23	0.87	-0.19	**1**	-0.22	-0.08	-0.05
LHD (GWh/km)	0.61	0.16	0.93	-0.22	**1**	0.30	-0.26
Max (kWh)	0.53	0.99	0.15	0.84	0.20	-0.01	-0.31
Min (kWh)	0.78	0.88	0.32	0.65	0.37	0.48	0.05
B/P	0.51	0.08	0.28	-0.08	0.30	**1**	0.70
HeatTot/Max	-0.04	-0.18	-0.20	-0.05	-0.26	0.70	**1**
HeatCoeff	0.75	0.14	0.76	-0.13	0.75	0.40	-0.08
ElCoeff	0.58	0.12	0.53	-0.09	0.54	0.25	-0.13
HeatDom/cap	-0.11	-0.08	-0.08	-0.04	-0.03	-0.09	0.11
Ter/HH	0.48	0.11	0.57	-0.08	0.46	0.19	-0.18
PeakW (kWh)	0.53	0.99	0.15	0.84	0.20	-0.01	-0.31
BaseW (kWh)	0.70	0.94	0.19	0.79	0.23	0.35	0.05
Peak-BaseW	0.51	0.98	0.14	0.83	0.20	-0.05	-0.34
AverageW (kWh)	0.55	1.00	0.14	0.85	0.19	0.05	-0.23
PeakT (kWh)	0.54	0.99	0.16	0.84	0.21	0.00	-0.29
BaseT (kWh)	0.71	0.94	0.22	0.76	0.26	0.35	0.03
Peak-BaseT	0.52	0.98	0.15	0.83	0.20	-0.03	-0.32
AverageT (kWh)	0.54	1.00	0.10	0.88	0.15	0.08	-0.16
PeakS (kWh)	0.55	0.99	0.10	0.89	0.14	0.11	-0.11

BaseS (kWh)	0.78	0.88	0.32	0.65	0.37	0.48	0.05
Peak-BaseS	0.53	0.98	0.08	0.90	0.12	0.07	-0.12
AverageS (kWh)	0.54	0.99	0.08	0.89	0.11	0.16	-0.06
P-B (kWh)	0.51	0.97	0.08	0.89	0.12	0.02	-0.15
HT/P-B (kWh)	0.55	0.18	0.42	-0.04	0.46	0.26	-0.15

(pink)	very high correlated	80%-100%
(yellow)	high correlated	60%-79%
(green)	medium correlated	40%-59%
(blue)	low correlated	20%-39%
(white)	not correlated	0%-19%

Table 4-4: Correlation Matrix for several parameters. Colours indicate the strength of correlation.

Classification of DE zones below 500 GWh heat demand per year

The classification has been performed using the k-means model[29]. By default, the k-Means method uses Euclidean distances for distances between class centroids, so the class centres are based on a least squares estimation. A centroid is a point of one class whose coordinate in each dimension is the mean of all observations of the class. Every iteration of the k-means algorithm reduces the test of 'least squares criterion' until convergence is achieved.

Before the k-means algorithm can be used, the correct number of classes needs to be obtained. This has been done using the Ward analysis. Figure 4-10present the results from this method.

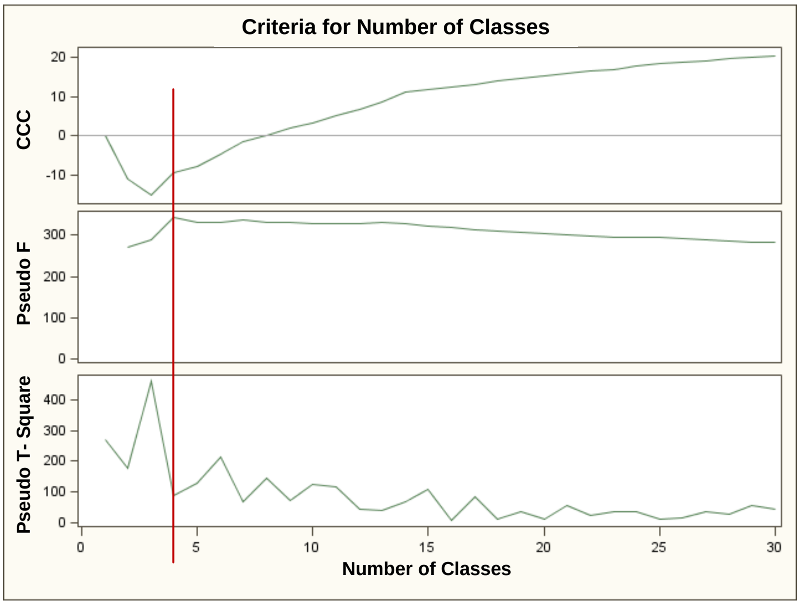


Figure 4-10: Graphs for CCC, Pseudo F, Pseudo t-squares

[29] For more information on the k-means algorithm, please see ...

The negative values of the CCC coefficient indicate that some outlier elements still exist.

In order to separate these existing outliers from the rest of the zones, the lowest possible number of classes has been chosen. In this case, the number 4 seemed to be a suitable number of classes and as a next step, the k-mean analysis was run with four classes as a target number.

The result was verified using the coefficient of variation. The numbers of elements for each class are 456, 208, 267 and 6 (see Table 4-5).

CLUSTER	_FREQ_	HeatTert_CV	HeatTot_CV	DensHeatTot_CV	HeatDom/cap_CV	B/P_CV	HeatTot/Max_CV
1	456	56,96	28,89	30,59	9,76	19,13	5,08
2	208	59,38	23,12	74,09	12,69	21,32	7,77
3	267	113,08	72,70	30,24	26,95	22,11	5,72
4	6	22,51	19,55	24,00	13,25	8,83	4,18

Table 4-5: Results of first classification step into four groups.

The smallest group in this set is class 4, with only 6 elements. The worst coefficient of variation for total heat demand can be found in group 3. The box and whiskers plot illustrated this result and show the outliers that exist within group 1 and 3 (see Figure 4-11). The lower edges of these boxes represent the lower quartiles (25^{th} percentiles) and upper edges represent the upper quartiles (75^{th} percentiles). Additionally, the median (50^{th} percentile) is represented by a central line that divides the box into two sections, the mean by diamond and the extreme values (minimum and maximum) by two whiskers. The box and whiskers plot summarized the distribution of set of the data, takes into account existing outliers as well (the empty circles are marked with red ellipses).

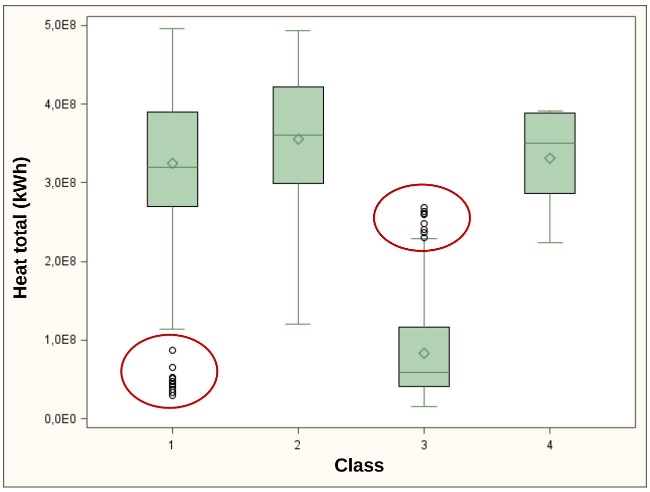


Figure 4-11: Box plot for the four classes of the first step of classification on the parameter total heat demand

The two graphs below (see Figure 4-12 and Figure 4-13) show the scatter plots of the four groups of DE zones of the first classification step. The x-axis in both graphs shows total heat demand; Figure 4-12 shows the total heat density and Figure 4-13 shows the Load factor (heat total divided by annual peak).

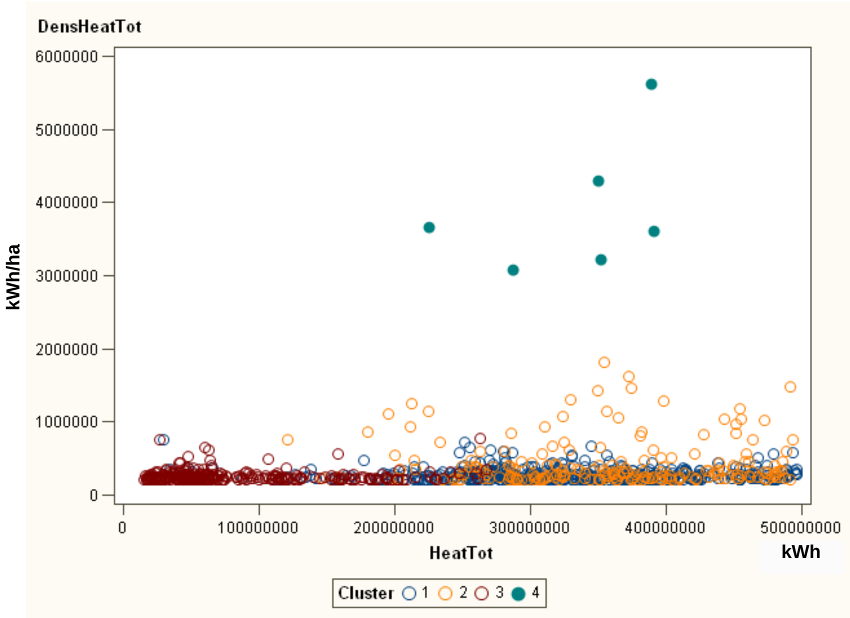


Figure 4-12: Scatter plot of first four groups of DE zones on total heat demand (kWh) and total heat density (kWh/ha)

The results in Figure 4-12 suggest that identifying outliers is possible using the total heat density and total heat demand dimensions. This is difficult for other pair of parameters as presented in Figure 4-13. On the other hand, the parameter pair of total heat demand and load factor shows a greater classification potential due to the larger range of values from the elements.

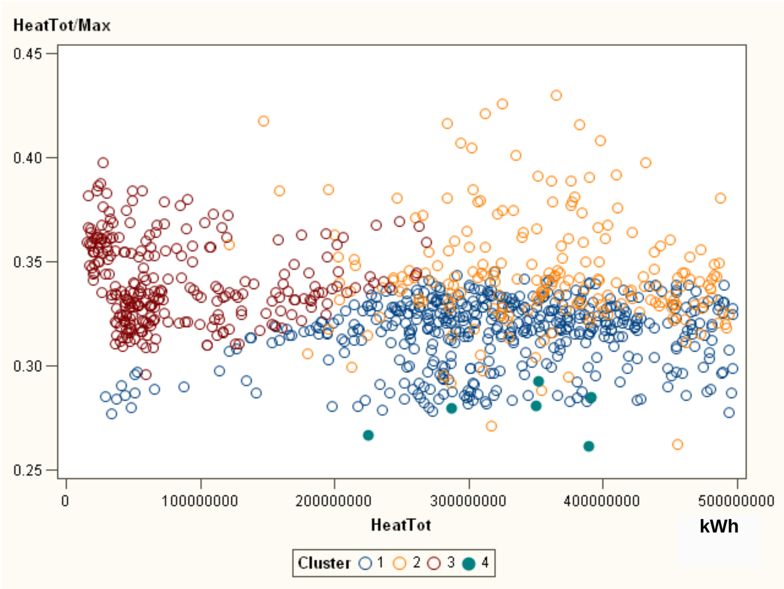


Figure 4-13: Scatter plot of first four groups of DE zones on total heat demand (kWh) and load factor

Before the next step (second step) of classification, class number 4 was separated and kept as the first final class of the classification process. The remainder sets were analysed once more using the Ward method. The result indicates that six classes are a good number of classes in this step.

One can read the number of classes from the dendrogram below (see Figure 4-14):

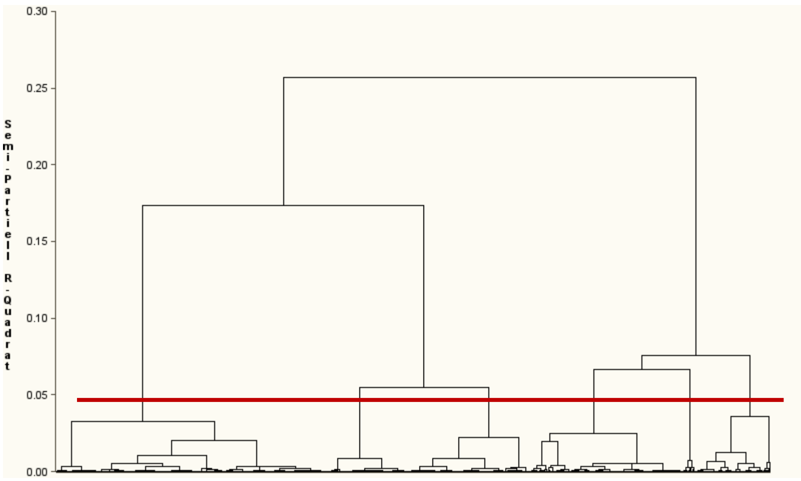


Figure 4-14: Dendogram of second classification step, indicating to classify into six classes at the red line

The vertical branches in the dendrogram indicate the distance between classes. The longer the branches, the greater are the differences between the classes. Dendrograms can only be generated for hierarchical methods of aggregation.

Statistical validation has shown that classes number 5 and 6 of step 2 have high values for the coefficient of variance of total heat demand (see Table 4-6).

Class	Frequency	Variance Coefficients (%)					
		Heat Tertiary	Heat Total	Total Heat Density	Heat Domestic per capita	Base over Peak	Load Factor
1	340	47,30	21,08	29,94	9,72	16,88	5,13
2	38	67,74	23,20	57,74	13,46	19,84	4,92
3	26	30,56	21,33	44,97	17,97	13,72	7,26
4	209	38,39	22,93	50,13	10,19	13,24	3,76
5	210	89,14	62,37	27,53	9,02	16,96	4,23
6	108	129,44	91,20	40,28	40,03	17,26	3,59

Table 4-6: Results of six classes of the second step of classification based on classes 1, 2 and 3 of the first step.

For these reasons, classes number five and six have been separately analyzed in the next step (third step). The best result was obtained by classifying on total heat demand only into four classes. The results are presented in the table below (see Table 4-7).

Class	Frequency	Variance Coefficients (%)				
		Heat Tertiary	Heat Total	Total Heat Density	Base over Peak	Load Factor
1	58	84,21	15,09	31,79	29,07	6,79
2	25	38,07	11,22	43,92	21,16	5,74
3	56	52,92	9,86	25,52	16,44	4,44
4	179	85,27	35,78	32,53	26,56	7,03

Table 4-7: Results of four classes of the third step of classification based on classes 5 and 6 of step two.

For a better presentation of the data, the scatter plot below (see Figure 4-15) shows the total heat demand plotted against Load factor:

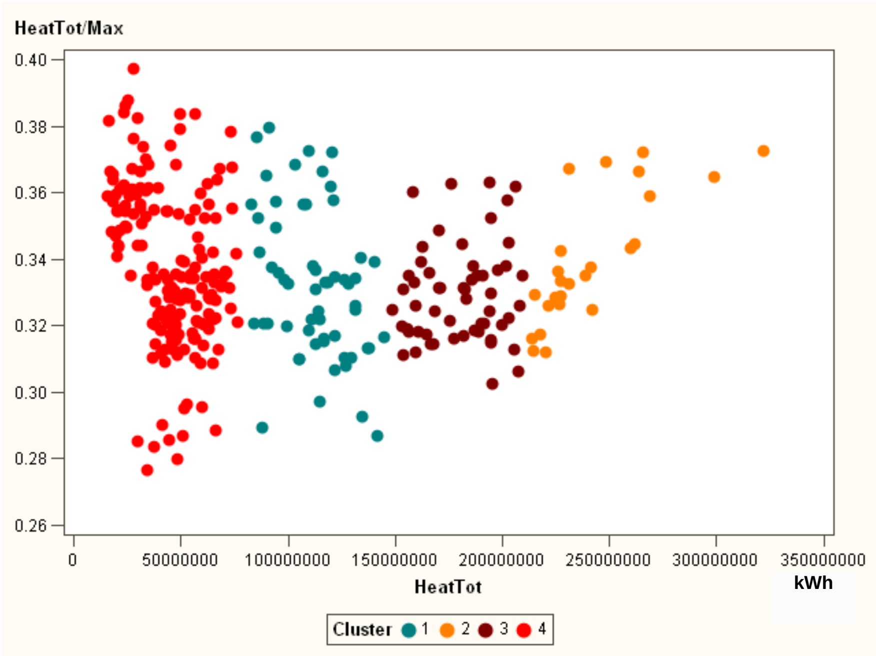


Figure 4-15: Scatter plot of the four classes of the third classification step for total heat demand (kWh) and load factor

The fourth class of the third step was classified once again in two classes by total heat demand. The table below shows the results of this fourth classification step (see Table 4-8).

Class	Frequency	Variance Coefficients (%)					
		Heat Tertiary	Heat Total	Total Heat Density	Heat Domestic per capita	Base over Peak	Load Factor
1	83	80,06	28,56	34,83	11,49	28,26	7,21
2	96	59,32	15,12	30,60	18,61	24,78	6,22

Table 4-8: Results of two classes of the fourth step of classification based on class 4 of step three.

The 1st class from the second step of classification was classified once again (third step). As a result, 4 classes were obtained. The total heat demand was used as a single classification factor (see Table 4-9).

Class	Frequency	Variance Coefficients (%)					
		Heat Tertiary	Heat Total	Total Heat Density	Heat Domestic per capita	Base over Peak	Load Factor
1	99	42,57	5,34	22,84	9,01	14,42	4,30
2	102	48,96	7,46	35,51	9,53	18,78	5,55
3	89	50,26	5,25	27,95	10,04	18,09	5,47
4	50	33,62	4,04	30,16	10,39	15,57	4,96

Table 4-9: Results of four classes of the third step of classification, based on class one of step two.

The fourth class from the second step was classified by total heat demand (third step). As a result three new classes were created (see Table 4-10).

Class	Frequency	Variance Coefficients (%)					
		Heat Tertiary	Heat Total	Total Heat Density	Heat Domestic per capita	Base over Peak	Load Factor
1	58	37,09	12,65	57,48	11,35	12,45	4,36
2	70	37,29	6,39	40,04	10,28	13,25	3,49
3	81	36,43	7,75	47,34	9,27	12,56	3,50

Table 4-10: Results of three classes of the third step of classification, based on classes four of step two.

As a result of all classification steps, 15 classes were created in total. Figure 4-16 below shows all classes for total heat demand. In order to represents the zones class by class in a plot diagram using only one classification parameter, the zones were sorted by classes and then indexed (ID in y-axis of Figure 4-16).

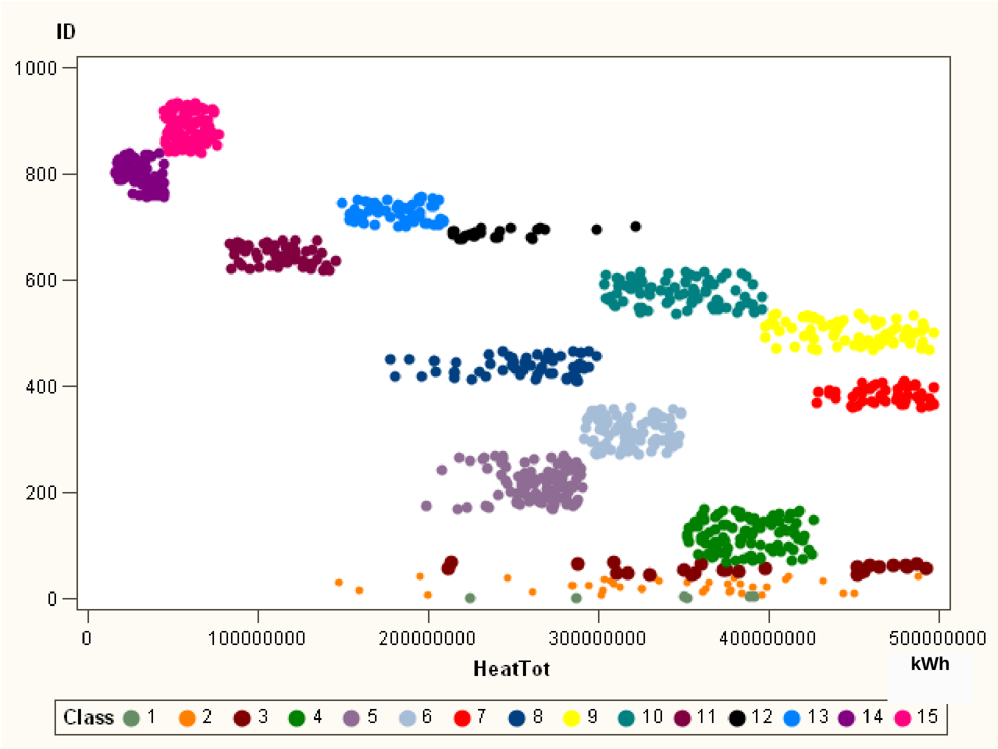


Figure 4-16: Distribution of total heat demand (kWh) within each class (DE zones have been sorted beforehand)

4.4 Classification results statistical evaluation and interpretation

As mentioned earlier, the classification process was supposed to create no more than 25 classes of DE zones. It was also hoped that the classification process would reduce the variance coefficients for total heat demand down to 10% or lower within each class. Later on, this value has been raised to 20 % and in return broadened to the further classification parameters except tertiary heat demand.

The quality of classification was verified in a first step using several statistical tools. With the help of the variance coefficients, the level of group consistency and similarity of elements within each class was tested. In order to check the difference between classes, the mean analysis was carried out. Classes with the same means in all dimensions (all parameters) should have been treated as the same classes. The distribution of elements within the class was measured using kurtosis and the class' symmetry using skewness. If a class possesses a high kurtosis value then objects of the set are concentrated around the mean.

The summary result of variance coefficients for the 15 created classes is presented below in Table 4-11.

Class	# zones in class	HeatTert (%)	HeatTot (%)	DensHeat Tot (%)	B/P (%)	HeatTot/ Max (%)	LHD (%)
1	6	22.51	19.55	24.00	8.83	4.18	28.26
2	38	67.74	23.20	57.74	19.84	4.92	48.21
3	26	30.56	21.33	44.97	13.72	7.26	34.97
4	99	42.57	5.34	22.84	14.42	4.30	13.90
5	102	48.96	7.46	35.51	18.78	5.55	21.46
6	89	50.26	5.25	27.95	18.09	5.47	18.66
7	50	33.62	4.04	30.16	15.57	4.96	19.99
8	58	37.09	12.65	57.48	12.45	4.36	43.01
9	70	37.29	6.39	40.04	13.25	3.49	26.68
10	81	36.43	7.75	47.34	12.56	3.50	26.96
11	58	84.21	15.09	31.79	29.07	6.79	25.32
12	25	38.07	11.22	43.92	21.16	5.74	50.09
13	56	52.92	9.86	25.52	16.44	4.44	24.91
14	83	80.06	28.56	34.83	28.26	7.21	31.83
15	96	59.32	15.12	30.60	24.78	6.22	21.78

Table 4-11: Summary statistics of the classification process for all 15 classes

In the fifteen created classes seven show a variance coefficient for total heat demand below 10%, five between 10% and 20% and three above 20%. The highest value is 28.5% for group 14. Attempts to improve the results by additional classification of the groups with high coefficients of variance (group 1, 2, 3 and 14) did not give satisfactory results.

The results from Table 4-11 above are presented in Figure 4-17 below. Variance coefficients for load factor is on the same low level for all fifteen classes, while for tertiary heat demand the variance coefficients have relatively high and variable values.

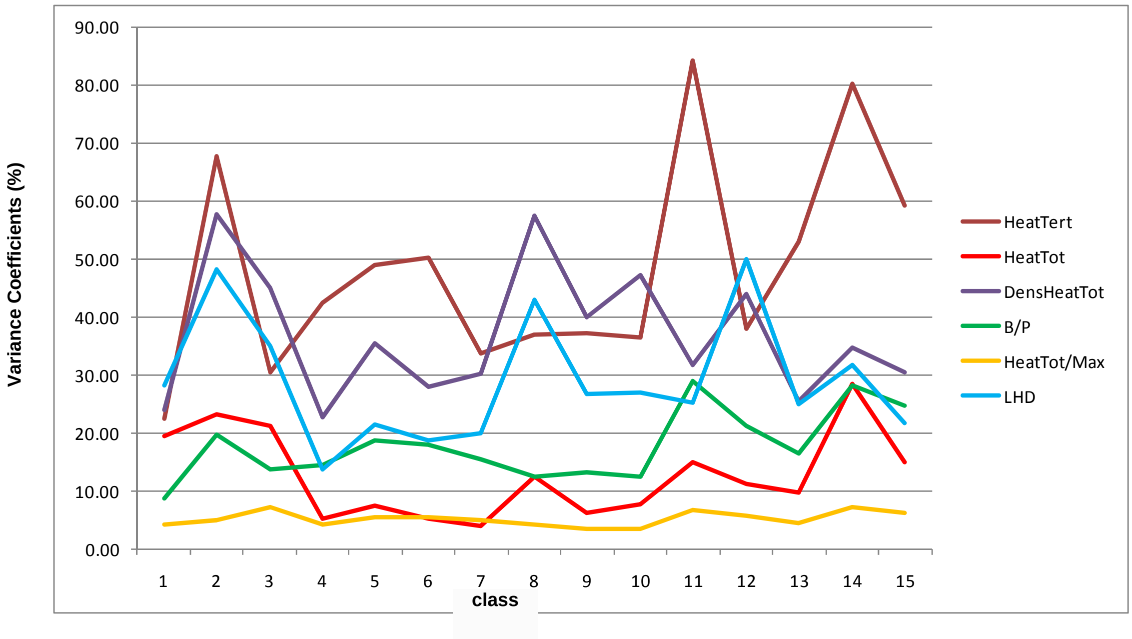


Figure 4-17: Values of variance coefficients for the five classification parameters and linear heat demand (LHD) for all 15 classes

In addition, a mean analysis was performed for all classes. Table 4-12 below contains the mean values for the main parameters used in classification.

Class	Frequency	HeatTert (kWh)	HeatTot (kWh)	DensHeat Tot (kWh/ha)	B/P	HeatTot/ Max	LHD (kWh/km)
1	6	2.78E+08	3.32E+08	3 924 596	0.05	0.28	17.75
2	38	8.52E+07	3.36E+08	352 249	0.06	0.39	2.89
3	26	2.30E+08	3.87E+08	956 288	0.05	0.32	6.91
4	99	3.95E+07	3.86E+08	268 894	0.03	0.32	2.52
5	102	3.13E+07	2.64E+08	286 796	0.03	0.31	2.69
6	89	3.63E+07	3.18E+08	263 450	0.03	0.31	2.50
7	50	5.13E+07	4.67E+08	292 394	0.03	0.31	2.68
8	58	7.57E+07	2.54E+08	363 189	0.04	0.33	3.27
9	70	9.89E+07	4.49E+08	296 340	0.04	0.33	2.82
10	81	8.67E+07	3.47E+08	303 758	0.04	0.34	2.80
11	58	1.48E+07	1.12E+08	255 790	0.03	0.33	2.44
12	25	2.04E+07	2.41E+08	268 823	0.03	0.34	2.65
13	56	2.36E+07	1.80E+08	239 639	0.03	0.33	2.48
14	83	2.10E+06	3.02E+07	263 403	0.03	0.35	2.22
15	96	6.82E+06	5.72E+07	264 201	0.03	0.33	2.36

Table 4-12: Mean values of classes for the five classification parameters and linear heat demand (LHD)

Figure 4-18 shows that some groups have similar means for total heat demand - for example, groups 1, 2, 6 and 10, groups 3 and 4, groups 5, 8 and 12 and groups 7 and 9, as shown on the table below (see Table 4-13).

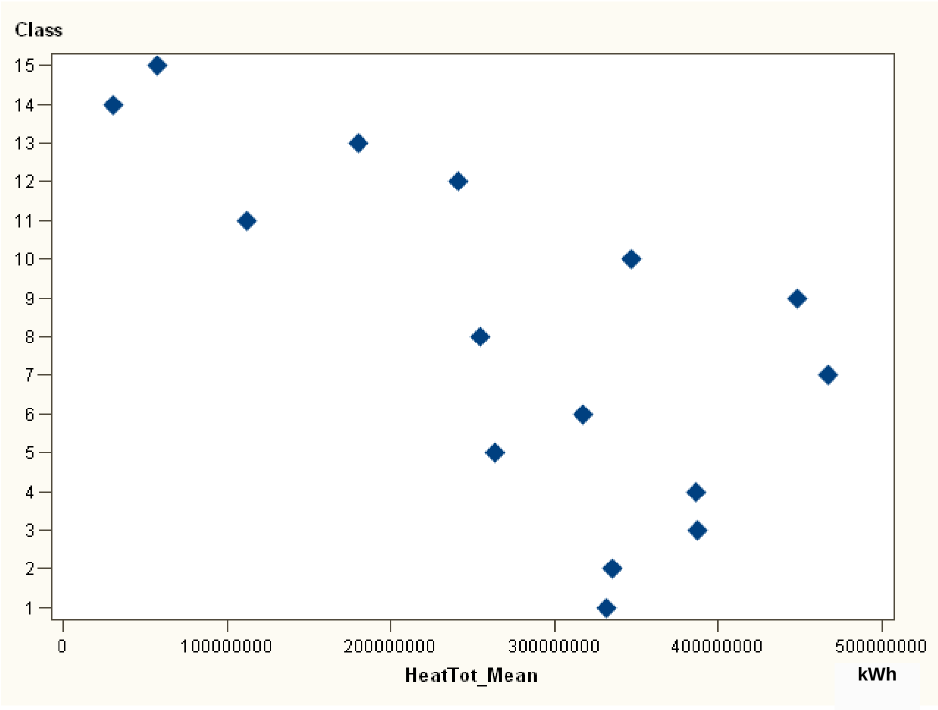


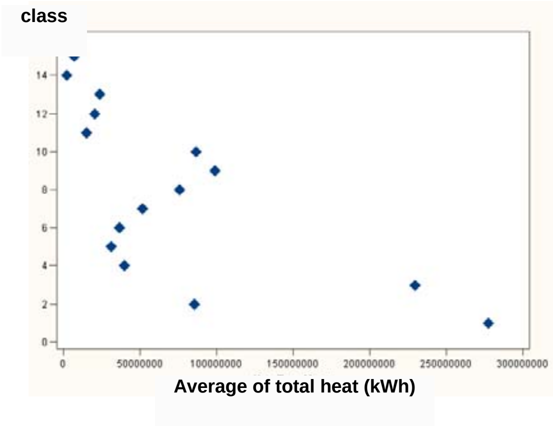
Figure 4-18: Mean values for total heat demand of all 15 classes

Group	Mean (kWh)	Group	Mean (kWh)	Group	Mean (kWh)	Group	Mean (kWh)
1	3.32E+08	3	3.87E+08	5	2.64E+08	7	4.67E+08
2	3.36E+08	4	3.86E+08	8	2.54E+08	9	4.49E+08
6	3.18E+08			12	2.41E+08		
10	3.47E+08						

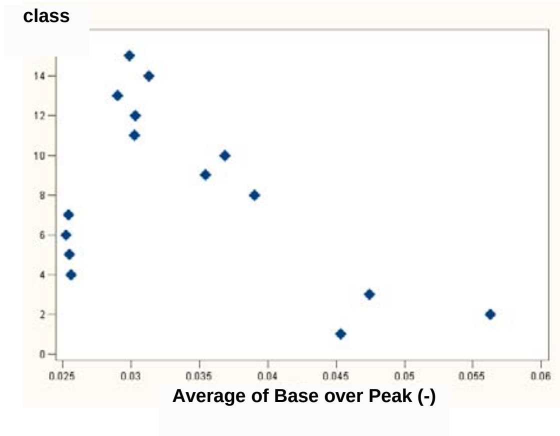
Table 4-13: Grouping of classes by similar means of total heat demand

Groups 11, 13, 14 and 15 have significant differently means: 111,908,997.07 kWh, 179,956,609.50 kWh, 30,227,384.24 kWh and 57,155,286.01 kWh respectively.

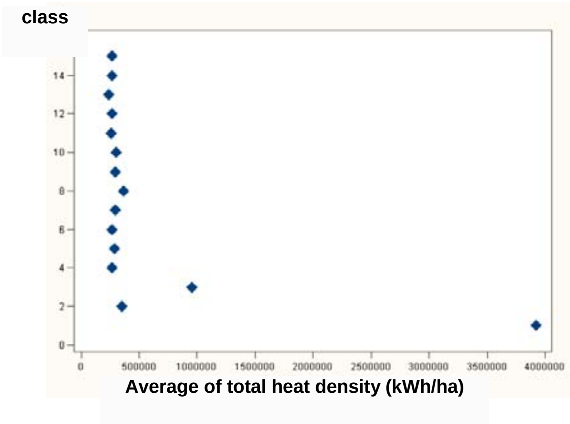
Further analysis of means for the remaining 4 classification parameters shows that classes with similar means for total heat demand differ from each other for other parameters. The scatter plots below (see Figure 4-19) present the means for particular groups for Heat Tertiary, Density Heat Total, Load Factor, and ratio of annual base to annual peak (B/P).



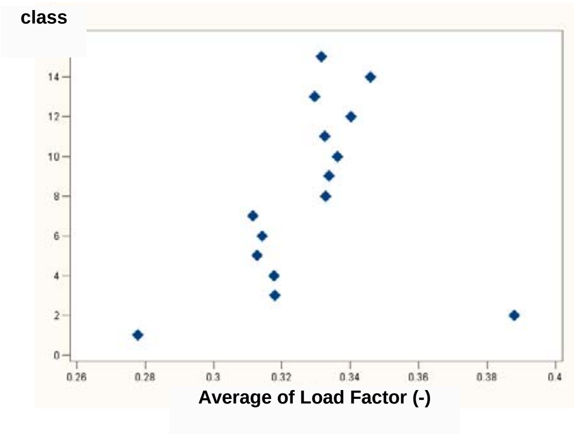
(a) Mean values of tertiary heat demand of all 15 classes



(b) Mean values for ratio of base over peak of all 15 classes



(c) Mean values of heat density of all 15 classes



(d) Mean values of the load factor of all 15 classes

Figure 4-19: Mean values of the four further classification parameters of all 15 classes

Comparing classes one, two, six and ten, which have very similar means for Heat Total one can observe that for Heat Tertiary, B/P and Load Factor their means differ significantly. In the case of classes three and four the difference can be observed for Heat Tertiary, B/P and Density Heat Total.

4.4.1 Distribution of zone within classes

Table 4-14 below shows the Kurtosis and Skewness coefficients calculated for all 15 classes:

Class number	Name	HeatTert	HeatTot	DensHeat	B/P	HeatTot/Max	LHD
1	Kurtosis	-1.370	0.102	2.056	-1.890	-1.082	-1.046
	Skewness	-0.423	-1.025	1.460	0.484	-0.393	0.683
2	Kurtosis	4.346	0.333	5.928	1.204	-0.386	-0.355
	Skewness	1.714	-0.607	2.361	1.213	0.564	1.105
3	Kurtosis	-0.251	-0.488	-0.786	0.511	0.129	-0.648
	Skewness	0.652	-0.545	0.044	0.546	-0.684	0.110
4	Kurtosis	0.159	-1.111	3.155	1.444	0.641	1.230
	Skewness	0.821	0.164	1.514	-0.906	-0.967	1.045
5	Kurtosis	1.719	0.760	5.398	0.167	-0.988	2.897
	Skewness	1.065	-1.009	2.230	-0.747	-0.533	1.549
6	Kurtosis	4.121	-1.200	3.638	-0.304	-0.889	3.548
	Skewness	1.730	0.221	1.874	-0.347	-0.513	1.702
7	Kurtosis	0.272	-0.790	3.044	1.126	-0.810	0.172
	Skewness	0.697	-0.260	1.724	-0.781	-0.310	0.860
8	Kurtosis	0.871	-0.331	5.382	-0.604	1.121	5.478
	Skewness	0.737	-0.741	2.277	0.094	-0.980	2.321
9	Kurtosis	-0.129	-1.124	8.184	0.878	1.109	3.381
	Skewness	0.532	-0.125	2.682	0.910	0.527	1.522
10	Kurtosis	0.637	-1.124	11.152	-0.584	-0.044	4.159
	Skewness	0.588	0.071	3.023	0.565	0.430	1.795
11	Kurtosis	5.644	-0.945	25.006	-0.148	-0.511	1.859
	Skewness	2.325	-0.062	4.520	0.401	0.268	1.052
12	Kurtosis	-1.017	2.282	15.298	-0.861	-1.038	15.587
	Skewness	0.409	1.513	3.631	0.716	0.441	3.739
13	Kurtosis	0.140	-1.302	16.292	1.692	0.039	8.738
	Skewness	0.828	-0.081	3.744	0.607	0.714	2.463
14	Kurtosis	0.128	-1.239	19.113	1.179	0.481	4.254
	Skewness	0.915	0.010	4.017	0.675	-0.719	1.618
15	Kurtosis	0.932	-0.928	8.652	1.016	0.624	0.390
	Skewness	1.055	0.308	2.687	0.742	0.447	0.562

Table 4-14: Kurtosis and Skewness values of various parameters for all 15 classes

Negative values of Kurtosis suggest that objects are evenly distributed within the class and the distribution is flat. A high value suggests that elements in the class are concentrated around the mean.

Negative values of Skewness indicate a left-side asymmetry of the class. Skewness values close to zero mean that the class has a symmetrical shape.

Analysing Kurtosis and Skewness for Heat Total one can observe that classes 1, 2, 3, 5, 7, 8, 9, 11 and 13 have negative Skewness, which means left-side asymmetry of the given classes. However, Skewness for class 10, 11 13 and 14 have values close to zero, which indicate quite good similarity with the normal distribution.

Figure 4-20 shows the example of distribution analysis for class 11 and 12. Class 11 represents the classes with negative Skewness and class 12 with positive. In the first case, the distribution has left-side asymmetry. The second one represents right-side asymmetry.

The Kurtosis of class 11 is close to minus one (-1), which is the special case of a flat-shape distribution, so-called semicircle distribution. Class 12 has a high and positive value of kurtosis and is an example of peaked shape distribution with value concentrated around the mean.

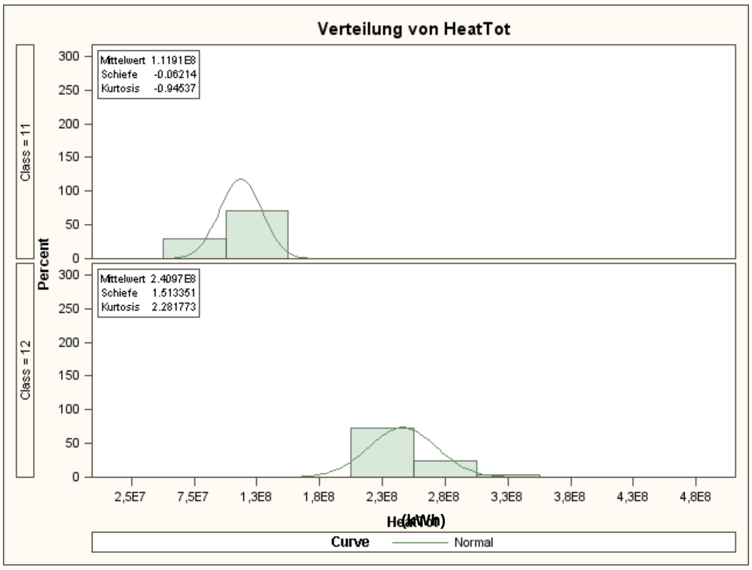


Figure 4-20: Distribution of total heat demand within class 11 (top) and class 12 (bottom)

The positive, close to one or higher value of Kurtosis indicates that the majority of elements within the group are concentrated around the mean and the group forms a compact aggregation. If the Skewness is also near plus or minus one this means that objects are concentrated on group's right or left side. A good example of this case is group no. 12 that has objects concentrated on the left side and with two detached elements on the right side (see Figure 4-21).

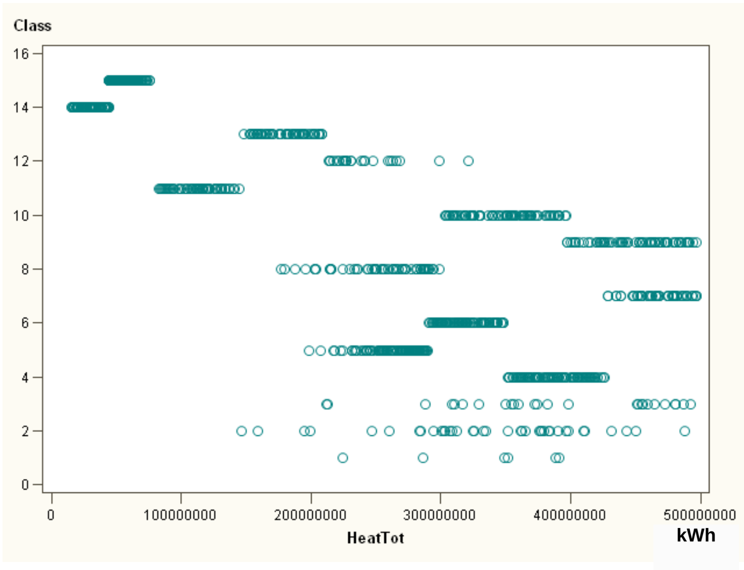


Figure 4-21: Distribution of total heat demand for all 15 classes

The Kurtosis for heat density shows that all classes, except class three, have positive values between 2.056 and 25. This suggests a high concentration of elements around the mean. Class no. 3 has a negative kurtosis of -0.786. The Skewness values for heat density are between 1.46 and 4.5, except once more class 3 with a Skewness close to zero (0.044).

This situation can also be observed in the scatter plot below (see Figure 4-22), which shows the distributions of heat density of all 15 classes. With the exception of class no 3, all other classes are characterized by right side Skewness which can be seen from the tails on the right side of the distributions.

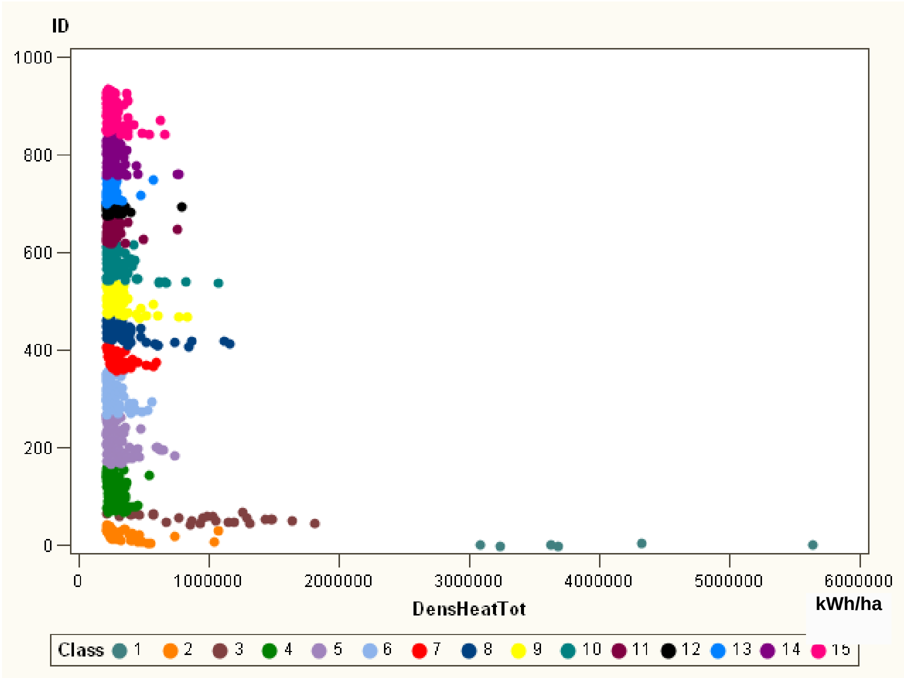


Figure 4-22: Scatter plot for heat density of all DE zones, sorted by classes

The concentration of classes was also analyzed using the ratio of the minimum and maximum values of parameters within the class, expressed as a percentage. From the graph below (see Figure 4-23) one can read that the best concentration of elements was obtained for load factor and total heat demand. The worst result is for tertiary heat demand. A high percentage means that minimum and maximum values are close to each other. In the case where the minimum is equal to the maximum, the ratio has a value of 100%.

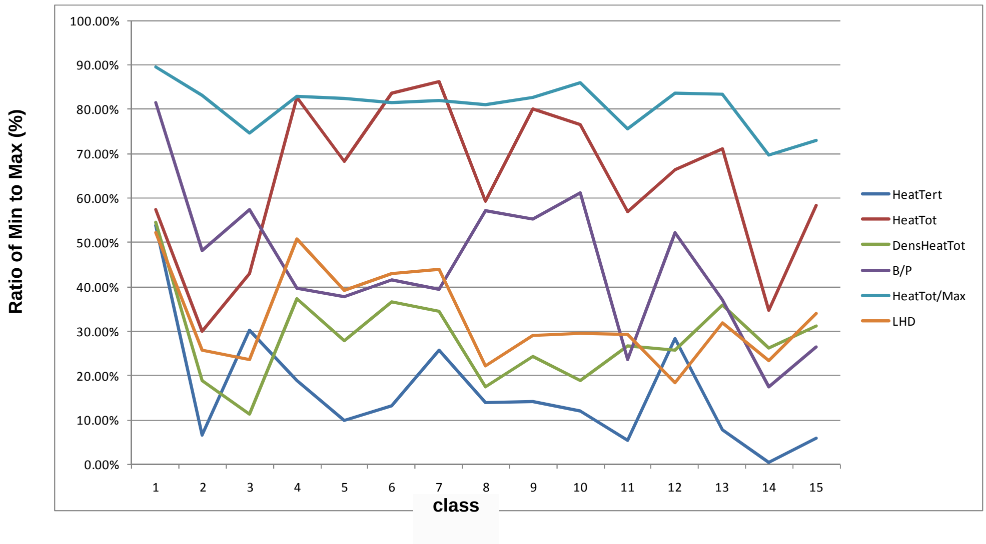


Figure 4-23: Ratio of minimum to maximum value of all 15 classes for the five classification parameters and linear heat density (LHD)

Another possible way to score the class concentration is with quartiles. Figure 4-24 shows the first, second (mean) and third quartiles for Heat Total. One can observe that for classes 1, 2 and 3 a significant difference between the quartiles' values exists.

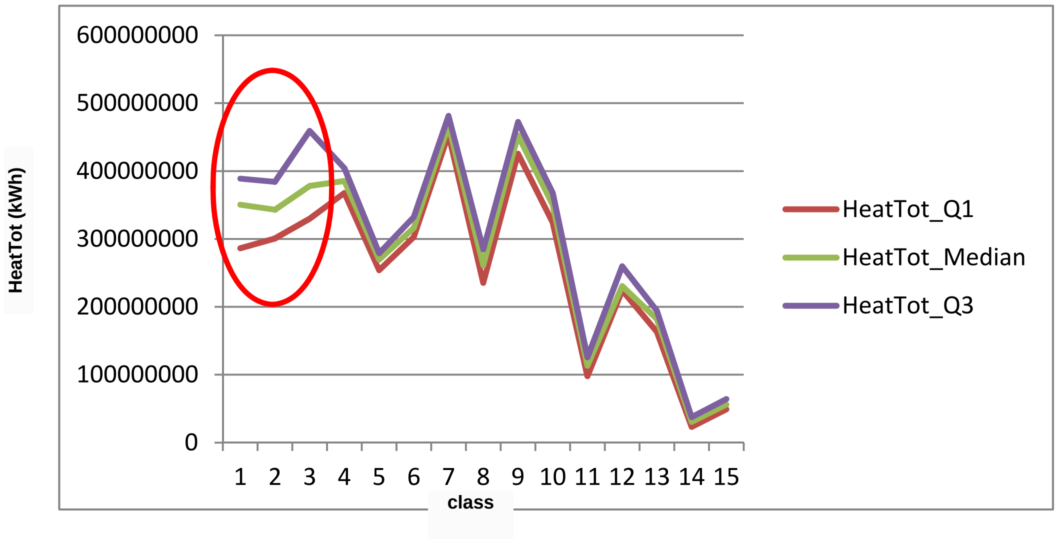


Figure 4-24: Comparison of the three zones representing quartile 1, median and quartile 3 of all 15 classes for total heat demand. The red oval emphasizes the greater difference in the first three classes

The quartiles analysis for Density Heat Total shows that the classes are fairly compact for this parameter, except for the first three classes (see red oval in Figure 4-24).

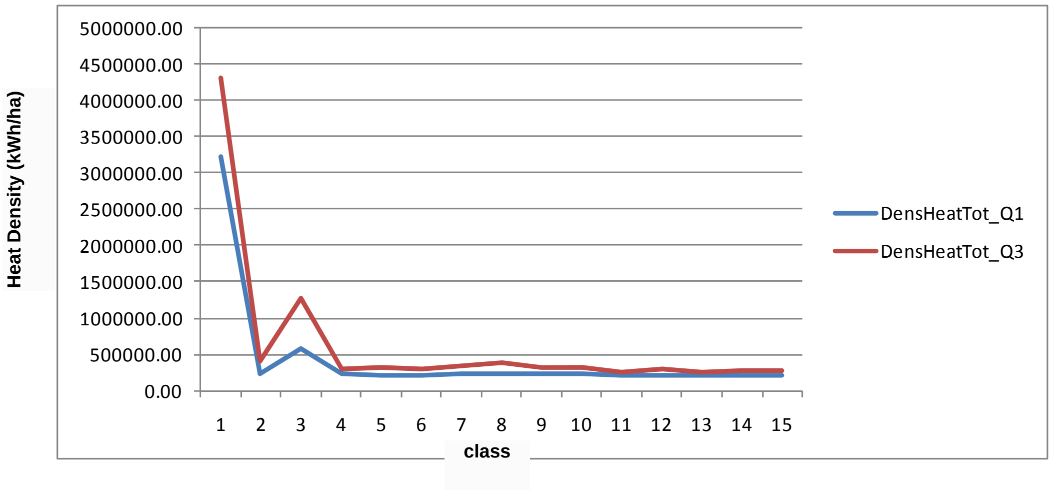


Figure 4-25: Comparison of the two zones representing quartile 1 and quartile 3 of all 15 classes for heat density

Quartiles can be also used for scoring the symmetry of class. If the difference between second and first quartiles is less than the difference between third and second, then the class is left side asymmetric.

From the statistical analysis results, classes no. 1, 2 and 3 show lower concentrations and for them less similarity can be expected between their members than for the other classes.

4.4.2 The choice of Characteristic Zones

The Characteristic Zones were chosen as those which are located closest to the centroid of their classes. The distances between each zone and their class centroid are available in a resulting sheet, generated by SAS. The chosen Characteristic Zones with some of their characteristic values are shown in section 4.5.4.

In order to verify the consistency and homogeneity of the classes, four additional zones were chosen from each class. These zones were indicated by 1st and 3rd quartile (Q1 and Q3) of total heat demand and of heat density. Table 4-14 below represents the list of quartile values of total heat demand and of heat density.

A list of all chosen quartiles zones are shown in section 4.5.5.

Class	Frequency	HeatTot Quartile 1 (kWh)	HeatTot Quartile 3 (kWh)
1	6	2,86E+08	3,89E+08
2	38	3,01E+08	3,84E+08
3	26	3,30E+08	4,59E+08
4	99	3,68E+08	4,04E+08
5	102	2,54E+08	2,79E+08
6	89	3,03E+08	3,32E+08
7	50	4,54E+08	4,81E+08
8	58	2,35E+08	2,84E+08
9	70	4,25E+08	4,72E+08
10	81	3,24E+08	3,68E+08
11	58	9,79E+07	1,26E+08
12	25	2,24E+08	2,60E+08
13	56	1,63E+08	1,94E+08
14	83	2,31E+07	3,76E+07
15	96	4,91E+07	6,41E+07

Table 4-15: Values of total heat demand of the DE zones representing quartile 1 and quartile 3 of this parameter and the number of zones of all 15 classes

The two zones closest to the quartile values given in Table 4-15 plus the equivalent zones for the quartile values of heat density and the Characteristic Zones have been sent to the University of Manchester for the homogeneity analysis. Based on their feedback it was decided to improve the classification on some classes as described in section 4.5.

4.5 The main goal for improvement of the intermediate results

The intermediate results of classification, covering the 15 classes and described in section 4.4, did not satisfy all expectations. The variation coefficients for both heat total and density heat total in several cases were much higher than 20%, which resulted in unsatisfying homogeneity of zones within classes (see Table 4-16). Thus, it has been decided to re-iterate the classification process for some of the classes which showed potential for improvements. For the following steps the number of parameters has been reduced two total heat and heat density. The effects on the other three parameters are of lesser importance and thus the rules for their variation coefficients have been loosened. For consistency the variation coefficients are still shown for all five parameters in the following Table 4-16 – Table 4-23.

Class	# Zones	Heat Total	Density Heat Total	Load Factor	Linear Heat Density
1	6	19.55%	24.00%	4.18%	28,26%
2	38	23.20%	57.74%	4.92%	48,21%
3	26	21.33%	44.97%	7.26%	34,97%
4	99	5.34%	22.84%	4.30%	13,90%
5	102	7.46%	35.51%	5.55%	21,46%
6	89	5.25%	27.95%	5.47%	18,66%
7	50	4.04%	30.16%	4.96%	19,99%
8	58	12.65%	57.48%	4.36%	43,01%
9	70	6.39%	40.04%	3.49%	26,68%
10	81	7.75%	47.34%	3.50%	26,96%
11	58	15.09%	31.79%	6.79%	25,32%
12	25	11.22%	43.92%	5.74%	50,09%
13	56	9.86%	25.52%	4.44%	24,91%
14	83	28.56%	34.83%	7.21%	31,83%
15	96	15.12%	30.60%	6.22%	21,78%

Table 4-16: Intermediate results of 15 classes are represented by sample size and different figures of variation coefficients. The colour code at the left and right is an indicator for the class homogeneity (green = good, yellow = medium, red = bad).

The experiments have been performed by Work package 4 team demonstrate that in some classes significant diversity between their members have been observed.

For this reason, the reclassification has been done with the goal of decreasing the value of variation coefficient below 20% for the two parameters total heat demand and heat density for all classes, especially the ones with a high demand share.

Finally, the process of selecting Characteristic Zones has been done a third time. The goal was to choose one zone from each class which represent the bulk of the population of their classes in the best way. In this stage the classes' asymmetry has been taken into account and statistical analysis using kurtosis and skewness has been carried out.

4.5.1 Phase 1 – Re-classification

The re-classification of the intermediate results has been carried out on six of the fifteen classes. These classes were indicated by the homogeneity analysis. To this set belong classes with number: 2, 3, 8, 9, 10 and 12 (see Table 4-16). The reminder nine classes have been left unchanged.

The six above mentioned classes have been re-classified using the same cluster analysis procedures (Ward, K-means) as described earlier, but the number of classification factors has been reduced to the total heat demand and the heat density which have proofed to be the most important parameters. As an intermediate result, three new classes have been obtained. Table 4-17 below shows the variation coefficient values of these classes, which are over the 20% for all classes for heat density and for class 3 for total heat demand.

CLUSTER	# Zones	Variation Coefficient (%)				
		Tertiary Heat	Total Heat	Heat Density	Base/Peak	Load Factor
1	136	60,37	16,25	39,72	24,59	7,21
2	135	48,27	11,11	36,62	20,72	5,95
3	27	56,48	23,84	24,63	22,34	10,86

Table 4-17: Intermediate results of the re-classification process. The red colour indicates the unsatisfying results on the two classification parameters.

In the second step of phase 1, the first two groups have been split further to improve the variation coefficients. In total, 7 groups have been created, thus the intermediate classes shown in Table 4-17 have been split into 3, 3, and 1.

Table 4-18 below presents the details of this step.

CLUSTER	# Zones	Variation Coefficient (%)				
		Tertiary Heat	Total Heat	Heat Density	Base/Peak	Load Factor
1a	12	74,53	18,42	16,79	12,26	7,79
1b	68	44,95	6,45	26,45	24,24	7,28
1c	56	67,35	11,57	25,08	26,39	5,94
2a	14	72,58	10,24	19,09	14,84	4,71
2b	64	43,34	5,26	21,74	21,20	6,76
2c	57	38,04	4,98	19,12	16,51	4,54
3	27	56,48	23,84	24,63	22,34	10,86

Table 4-18: Results of phase 1 of the re-classification process, showing seven classes. The red colour indicates the unsatisfying results on the two classification parameters.

The classes obtained in this stage were put together with the remaining nine classes of the intermediate iteration, resulting in 16 new classes (see Table 4-19).

Class	# Zones per Class	Variation Coefficient (%)					
		Heat Tertiary	Heat Total	Heat Density	Base / Peak	Load Factor	LHD
1	6	22,51	19,55	24,00	8,83	4,18	28,26
2	68	44,95	6,45	26,45	24,24	7,28	30,37
3	56	67,35	11,57	25,08	26,39	5,94	24,09
4	99	42,57	5,34	22,84	14,42	4,30	13,90
5	102	48,96	7,46	35,51	18,78	5,55	21,46
6	89	50,26	5,25	27,95	18,09	5,47	18,66
7	50	33,62	4,04	30,16	15,57	4,96	19,99
8	12	74,53	18,42	16,79	12,26	7,79	31,73
9	14	72,58	10,24	19,09	14,84	4,71	21,67
10	64	43,34	5,26	21,74	21,20	6,76	21,06
11	58	84,21	15,09	31,79	29,07	6,79	25,32
12	57	38,04	4,98	19,12	16,51	4,54	20,37
13	56	52,92	9,86	25,52	16,44	4,44	24,91
14	83	80,06	28,56	34,83	28,26	7,21	31,83
15	96	59,32	15,12	30,60	24,78	6,22	21,78
16	27	56,48	23,84	24,63	22,34	10,86	24,59

Table 4-19: Intermediate results of all 16 classes after phase 1 of the re-classification. Variation coefficients on the two parameters total heat demand and heat density are marked red if above 20%.

4.5.2 Phase 2 – Discriminant analysis[30]

In order to improve the results the discriminant analysis has been executed on the two parameters total heat demand and total heat density.

The discriminant analysis is a statistical method, which allocates elements to already existing classes with usage of created classifier (discriminant function). Based on its value the probability of affinity to the particular class is tested. In practice, the distribution of analysed class attributes is analysed and probability density function is used as discriminant function. As a result, possible new localisations for all elements are found. If the value of the classifier of a given element is higher in a new class than in the original one, then the element is moved to the new class. The discriminant analysis can be performed with more than one parameter.

The improved results are shown in Table 4-20.

[30] Discriminant analysis is a general method to separate a data set (objects) into a defined number of classes based on a combination of variables.

Cluster	_FREQ_	HeatTert_CV	HeatTot_CV	DensHeatTot_CV	B/P_CV	HeatTot/Max_CV
1	6	22,51	19,55	24,00	8,83	4,18
2	79	70,82	3,34	28,89	36,09	8,99
3	50	60,43	5,28	22,93	25,61	6,18
4	81	62,36	2,93	21,23	30,68	7,34
5	93	73,61	3,94	23,33	30,62	7,08
6	78	60,53	3,39	22,20	27,60	6,40
7	51	60,43	2,08	21,37	24,09	6,04
8	26	82,02	16,17	26,20	26,18	9,28
9	19	81,58	6,96	28,56	20,38	4,90
10	81	64,95	2,42	24,75	28,15	6,79
11	56	82,76	14,40	32,28	29,31	6,83
12	55	47,26	3,15	18,89	24,70	6,25
13	64	72,70	9,86	28,58	29,34	6,48
14	78	79,74	28,17	36,21	28,06	7,12
15	103	60,64	16,91	28,72	24,69	6,28
16	17	43,62	17,50	22,44	22,79	11,09

Table 4-20: Results of the 16 classes after the discriminant analysis on the parameters total heat and heat density

As mentioned above and in section 4.4.1, the created classes are asymmetric. This asymmetry has influence on the value of the variation coefficients. Therefore, for the determination of the Characteristic Zones and the calculations of the quartiles, it was decided to leave outlying zones out in an extra analysis and to recalculate the variation coefficient for all classes. An improvement of the results has been expected which can be seen in Table 4-21. It shows the variation coefficient for cut and uncut classes:

	Iteration 3			Cutting Density		
Class	# Zones	Heat Total	Heat Density	# Zones	Heat Total	Heat Density
1	6	19,55%	24,00%	5	20,44%	13,43%
2	**79**	**3,34%**	**28,89%**	**78**	**3,16%**	**27,26%**
3	50	5,28%	22,93%	49	5,24%	20,07%
4	*81*	*2,93%*	*21,23%*	*79*	*2,92%*	*18,06%*
5	**93**	**3,94%**	**23,33%**	**92**	**3,92%**	**22,82%**
6	*78*	*3,39%*	*22,20%*	*77*	*3,39%*	*20,37%*
7	51	2,08%	21,37%	50	2,09%	19,88%
8	*26*	*16,17%*	*26,20%*	*24*	*14,93%*	*19,31%*
9	19	6,96%	28,56%	17	7,29%	22,55%
10	**81**	**2,42%**	**24,75%**	**80**	**2,42%**	**22,60%**
11	*56*	*14,40%*	*32,28%*	*54*	*14,62%*	*15,05%*
12	55	3,15%	18,89%	55	3,15%	18,89%
13	64	9,86%	28,58%	61	9,87%	13,55%
14	**78**	**28,17%**	**36,21%**	**73**	**27,87%**	**16,43%**
15	*103*	*16,91%*	*28,72%*	*100*	*17,13%*	*19,11%*
16	17	17,50%	22,44%	16	17,96%	19,89%

Table 4-21: Improved results, both for uncut and cut classes on heat density. Orange ovals show the classes that will be improved, green ones that were considered for improvements, but left unchanged in the end.

4.5.3 Final Iteration

Due to the still unsatisfactorily high values of variation coefficient for some of the classes, an additional classification step was executed for classes with first either high number of classes or large share of the calculated GB demand within all zones and secondly a high variation coefficient for either heat density or total heat demand. For this step, the complete set of the uncut classes have been used.

Therefore, it was decided that classes 1, 3, 7, 9, 12, 13 and 16 show good results or cannot be improved further and thus have been left without changing. Classes 2, 5, 10 and 14 needed updating because of their high variation coefficients of total heat (class 14) or of heat density (classes 2, 5 and 10), while classes 4, 6, 8, 11 and 15 were considered to be updated, but finally were left unchanged, because no improvements could be achieved.

The step has been performed according to the schema presented in Table 4-22. Class 2 has been split at a density above and below 310,000 kWh/ha. The splitting point in class 5 could be found at 300,000 kWh/ha, while 280,000 kWh/ha has been used for class 10. Class 14 has been split on the parameter total heat. Here the value of 32 GWh has been used.

		Variation Coefficients						
Class2	# Zones	Total Heat	Heat Density	Tertiary Heat	Peak	Base/Peak	Load Factor	LHD
DHT>=310000	17	3,82	12,83	64,28	11,78	40,82	11,67	25,67
DHT<310000	62	2,58	12,88	71,48	8,54	32,56	8,17	22,71
Class5	# Zones	Total Heat	Heat Density	Tertiary Heat	Peak	Base/Peak	Load Factor	LHD
DHT>300000	36	4,15	11,57	73,41	9,94	35,26	8,77	23,45
DHT<=300000	57	3,82	11,93	68,94	6,86	23,88	5,85	13,99
Class10	# Zones	Total Heat	Heat Density	Tertiary Heat	Peak	Base/Peak	Load Factor	LHD
DHT>=280000	26	2,09	17,38	70,39	8,64	33,35	7,82	19,23
DHT<280000	55	2,58	10,69	61,71	6,53	25,59	6,27	17,15
Class14	# Zones	Total Heat	Heat Density	Tertiary Heat	Peak	Base/Peak	Load Factor	LHD
HT>=32000000	30	8,76	22,24	60,66	12,90	26,33	7,16	24,30
HT<32000000	45	20,58	15,30	94,71	20,48	23,65	3,59	25,00

Table 4-22: Results of the split classes 2, 5, 10 and 14 into two groups each

Finally, the set of the 20 final new classes has been obtained and are shown in Table 4-23 together with their values of variation coefficients.

		Variation Coefficients (%)				
Class	# Zones	Total Heat	Heat Density	Tertiary Heat	Base/Peak	Load Factor
1	6	19,55	24,00	22,51	8,83	4,18
2	17	3,82	12,83	64,28	40,82	11,67
3	50	5,28	22,93	60,43	25,61	6,18
4	81	2,93	21,23	62,36	30,68	7,34
5	36	4,15	11,57	73,41	35,26	8,77
6	78	3,39	22,20	60,53	27,60	6,40
7	51	2,08	21,37	60,43	24,09	6,04
8	26	16,17	26,20	82,02	26,18	9,28
9	19	6,96	28,56	81,58	20,38	4,90
10	26	2,09	17,38	70,39	33,35	7,82
11	56	14,40	32,28	82,76	29,31	6,83
12	55	3,15	18,89	47,26	24,70	6,25
13	64	9,86	28,58	72,70	29,34	6,48
14	30	8,76	22,24	60,66	26,33	7,16
15	103	16,91	28,72	60,64	24,69	6,28
16	17	17,50	22,44	43,62	22,79	11,09
17	62	2,58	12,88	71,48	32,56	8,17
18	57	3,82	11,93	68,94	23,88	5,85
19	55	2,58	10,69	61,71	25,59	6,27
20	45	20,58	15,30	94,71	23,65	3,59

Table 4-23: Results of all 20 classes, showing the number of zones in each class and the classes' variation coefficients for the 5 classification parameters

The scatter plot below (see Figure 4-26) shows total heat demand distribution between 20 classes. It can be seen that nearly all of the classes form close batches of zones at different levels of heat demands. Only the classes 1, 8, 16 and maybe 9 have a significant distribution within the total heat demands which could have been assumed by the relatively high variation coefficients for heat total (see Table 4-23). The relatively high values on the variation coefficients of this parameter for the classes 11 and 15 can be explained by the small heat demands per zone in these classes. Thus, the variation coefficients are higher, although the scatter plot shows lesser deviation.

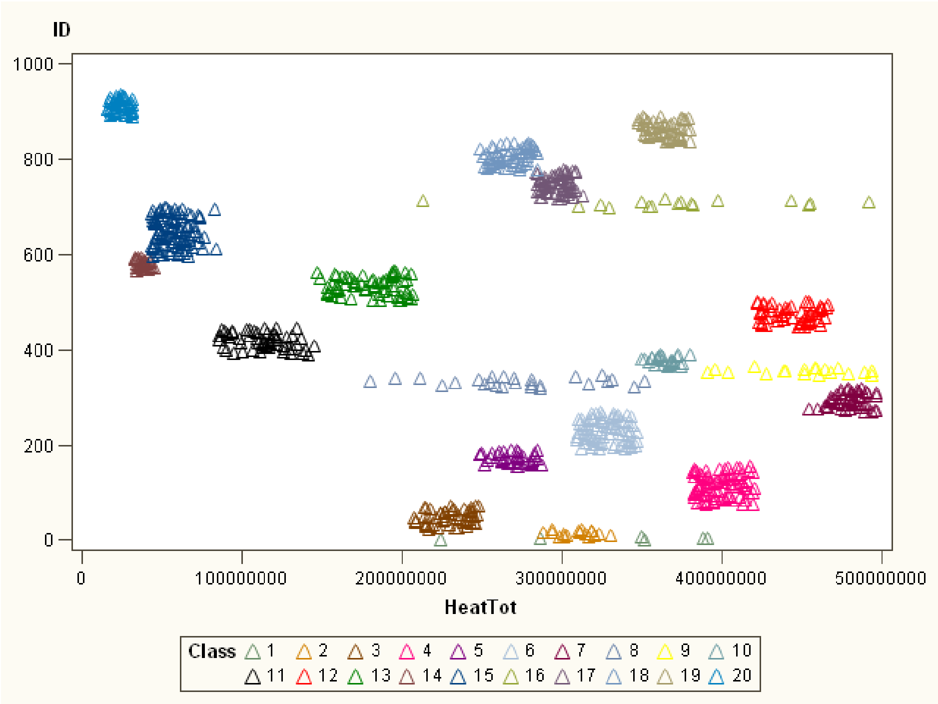


Figure 4-26: Scatter plot of zones in all classes by total heat demand

4.5.4 Determining of Characteristic Zones

The characteristic of all classes is their asymmetry. This effect is easy to observe using scatter plots like Figure 4-26 and Figure 4-27 or analysing the skewness and kurtosis values in the classes. The outcome is especially visible for the parameter heat density in individual classes (see columns for DensHeatTot in Table 4-24). High and positive kurtosis' values indicate strong concentration of heat density values around the mean (e.g. classes 11, 13 and 15) and positive values of skewness signify right asymmetries (e.g. same classes 11, 13 and 15).

	HeatTot		DensHeatTot	
Class	**Kurtosis**	**Skewness**	**Kurtosis**	**Skewness**
1	0,10	-1,02	2,06	1,46
2	-0,88	-0,11	1,01	0,57
3	-122012,00	-0,19	2,72	1,59
4	-1,23	0,21	2,58	1,48
5	-0,95	-0,24	-0,57	0,61
6	-1,25	0,05	1,96	1,39
7	-0,33	-0,34	0,15	0,75
8	-0,26	-0,13	1,51	1,30
9	-0,55	-51225,00	-0,11	0,97
10	-0,52	0,00	2,15	1,47
11	-0,89	-0,06	24,07	4,44
12	-1,36	-0,13	0,28	1,00
13	-1,30	-0,22	11,77	3,33
14	-0,99	0,19	1,35	1,30
15	-0,40	0,50	11,84	3,05
16	1,36	-0,41	-0,18	0,51
17	-1,17	-0,08	0,18	1,11
18	-1,20	-0,06	-0,96	0,44
19	-1,14	-0,01	-1,41	0,33
20	-1,15	0,11	1,25	1,29

Table 4-24: Kurtosis and Skewness for all 20 classes on the two parameters total heat demand and heat density

That means that the majority of zones within the classes create the concentrated agglomeration in area of the lowest values of density heat total and only a very small amount are detached, creating the “tail”.

Figure 4-27 below illustrates this situation:

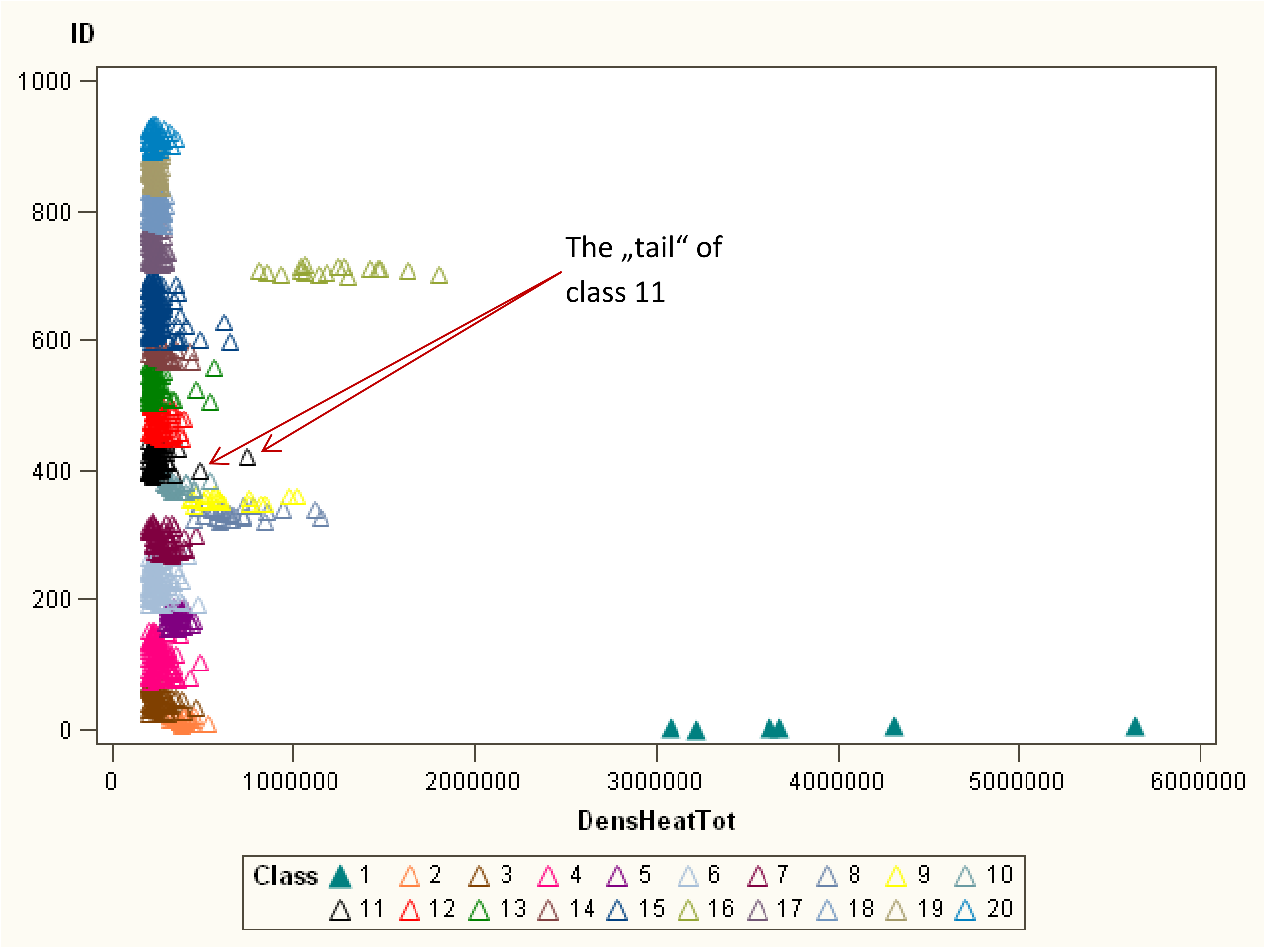


Figure 4-27: Scatter plot of all 20 classes for energy density, focusing on the tails of the classes

Hence two processes of determining the Characteristic Zones were created:

1. Using all existing class members for the determination of the Characteristic Zones,
2. the choice of Characteristic Zones after cutting elements of the tails.

The first option means the selection of representatives out of all zones in each class considering possible weakening influences on the results by detached elements. The second option guarantees stronger representativeness of potential Characteristic Zones but with lesser agreement on the tails.

The initial estimation shows that around 3 to 5% of the population of zones form these tails. Based on this estimation, the second option has been selected to determine the Characteristic Zones. The full compatibility of Characteristic Zones with the majorities of class members in the best possible way has been expected.

For the 20 classes the analysis of kurtosis and skewness shown in Table 4-24 has been used to decide on the cutting of tails. Because the results for heat density are more meaningful than for the other parameters, the cutting of tails has been done on this parameter. Each class has been analysed separately, without using a predefined amount of eliminated zones. Figure 4-28 shows example cuttings for the two classes 11 and 13. As the skewness values for total heat density are positive for all classes, the cut zones of the tails can always be found on the higher dense end of the classes.

Thus, the mean values for heat density per class are slightly reduced which lead to a stronger assumption that all members of a class are economic viable if the Characteristic Zones are viable themselves.

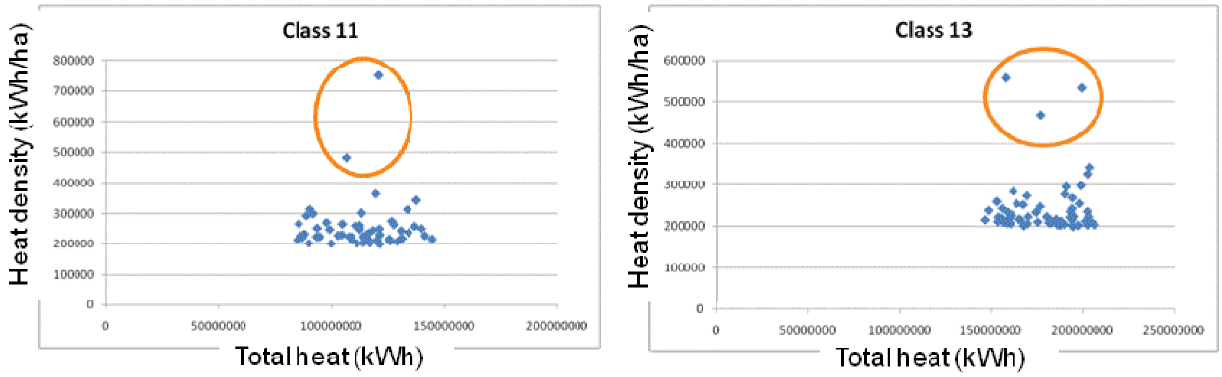


Figure 4-28: Cut zones tails in the two example classes 11 and 13

Table 4-25 below contains the values of variation coefficient of 20 cut classes.

		Variation Coefficients (%)				
Class	**# Zones**	Total Heat	Heat Density	Tertiary Heat	Base/Peak	Load Factor
1	5	20,44	13,43	22,37	8,74	3,37
2	16	3,51	10,46	60,83	42,07	11,97
3	49	5,24	20,07	60,75	25,17	6,16
4	79	2,92	18,06	61,30	29,99	7,20
5	36	4,15	11,57	73,41	35,26	8,77
6	77	3,39	20,37	61,05	27,74	6,36
7	50	2,09	19,88	56,85	23,21	6,00
8	24	14,93	19,31	85,20	27,27	9,67
9	17	7,29	22,55	78,75	21,31	5,10
10	24	2,12	13,03	73,11	34,58	8,13
11	54	14,62	15,05	79,14	28,22	6,81
12	55	3,15	18,89	47,26	24,70	6,25
13	61	9,87	13,55	67,13	28,68	6,42
14	28	8,77	16,77	62,87	26,94	7,23
15	100	17,13	19,11	60,57	24,94	6,27
16	16	17,96	19,89	45,27	23,31	11,06
17	60	2,59	13,05	71,67	32,60	8,14
18	57	3,82	11,93	68,94	23,88	5,85
19	55	2,58	10,69	61,71	25,59	6,27
20	41	20,52	10,87	94,29	23,34	3,51

Table 4-25: Results of all 20 classes after cutting, showing the number of zones per class and the variation coefficients for the 5 classification parameters

In total, the cut set of classes consist of 904 zones that represent the 96.79% of all initial zones. These updated classes were used to determine the final set of Characteristic Zones, according to the following procedure:

- Defining the centroids of all classes,
- Calculating the Euclidian distance between centroids and every zone of the class,
- Selecting the points (zones) closest to the centroids (Characteristic Zones).

The centroid for each class has been calculated as a value equal to the mean of each dimension. Table 4-26 below shows the values of means for the 5 classification parameters:

Class	Mean of total heat	Mean of Heat Density	Mean of Tertiary Heat	Mean of Base/Peak	Mean of Load Factor
1	3,20E+08	3,58E+06	2,64E+08	0,05	0,28
2	3,08E+08	3,91E+05	7,26E+07	0,04	0,33
3	2,30E+08	2,51E+05	3,42E+07	0,03	0,33
4	3,99E+08	2,56E+05	5,83E+07	0,03	0,33
5	2,69E+08	3,56E+05	5,22E+07	0,03	0,32
6	3,28E+08	2,57E+05	5,20E+07	0,03	0,33
7	4,80E+08	2,78E+05	8,29E+07	0,03	0,32
8	2,78E+08	6,42E+05	7,29E+07	0,04	0,32
9	4,49E+08	5,87E+05	1,08E+08	0,04	0,33
10	3,65E+08	3,38E+05	6,46E+07	0,03	0,33
11	1,13E+08	2,43E+05	1,42E+07	0,03	0,33
12	4,44E+08	2,65E+05	7,91E+07	0,03	0,33
13	1,80E+08	2,30E+05	2,98E+07	0,03	0,33
14	3,75E+07	2,64E+05	2,59E+06	0,03	0,33
15	5,68E+07	2,48E+05	6,47E+06	0,03	0,33
16	3,75E+08	1,18E+06	1,97E+08	0,05	0,32
17	2,97E+08	2,32E+05	5,05E+07	0,03	0,33
18	2,68E+08	2,40E+05	3,80E+07	0,03	0,32
19	3,64E+08	2,34E+05	5,77E+07	0,03	0,33
20	2,32E+07	2,33E+05	1,51E+06	0,03	0,36

Table 4-26: Class means of all five classification parameters for all classes, the two main parameters, total heat and heat density, are highlighted

The distances from the classes' centroids of each zone have been calculated using the Euclidean distance on the two parameters total heat demand and density heat total. The equation has been presented in section 4.2.3.

The zones with the shortest distance to the classes' centroids have been selected as Characteristic Zones. These 20 chosen zones with a set a characteristic values are shown in Table 4-27. This table has been sent to the University of Manchester together with the calculated time band for heat and electrical demand to be used for the benefit case.

Characteristic Zones (Total Heat & Heat Density)										
Class Nr.	1	2	3	4	5	6	7	8	9	10
Zone Nr.	154	711	239	619	70	509	127	145	29	138
Heat Total (kWh)	3,49E+08	3,10E+08	2,30E+08	4,00E+08	2,70E+08	3,28E+08	4,80E+08	2,81E+08	4,52E+08	3,64E+08
Electric Total (kWh)	2,42E+08	8,97E+07	7,41E+07	2,37E+08	1,93E+08	1,02E+08	1,36E+09	1,12E+08	4,37E+08	1,36E+08
Heat Density (kWh/ha)	4,31E+06	3,18E+05	2,23E+05	2,29E+05	3,28E+05	2,22E+05	2,29E+05	6,01E+05	8,47E+05	3,46E+05
Elec. Density (kWh/ha)	2,99E+06	9,20E+04	7,20E+04	1,36E+05	2,35E+05	6,90E+04	6,46E+05	2,39E+05	8,19E+05	1,29E+05
Zone Population	15089	42350	28259	36680	37325	45254	64948	41887	47184	67152
Zone Area (hectare)	81	976	1027	1749	821	1478	2097	467	533	1051
# of Households	8361	18551	11605	14890	15296	18787	27243	18779	21924	28590
# of Tertiary	3206	892	2052	3848	8388	1842	5657	2999	15295	2981
# of MLSOAs	2	6	4	5	6	6	9	6	7	9
Thermal Peak (kWh)	1,42E+05	1,08E+05	7,71E+04	1,40E+05	1,04E+05	1,13E+05	1,90E+05	1,07E+05	1,60E+05	1,34E+05
Thermal Base (kWh)	6,20E+03	2,60E+03	2,70E+03	4,90E+03	2,80E+03	3,10E+03	3,90E+03	3,10E+03	8,30E+03	4,20E+03
Class Nr.	11	12	13	14	15	16	17	18	19	20
Zone Nr.	566	313	686	535	734	148	23	547	701	909
Heat Total (kWh)	1,13E+08	4,44E+08	1,81E+08	3,76E+07	5,67E+07	3,74E+08	2,97E+08	2,68E+08	3,64E+08	2,33E+07
Electric Total (kWh)	2,78E+07	1,83E+08	5,61E+07	1,16E+07	2,04E+07	5,02E+08	9,94E+07	9,01E+07	1,23E+08	7,30E+06
Heat Density (kWh/ha)	2,48E+05	2,23E+05	2,24E+05	2,52E+05	2,22E+05	1,46E+06	2,79E+05	2,78E+05	2,07E+05	2,14E+05
Elec. Density (kWh/ha)	6,10E+04	9,20E+04	7,00E+04	7,80E+04	8,00E+04	1,96E+06	9,40E+04	9,30E+04	7,00E+04	6,70E+04
Zone Population	15128	49510	21070	5510	7724	22075	36020	23745	41394	3222
Zone Area (hectare)	455	1993	806	149	255	256	1063	964	1757	108
# of Households	6461	19912	8768	2312	3150	10586	15063	10543	18045	1576
# of Tertiary	476	3807	323	158	563	12227	3551	2154	1968	70
# of MLSOAs	2	7	3	1	1	3	5	3	5	1
Thermal Peak (kWh)	3,82E+04	1,53E+05	5,98E+04	1,36E+04	1,98E+04	1,45E+05	1,19E+05	8,86E+04	1,21E+05	7,30E+03
Thermal Base (kWh)	1,10E+03	4,90E+03	2,10E+03	4,00E+02	5,00E+02	7,00E+03	2,30E+03	3,40E+03	4,30E+03	3,00E+02

Table 4-27: Values of Characteristic Zones

4.5.5 Additional Zones (quartiles)

In addition to the Characteristic Zones, a set of 4 further zones from each class has been selected. The role of these zones is an additional verification of coherence of each class. The lesser the difference between these points, the more homogeneous and compacter are the classes. This means that Characteristic Zones represent their own classes in a good way.

In order to define the additional points from classes, the quartiles for the two parameters total heat and heat density for all 20 classes have been determined. This analysis has been executed on the cut classes, the same as for determination of the Characteristic Zones.

Table 4-28 below contains the values of first (Q1) and third (Q3) quartiles of final set of classes. The last two columns show the ratio between Q1 and Q3 that represents the similarity degree between quartiles, expressed by percentage. In case of the same values (Q1=Q3) the ratio is 100%. In general, this ratio for total heat is very high. 17 out of 20 classes have higher values than 80%, and the last three are between 70% and 80%. For heat density, the ratio is as high, but still all values are higher than 70%. Thus, it can be concluded that the Characteristic Zones can good representatives for their classes.

Class	Quartile 1 of Total Heat (kWh)	Quartile 1 of Heat Density (kWh)	Quartile 3 of Total Heat (kWh)	Quartile 3 of Heat Density (kWh)	Ratio of Q1 to Q3 for total heat	Ratio of Q1 to Q3 of heat density
1	2,86E+08	3,22E+06	3,51E+08	3,68E+06	81,50%	87,64%
2	2,99E+08	3,63E+05	3,17E+08	4,26E+05	94,37%	85,30%
3	2,20E+08	2,09E+05	2,41E+08	2,81E+05	90,98%	74,38%
4	3,89E+08	2,19E+05	4,10E+08	2,93E+05	95,03%	74,88%
5	2,61E+08	3,24E+05	2,79E+08	3,83E+05	93,74%	84,48%
6	3,18E+08	2,15E+05	3,37E+08	2,84E+05	94,40%	75,64%
7	4,73E+08	2,29E+05	4,88E+08	3,15E+05	96,99%	72,59%
8	2,56E+08	5,69E+05	3,12E+08	7,22E+05	81,88%	78,81%
9	4,28E+08	5,04E+05	4,68E+08	6,08E+05	91,28%	82,96%
10	3,60E+08	3,02E+05	3,70E+08	3,59E+05	97,29%	83,97%
11	9,92E+07	2,17E+05	1,26E+08	2,61E+05	78,77%	83,14%
12	4,30E+08	2,23E+05	4,55E+08	2,85E+05	94,50%	78,16%
13	1,64E+08	2,10E+05	1,94E+08	2,39E+05	84,36%	87,65%
14	3,43E+07	2,32E+05	4,03E+07	3,02E+05	85,08%	76,85%
15	4,82E+07	2,12E+05	6,45E+07	2,64E+05	74,75%	80,19%
16	3,40E+08	1,04E+06	4,20E+08	1,36E+06	80,75%	76,10%
17	2,90E+08	2,08E+05	3,04E+08	2,42E+05	95,48%	85,96%
18	2,59E+08	2,16E+05	2,77E+08	2,57E+05	93,71%	84,17%
19	3,55E+08	2,13E+05	3,73E+08	2,60E+05	95,34%	82,05%
20	1,96E+07	2,13E+05	2,72E+07	2,42E+05	72,08%	87,96%

Table 4-28: Results for quartiles 1 and 3 for all 20 classes for the two parameter total heat and heat density, including their Q1/Q3 ratios

Chapter 5 — Application of DE design tool to evaluate characterisation of zones

5.1 Introduction

The ETI Macro DE project aims to evaluate future opportunities for application of macro-scale distributed energy for district heating, with electricity generation in the range 100 kWe to 50 MWe. To this end, the Macro DE project aims to characterise energy consumption in Great Britain (GB). Once characteristic zones have been defined, the DE design tool developed within the project will be used to design the energy centre and cost DE solutions (the energy centre and district heating network) to meet the heat requirements of each characteristic zone. Provided that the characteristic zones are indeed representative of classes of geographic areas (or 'zones'), the design results can be used to evaluate the performance of all zones in a class. This means that a limited number of zones – only the characteristic zones – need to be assessed to build an understanding of the potential of DE for GB.

In Work Package 4 (WP4), the University of Manchester has developed a methodology for the design of energy centres and has implemented this methodology in a software tool. Deliverable 4.1 (February 2011) describes the modelling and optimisation approach and the assumptions applied in the modelling and demonstrates that the designs generated by the tool are well aligned with industrial practice and with important economic and performance trends

A summary of the application of the design methodology to characteristic zones and other zones identified in the clustering and classification process of Work Package 2 is presented. The design tool is applied to investigate how different the performance of DE solutions is between classes, and how similar the performance of DE solutions is within each class. The analysis of these design results allows qualitative and quantitative appreciation of the effectiveness of the classification procedure and provides confidence in the concept of using characteristic zones to represent a large number of diverse zones in GB.

5.2 DE design methodology – summary

The objective of the design methodology is to optimise the DE solution (energy centre) for a given zone with a known heat demand. This means selecting a suite of DE systems (e.g. engines and boilers of given capacities, and thermal storage units) and their operating schedule to meet the heat demand with the minimum operating and capital costs. The cost of the associated district heating network (DHN) is accounted for using a simple cost algorithm, although the heat network itself is not designed or optimised. The CO2 emissions generated by the DE solution are computed, taking into account fuel combustion and avoided emissions associated with electricity generated. The heat demand is expected to vary with the season and time of day; 'time bands' are defined to represent typical use on a given type of day (weekday or weekend), season (summer, winter or 'transition'), and time of day, as shown in Table 5-1.

	Weekdays (WD)	Hours	Winter (W)	Summer (S)	Transition (T)
	0:00 - 04:59	5	53	10.6	30.5
	5:00 - 6:59	2	105.8	35.4	71
	7:00 - 9:59	3	104.6	32.8	69.1
	10:00 - 16:59	7	89.3	22.7	54.5
	17:00 - 18:59	2	98.7	23.8	61.3
	19:00 - 21:59	3	94.2	25.8	60.9
	22:00 - 23:59	2	56.3	13.6	33.8
	Baseload		12.8	11.2	4.1
Thermal, MW	**Peak**		167.7	124.2	65.8
	Days		87	86	88
	Weekends (WE)	**Hours**			
	0:00 - 04:59	5	54	30	9
	5:00 - 6:59	2	107.9	70.9	32.3
	7:00 - 9:59	3	105.8	68.7	29.7
	10:00 - 16:59	7	91.2	54.3	20.3
	17:00 - 21:59	5	99.4	61.9	22.4
	22:00 - 23:59	2	58.9	34.7	12.4
	Days		34	36	34

Table 5-1: Example of heat demand of a zone by time band

The design objective is to minimise the cost of meeting the heat demand of a zone by the DE solution (energy centre and district heating network). The costs include capital costs and operating costs – i.e. fuel costs, electricity costs for pumping and maintenance costs; electricity sold generates revenue. Maintenance costs include variable maintenance costs that depend on the output of the unit (e.g. MWh of heat produced) and fixed maintenance costs, to allow the asset to operate for its full lifetime, e.g. including replacement of parts.

Important trade-offs in the design problem relate to:

I) the number, type and size of DE systems – where larger units benefit from economies of scale and improved efficiencies, but have higher capital costs;

II) the operating schedule of the centre – as revenue from electricity generation depends in principle on the season and time of day;

III) the type of primary energy sources – fossil fuels cost less than biofuels, but have higher CO2 emissions.

The design methodology applies optimisation techniques to account for these trade-offs while taking into account relevant constraints.

The design problem – the choice of DE systems and their operating schedules – is formulated using integer variables and linear relationships to represent the number and type of DE systems and their performance, e.g. fuel consumption, electricity generation, thermal storage requirements, as well as relevant constraints. Such a formulation (a mixed-integer linear program, or MILP) facilitates optimisation, in spite of the relatively large number of variables and relationships involved. Details are presented in Deliverable 4.1.

Constraints applied in the design pertain to practical limitations, such as the minimum part-load of combined heat and power (CHP) units, suitable redundancy in the design and the daily energy balance. Design assumptions are summarised in Table 5-2. Deliverable 4.1 presents the design assumptions in more detail; some new assumptions have been made to provide more practical designs and address more realistic scenarios.

Heat and electrical power demand	The energy centre must satisfy maximum peak thermal demand. Heat demand is characterised using time bands during which the demand is assumed constant. Only new DE solutions (energy centres and district heating networks) are assessed.
District Heating Network (DHN)	All heat demand in a zone is produced by the energy centre and delivered via the district heating network (DHN). To cover thermal losses from the district heating network, heat production exceeds heat delivered by 10% (in each time band) The temperature difference between supply and return is assumed to be 40°C. The methodology costs, but does not design, the district heating network. Costs cover the piping network from the energy centre to dwellings and heat interface units (HIU) in dwellings. The cost model used for the district heating network will be reported in a separate deliverable. A summary is presented in Table 5-3. The general formula is: DHN capital cost (£, 2010) **=** Heat Demand (GWh/y) · [Length of transmission pipe (m/GWh) · Cost of transmission pipe (£/m) **+** Length of distribution pipe (m/GWh) · Cost of distribution pipe (£/m)] **+** Cost of domestic heat interface units **+** Cost of tertiary* heat interface units *'tertiary' = commercial and industrial users Electricity required for pumping (kWh) is 2% of the heat demand (in kWh) in each time band; the electricity will be supplied by the national grid. Refurbishing of houses and buildings is out of scope.
Revenue and expenditure	The basis for design cost calculations is 2009/2010. User inputs related to some other scenarios can be accommodated. The cost of generating heat will be determined (income from heat generation will not be included). **All** electricity generated will be sold to the grid. Electricity sales prices and costs are based on an average value of 6.36 p/kWh (DECC IAG variable element: industrial for 2010)[31]: • The peak rate (07:00 -24:00) is taken to be 7.0 p/kWh (avg. +10%) • The off-peak rate (00:00 – 07:00) is taken to be 4.8 kWh/p (avg. –25%). For purchased electricity, the same costs are applied *i.e.* the electricity used in the

[31] Source: DECC Interdepartmental Analysts Group (IAG) (www.decc.gov.uk/en/content/cms/what_we_do/lc_uk/valuation/valuation.aspx) Supply cost reflects long-term variable cost components of energy supply, and excludes costs that do not change with level of demand. The DECC IAG variable costs include costs of energy inputs; conversion, transmission and distribution costs, including costs of capital and operating expenditure associated with the supply network. The wholesale price of electricity (paid by the industrial sector) is applied; losses on the local distribution network are taken to be 5%. Most transmission costs (90%) are included, but only 10% of distribution costs are included as variable costs. Note that this electricity price is above that which is commercially available in the current market; however, it aims to represent the value of the electricity generated in terms of the national economy.

	energy centre reduces the net revenue from electricity sold. The gas price is assumed to be 2.64 p/kWh (DECC IAG retail: industrial price for 2010, gross calorific value basis *i.e.* higher heating value basis). Any capital items with a lifespan of less than the project life will need to be replaced; replacement costs will be incurred at the project outset. If more than one unit of a supply technology is required during the project lifetime, the capital cost of replacement units is reduced by 30% (assuming new foundations, piping, instrumentation, etc. would not be required). The cost of housing the energy centre and service facilities will account for 25% of the total fixed investment (capital expenditure). Engineering and contingency costs will account for 20% of the total fixed investment (capital expenditure). Annualised capital investment is spread over the project life, by applying an annual capital charge ratio based on an interest rate of **8%** and project life of **25 years** for both the energy centre and the DHN.
Supply technologies	To provide a practical level of redundancy in the energy centre: i) The maximum peak thermal demand must be met by boilers alone – *i.e.* without relying on CHP units, thermal storage units or solar heaters. ii) The maximum capacity of boilers should be less than 50% of the maximum peak thermal demand. The minimum load of DE systems is accounted for: engines: 50%; turbines: 50%; fuel cells: 50 %; boilers: 5% Fuel (natural gas) consumption, reported in terms of net calorific value (lower heating value) in the models, is converted to the gross CV / higher HV using a factor of 1.109 [http://chp.decc.gov.uk/cms/fuel-calorific-value/]. The annual amount of electricity generated is not limited. Table 5-4 summarises the DE systems applied for evaluating characteristic zones.
Thermal storage	Heat can be stored as hot water at the supply temperature (90°C); the energy balance is closed for each 24 hour period. Longer term storage solutions and other storage technologies are not considered. Thermal storage units are sized by assuming that 10% of the total volume of the tank is not available for heat supply. Thermal losses per day are assumed to be 2% of the total energy content of the storage unit. The default is to include thermal storage units in the design.
CO_2 emissions	CO_2 generated and emitted includes: • CO_2 generated through fuel combustion (WP3 data) • CO_2 associated with generation of imported electricity (0.485 kg/kWh) For any exported electricity, equivalent credit (0.485 kg/kWh) is given to account for emissions avoided by reducing centralised power production. 'Embodied' carbon, *i.e.* CO_2 emissions associated with the manufacture of DE systems as given in WP3 data, is *not* included in the overall emissions (to allow like-for-like comparisons to be made with the baseline). CO_2 emissions from the DE solution (emissions from fuel combustion less avoided emissions claimed for exported power) are compared with the baseline CO_2 emissions from fuel burnt in individual boilers.

Table 5-2: Design assumptions for analysis of characteristic zones

Zone type	Annual heat demand (£, 2010)
Low density (LD) < 0.35 GWh/hectare	HeatDemand (GWh/y) x (**725.3** m/GWh x 512.5 £/m + **556.6** m/GWh x 132 £/m) + NumberHH x 1,750 £/meter + NumberTertiary x 4,800 £/meter
High density (HD) > 0.35 GWh/hectare, < 0.04 GWh/meter	HeatDemand (GWh/y) x (**543.8** m/GWh x 512.5 £/m + **360.9** m/GWh x 132 £/m) + NumberHH x 1,750 £/meter + NumberTertiary x 4,800 £/meter
High density, high intensity (HD/HI) > 0.35 GWh/hectare, > 0.04 GWh/meter	HeatDemand (GWh/y) x (**482.2** m/GWh x 512.5 £/m + **381.3** m/GWh x 132 £/m) + NumberHH x 1,750 £/meter + NumberTertiary x 4,800 £/meter

Table 5-3: Simple algorithm for estimating capital cost of district heating network
NumberHH: number of households; meter: gas meter; NumberTertiary: number of tertiary meters

Technology		Available sizes (thermal capacity, MWh_{th})					
1	Gas engine	0.47	1.62	2.29	4.42	5.24	7.00
2	Gas turbine	8.38	15.23	18.39			
3	Gas boiler	0.25	1.40	3.50	7.00	10.00	20.00
4	Diesel engine	0.52	1.39	5.15			
5	Biodiesel engine	0.52	1.39	5.15			
6	Landfill gas engine	0.47	1.39	2.74			
7	Fuel cell (natural gas)	0.50					
8	Solar heater	0.1					
9	Heat pump (waste heat)	3.17	4.15				

Table 5-4: Summary of DE systems included in zone characterisation study

5.3 DE solution design model

Two of the main constraints of the problem are the system energy balances for power and heat. The total heat produced within the energy centre must be at least equal to total heat demand. Producing more heat than the required is undesirable because fuel consumption and CO2 emissions will increase, but can bring benefits in terms of increased revenue from exported electricity. The design model allows the option of 'dumping' excess heat, but in the results presented in this analysis, no heat dumping is allowed.

A thermal storage tank can be included by the user in order to allow CHP units to keep working when thermal demand falls but electricity generation is profitable. A thermal storage tank is a heat accumulator used for storage of water-based energy. The difference between the total heat produced by the energy centre and the heat demand in each time band defines the amount of water (heat) being diverted to the thermal storage or extracted from it. To avoid computational complexity, the heat accumulation cycle is taken to be 24 hours. The maximum heat storage capacity of the tank is computed by the design tool to minimise operating costs.

Figure 5-1 provides an example of an optimised DE solution to satisfy the demand presented in Table 5-1. It may be seen that heat generated in a time band sometimes exceeds demand, and sometimes does not meet demand: the thermal storage unit is used to transfer heat between time bands. It may be noted that electricity production is lowest in time bands 1 and 2 (between midnight and 07:00) when the sales prices of electricity is lowest and heat demand is also relatively low.

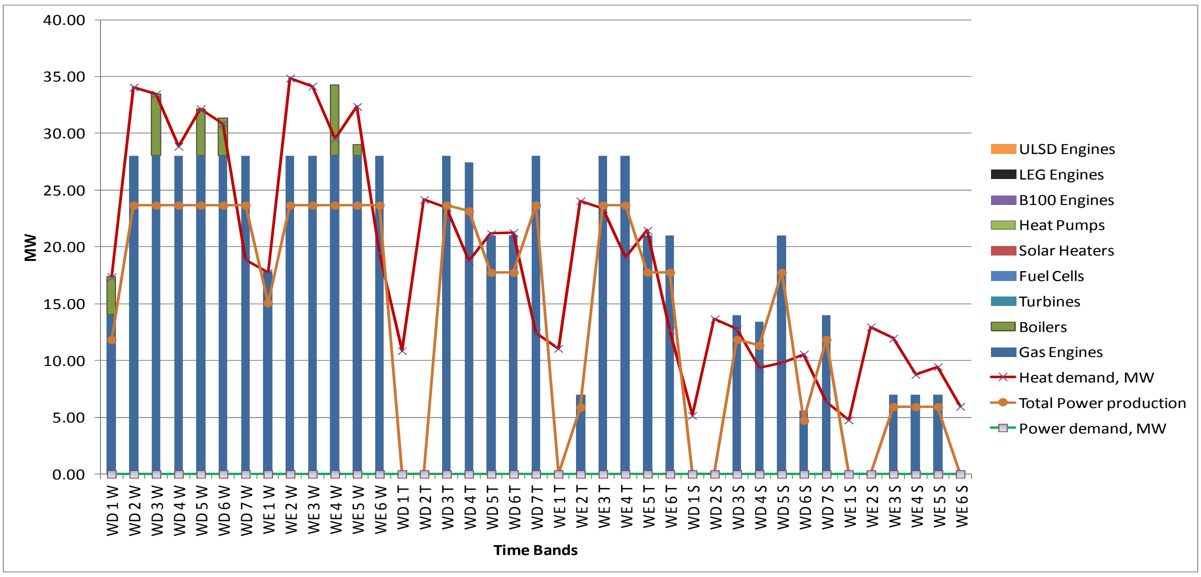


Figure 5-1: Optimised DE solution, showing DE systems and operating schedule by time band

To provide redundancy to allow the heat demand of a zone to be satisfied, the sum of the total thermal capacity of the DE systems, excluding solar heaters and the thermal storage unit, must be at least larger than the heat peak demand of the zone. Furthermore, it must be possible to satisfy the peak heat demand by boilers alone (rather than CHP units), none of which should provide more than 50% of the peak heat demand.

The 'baseline' for performance analysis is that all electricity is supplied by the main power grid and the heat demand is met by gas boilers (80% efficient, 8 year life span, as defined in Deliverable 1.1). The design tool compares the performance of optimised DE solutions to this baseline.

5.4 Software Tool Application

The design methodology has been implemented in Microsoft Excel, with the add-in WhatsBest! for optimisation. The optimiser minimises the total annualised cost of the DE solution and calculates the CO_2 emissions associated with operating the energy centre. Constraints may be imposed by the user, e.g. to select or deselect certain DE systems, to achieve a given level of CO_2 emissions less than a specified limit, etc. To optimise the design and operation of the energy centre to satisfy the thermal demand of a zone takes around 5 minutes computing time on a desktop PC.

5.5 Evaluating DE solutions of characteristic zones and classes of zones

The overall DE solution performance will be evaluated in terms of its economic performance and CO_2 emissions, including CO_2 emissions associated with the energy centre and those associated with imported and exported electricity (emissions or avoided emissions from centrally generated electricity).

The concept of a 'characteristic zone' in principle allows the potential for GB of macro-scale DE to be evaluated in terms of cost and CO_2 emissions. DE solutions are generated for a limited number of 'characteristic zones' and the performance of the DE solutions is evaluated; this performance is then ascribed to all zones in that class, allowing all classified zones to be evaluated.

It is appropriate to interrogate the classes developed through the clustering and classification approach of this project. Key questions include:

- How *distinctive* are the classes of zones? (If several zones have very similar features and performance, it might make more sense to lump them into a single class.)
- How *homogeneous* are the classes, *i.e.* how similar are the zones within a class? (If zones within a class have DE solutions that perform very differently, then the characteristic zone cannot be said to be representative of other members of the class.)

To address the first question, the key performance indicators of the DE solution for each characteristic zone are compared: these are the total annualised cost of the DE solution, the district heating network and of the overall installation; the CO_2 emissions from fuel combustion and 'global' CO_2 emissions – those from fuel combustion less those avoided.

To address the second question, relating to the homogeneity of classes of zones, DE solutions are generated for a small number of zones in a class and their performance is compared. Two significant classification parameters are the total annual heat demand (GWh) of a zone and the thermal demand density (or 'heat density') of the zone (GWh/ha) as the first impacts on the capital and operating costs of the energy centre and both parameters impact on the capital cost of the DHN. The zones to be compared are selected to represent the range of these classification parameters within a class – therefore two zones, representing the first and third quartiles with respect to heat demand, and two zones representing the first and third quartiles with respect to heat density, are selected.

In order to assess how similar the performance of these four 'quartile zones' are to that of the 'characteristic' zone of a class, DE solutions are generated for five zones in each class and their performance is compared by plotting various performance indicators against annual thermal demand. To remove effects of scale from the analysis, the performance of the zones in each class is also compared in terms of costs and CO2 emissions per unit of heat delivered. The maximum difference in performance (per unit of heat delivered) between any 'quartile' zone and the characteristic zone is used as a quantitative measure of homogeneity. These specific performance indicators (i.e. per unit of heat delivered) will be used in WP5 to extrapolate performance results for the characteristic zones to the performance of whole classes, based on the total thermal demand of the class. The GB case will be analysed based on performance results for all classes.

Given that the analysis is being carried out with quartile zones, the analysis deliberately aims to consider a significant range of characteristic behaviour in each class. The selection of heat demand and heat density to define relevant quartiles follows from the understanding that heat demand impacts directly and significantly on a range of economic issues (capital cost of DE centre and DHN, fuel cost, etc.) and that heat density impacts significantly on the DHN cost, which is expected to dominate the process economics.

Investigation	Performance indicators
Distinctiveness and homogeneity of classes	• total annualised cost of DE centre (£/yr) • DHN annualised capital cost (£/yr) • annualised capital cost of the DE centre (£/yr) • total annualised cost of DE solution (£/yr) • fuel CO_2 emissions (t/yr) • global CO_2 emissions (t/yr) (including fuel emissions and emissions credits for electricity export)
Homogeneity of a class (Appendix H —)	• total annualised cost of DE centre per unit of heat delivered (£/kWh) • DHN cost per unit of heat delivered (£/kWh) • total annualised cost of DE solution per unit of heat delivered (£/kWh) • annualised capital cost of DE centre per unit of heat delivered (£/kWh) • fuel CO_2 emissions per unit of heat delivered (t/MWh) • CO_2 emissions avoided per unit of heat delivered (t/MWh) • proportion of heat produced by CHP units as an indicator of the technical solution selected by the model.

Table 5-5: Performance indicators for distinctiveness and homogeneity

5.5.1 Results of homogeneity studies

The design methodology and assumptions have been briefly described above (and Deliverable 4.1 provides a more detailed explanation). Design results should agree with expected trends. For example, zones with the highest annual thermal demand should have the largest cost and highest CO2 emissions.

Designs were generated for the characteristic zone and four quartile zones for all 20 classes using the software tool developed in WP4 by the University of Manchester. Relevant input data for the 20 characteristic zones are presented in Appendix G —, in Table G-1: Design inputs for characteristic zones; design results are summarised in Table G-2: Design results for characteristic zones. Figure 5-2 presents the annual thermal demand and heat density for the characteristic zones.

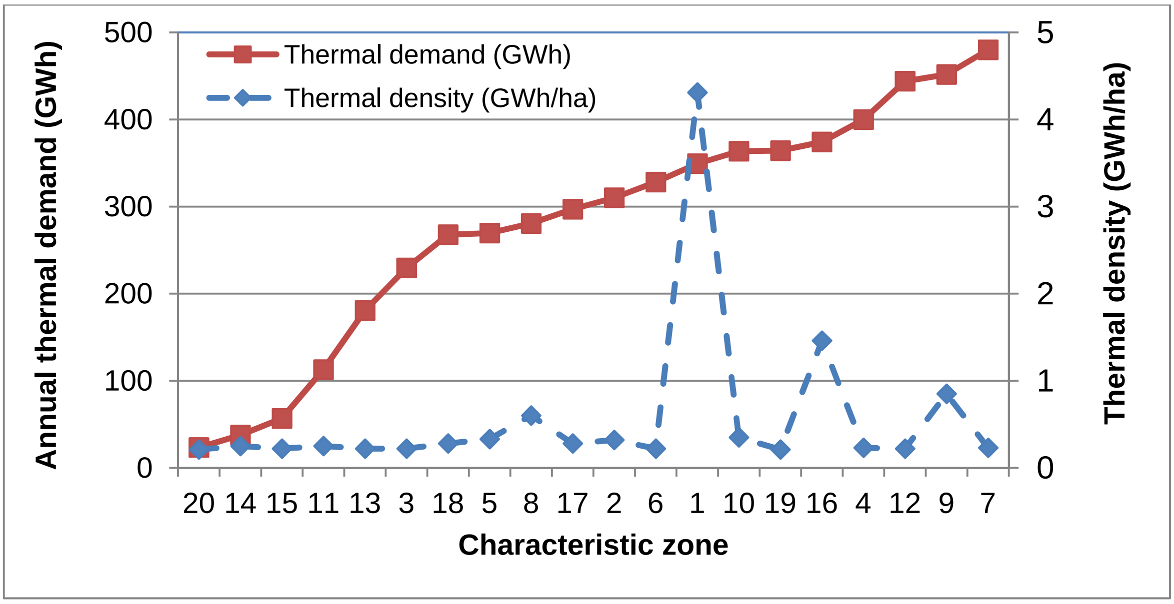


Figure 5-2: Thermal demand and thermal density of characteristic zones

Table 5-6 indicates how the two most important classification parameters – thermal demand and thermal demand density – impact on the design results, in terms of cost and CO2 emissions. These two parameters were used (in addition to three others) for classification of zones by EIFER in WP2, to generate classes of zones. The parameters were also used to identify a representative ('characteristic') zone for each class and to identify 'quartile zones' with respect to each parameter, i.e. zones with a thermal demand around the top and bottom quartile value, and zones with thermal demand density around the top and bottom quartile value. Since the design inputs impact on the design results, the results can be used to answer questions about distinctiveness of classes and homogeneity within each class.

Design Inputs	Design Performance
1. Thermal density	• DHN annualised capital cost*
2. Annual thermal demand**	• Capital cost of the DE centre • Operating (fuel and maintenance) cost of the DE centre • Capital cost of the DHN • Fuel CO_2 emissions from the DE centre

Table 5-6: Classification parameters (design inputs) impact on design performance
*** The DHN cost is not an integral part of the optimisation.**
****The peak thermal demand, which is strongly correlated with the annual thermal demand, affects the size of DE solutions.**

5.5.2 Assessing distinctiveness of classes: Design results

Table 5-5 presents the performance indicators used to assess the distinctiveness of classes identified in WP2. These design inputs are represented in Table G-1. Figure 5-3 and Figure 5-4 present results for the characteristic zones. In section 5.5.3, the homogeneity of each class is assessed by considering how similar the performance is for the five zones in each class. Section 5.5.4 presents a quantitative assessment of homogeneity.

The capital cost of distributing heat produced in the DE centre to the end-users is shown in Figure 5-3. The highly simplified cost algorithm presented in Table 5-3 was used to calculate the cost of the district heating network. It may be seen that the district heating network dominates the costs. The twenty zones have DHN costs that range from around £1 million/year to around £25 million/year as heat demand varies from 23 GWh to 480 GWh. That is, there are significant differences between DHN costs for different characteristic zones. In Figure 5-4 the annualised cost of producing and distributing heat is shown to increase with thermal demand, as expected.

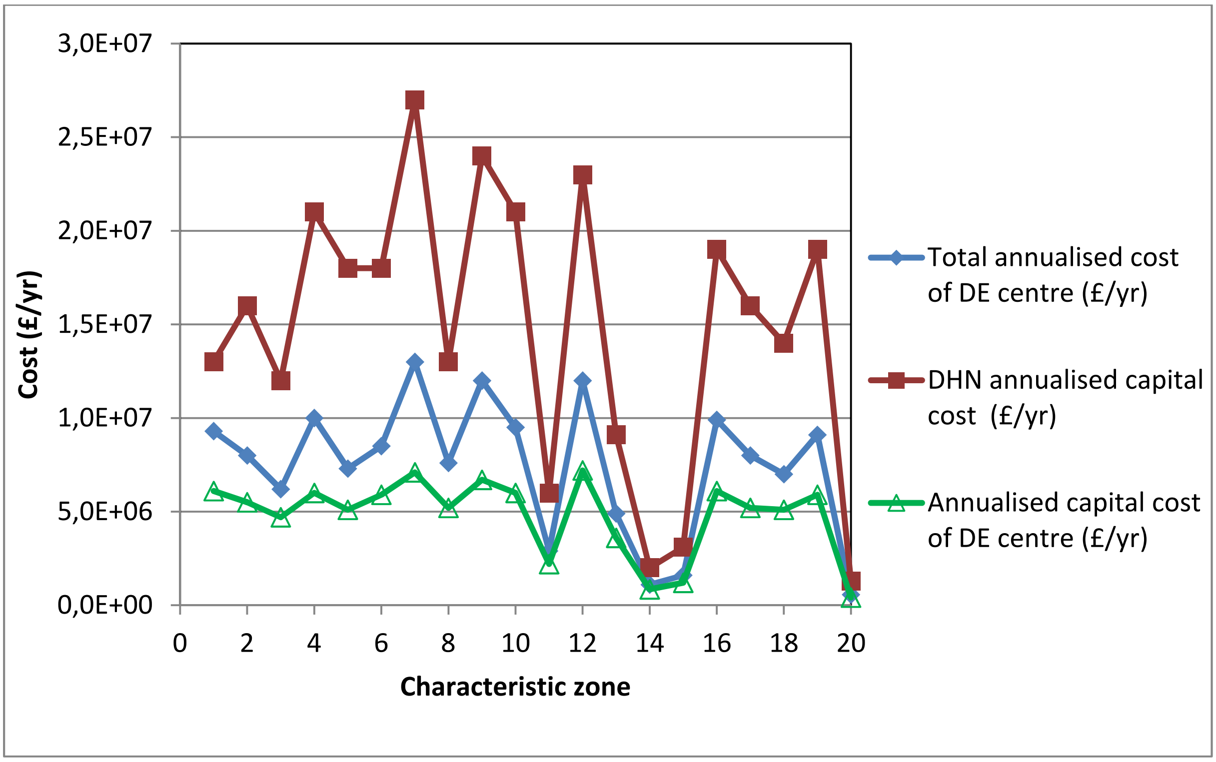


Figure 5-3: Annualised costs of DE centre and district heating network for characteristic zones

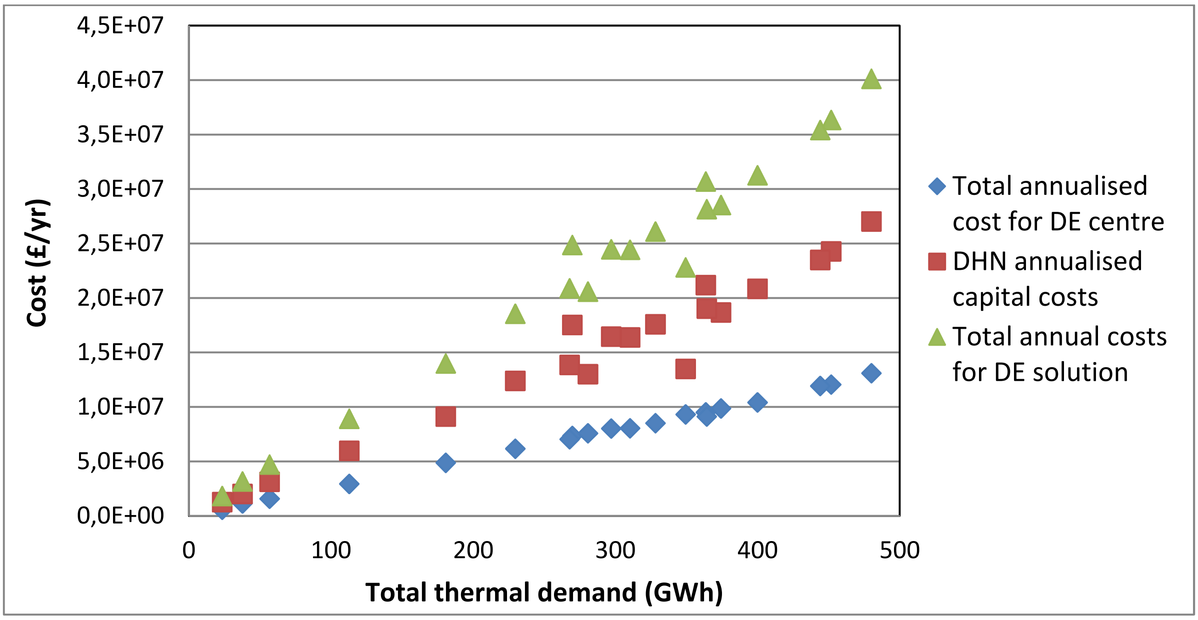


Figure 5-4: Annualised costs of DE centre and district heating network of characteristic zones with respect to thermal demand

Figure 5-3 and Figure 5-4 also depict the total annualised cost of the DE centre that produces heat to satisfy the demand of zones. These costs include the annualised capital cost of DE systems in the energy centre, revenue from electricity sales, fuel costs and maintenance costs. Class 7, with the highest cost, corresponds to the class with the maximum annual thermal demand, as can be seen in Table G-1 and Table G-2. Again, there is a wide variety in the performance of the classes represented in Figure 5-3 and Figure 5-4, which provides further evidence of the distinctiveness of the classes. Figure 5-5 shows the breakdown of the costs comprising the DE centre total annualised costs. It may be observed that the revenue from electricity sales is of similar magnitude to the operating costs, which implies that the electricity sales tariff is a very significant variable in the design.

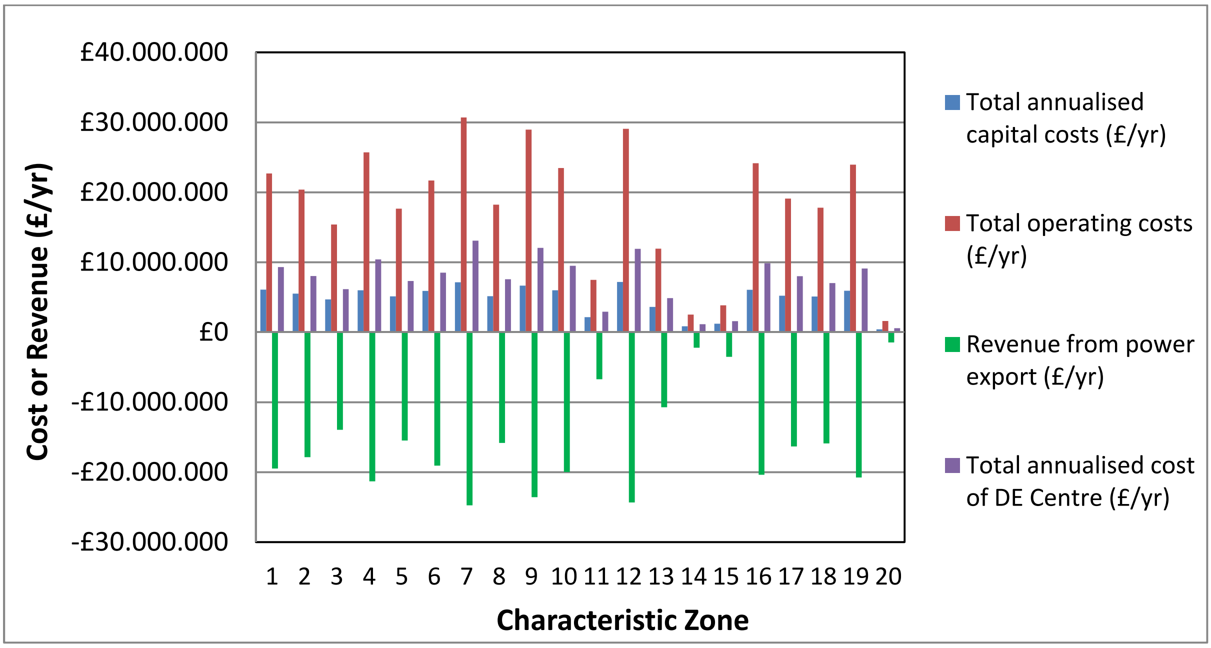


Figure 5-5: Contributions to total annualised cost of DE centre (annualised capital cost of DE systems, revenue from electricity sales, operating costs are fuel costs and maintenance costs)

The annualised capital cost of the DE systems selected to satisfy the thermal demand is shown in Figure 5-3 and Figure 5-4. The type, size, and number of DE systems selected and the operating schedule of the DE systems are an output of the design tool (

Figure 5-1) and impact on the capital cost of the DE systems selected. As expected, characteristic zone 20, which has the lowest total demand, has the lowest capital cost of the DE systems. The annualised capital cost of the district heating network for characteristic zone 7, which has the highest heat demand, is £27 million/yr, while the corresponding annualised capital cost of the DE centre is £13 million/yr; these are the highest costs over all classes.

Other performance indicators consider CO2 emissions. Figure 5-6 presents the 'global' CO2 emissions i.e. those from fuel combustion, less the credit applied for exported electricity that allows emissions from a central power station to be avoided and in Figure 5-7 the global CO2 emissions are plotted against the thermal demand of the characteristic zones. Note that the emissions figures are very low, around 0.03 t/MWh, reflecting the significant benefit of avoiding high-carbon electricity generation in central power stations. The grid emissions factor of 0.485 t/MWh (see Table 5-2) used to estimate the avoided emissions is expected to decrease in the future, as the grid is increasingly decarbonised.

In conclusion, Figure 5-2 to Figure 5-7 show that the characteristic zones of the 20 classes do perform differently, supporting the idea that zones generated in WP2 are distinctive. It may also be observed that a number of expected trends are followed, e.g. high costs are associated with high demand.

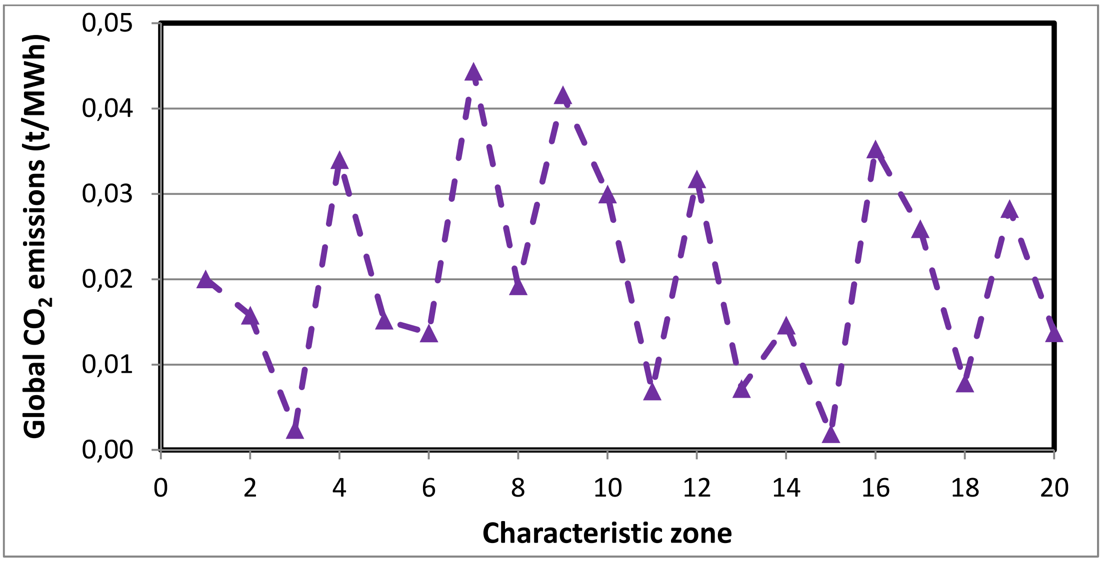


Figure 5-6: Global CO_2 emissions – fuel emissions less credit for avoided emissions

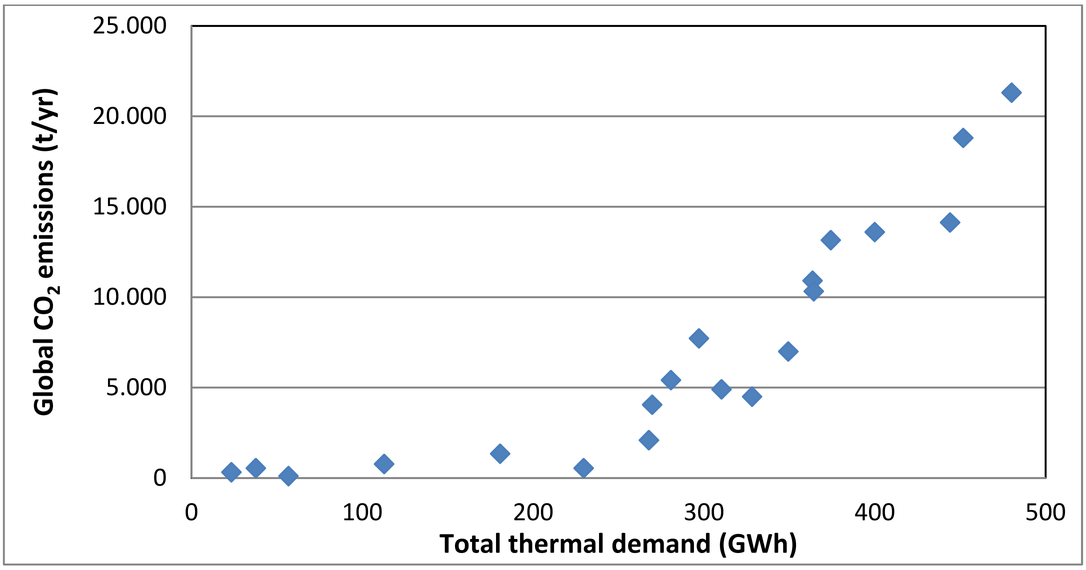


Figure 5-7: Global CO_2 emissions of characteristic zones with respect to annual thermal demand

5.5.3 Assessing homogeneity of classes: Design results

Establishing a case for Great Britain involves extrapolating results for 20 characteristic zones, representing 20 classes, to determine the performance for each class, and then summing up the behaviour over all the classes to represent the potential for DE in Great Britain. Therefore it is important that the zones in a class are similar to each other if this extrapolation is to be sound.

The homogeneity of the classes generated is assessed graphically, by considering the performance of the characteristic zone and the four 'quartile zones' of each class using the performance indicators presented in Table 5-5. Figure 5-8 to Figure 5-16 plot these indicators against annual thermal demand. Appendix H — presents performance results per unit of heat delivered.

Section 5.5.3 presents a quantitative analysis, where the difference between the performance (per unit of heat delivered) of the characteristic zone and the least similar quartile zone of that class is reported. Results that are more than 20% different to each other are highlighted, to indicate lower confidence in the homogeneity of the class.

Figure 5-8 presents the total annualised cost of the DE centre (i.e. annualised capital cost of DE systems, revenue from electricity sales, fuel costs and maintenance costs) for five zones in each class. As shown in Figure 5-8, the total cost of heat production ranges between £0.5 million/yr and £13.3 million/yr. Costs for the five zones in each class can be seen to be similar. For example in class 18 the range for cost is £6.6 million/yr to £7.3 million/yr, while thermal demand ranges between 250 GWh and 270 GWh. Figure H-1, which presents the total annualised cost of the DE centre per unit of heat delivered, supports the observation that costs are similar within each class.

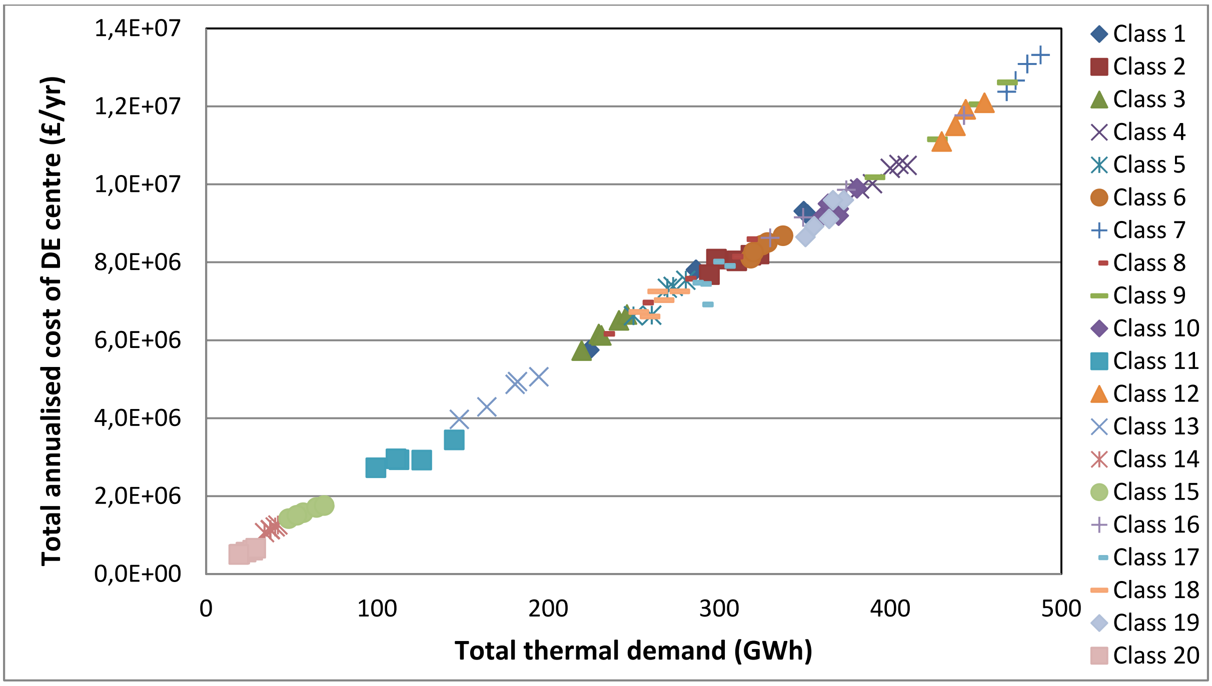


Figure 5-8: Total annualised cost of DE centre of five zones in each class

Figure 5-9 presents the annualised capital cost of the DE centre, i.e. capital investment for heat generation and Figure H-2 presents these costs per unit of heat delivered. The similar costs within each class reflect the similar DE solutions generated, with gas engines, gas boilers and thermal storage units providing the DE solution. Capital cost is the most significant contributor to total annualised cost of the DE centre, since revenue from electricity sales offsets the cost of fuel and maintenance significantly, as shown in Figure 5-15.

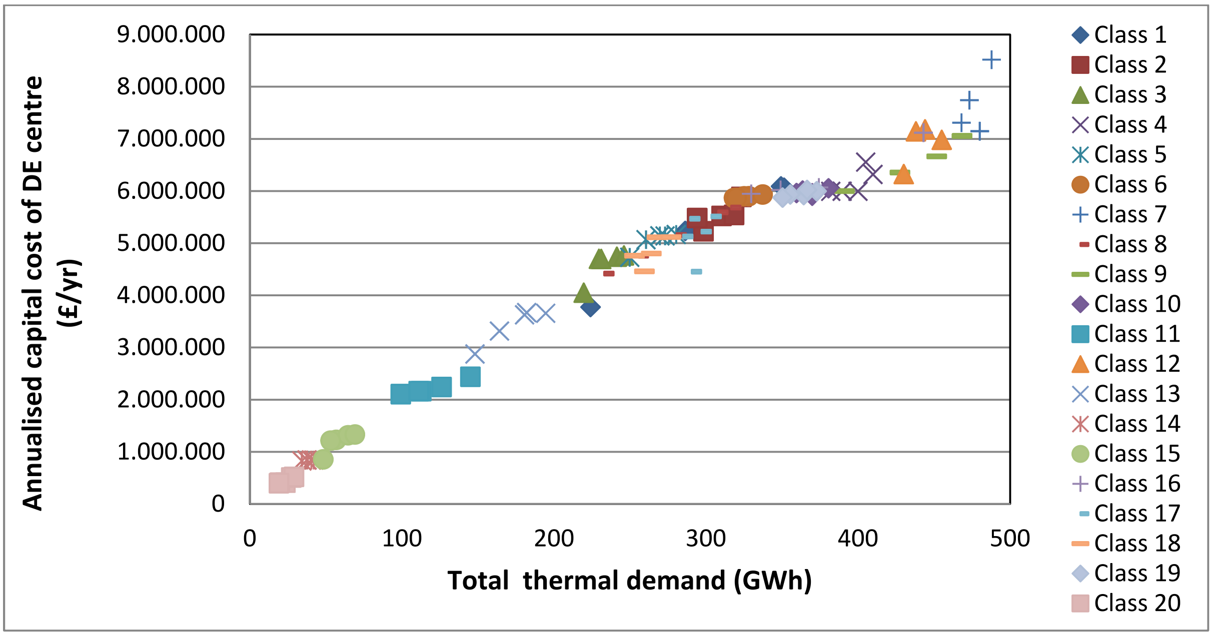


Figure 5-9: Annualised capital cost of DE centre of five zones in each class

Figure 5-10 presents the cost of the DHN plotted against the annual thermal demand; wide variation in costs is evident in certain classes (e.g. 1, 2, 5, 8, 10 and 19). Figure H-3 presents these results per unit of heat delivered. Note that costing of the DHN is not an integral part of the design optimisation. The cost model presented in Table 5-3 is highly simplistic; the costs it predicts depend strongly on whether the thermal density of a zone lies above or below 0.35 GWh/ha and whether the demand intensity (average heat demand per gas meter) is above or below 0.04 GWh/meter. In some classes, as shown in Table G-3, all five zones being analysed are of the same type. In the cases that different density/intensity zone types exist, greater scatter is observed. Figure H-4 and Appendix I — present further details related to the DHN cost model as applied to the five zones in each class.

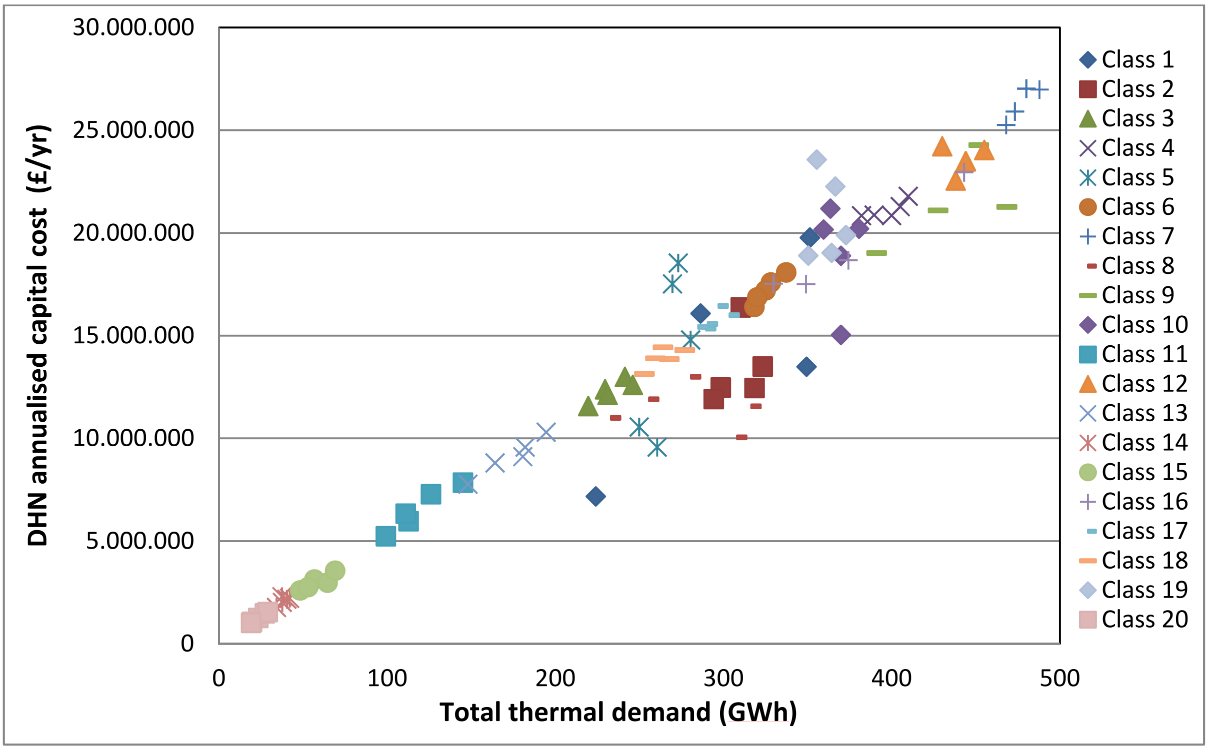


Figure 5-10: DHN annualised capital cost of five zones in each class

Figure 5-11 presents the annualised capital cost of the district heating network recalculated by taking all zones in a class to be of the same 'type' (low or high density and low or high intensity), i.e. costs are recalculated for zones 1, 2, 5, 8, 10 and 19. Comparing Figure 5-10 and Figure 5-11, it may be seen that the discrete nature of the DHN cost model presented in Table 5-3 is responsible for some of the variation in costs. Further investigations of the differences in cost for different zones in a class are presented in Appendix I —; it may be seen that in some classes the number of tertiary gas meters (i.e. heat users) and the number of domestic gas meters have some non-homogeneity, affecting the cost of heat interface units (HIU) in the DHN.

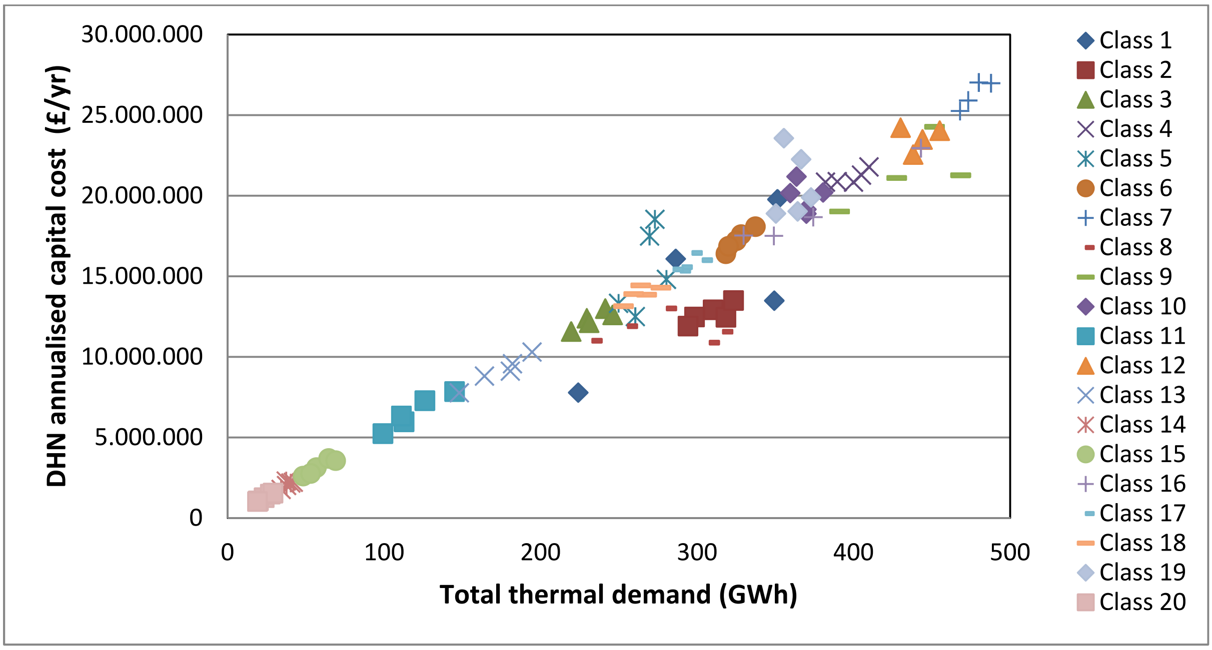


Figure 5-11: DHN annualised capital cost – recalculated taking all zones in a class to be of the same type'

Figure 5-12 plots the total annualised cost of the DE solution (DE centre and DHN) of five zones in each class against the annual thermal demand. The total annualised cost of heat production and generation indicates the minimum sales price for heat that would be economic, around £0.08/kWh – see also Figure H-5 in Appendix H —.

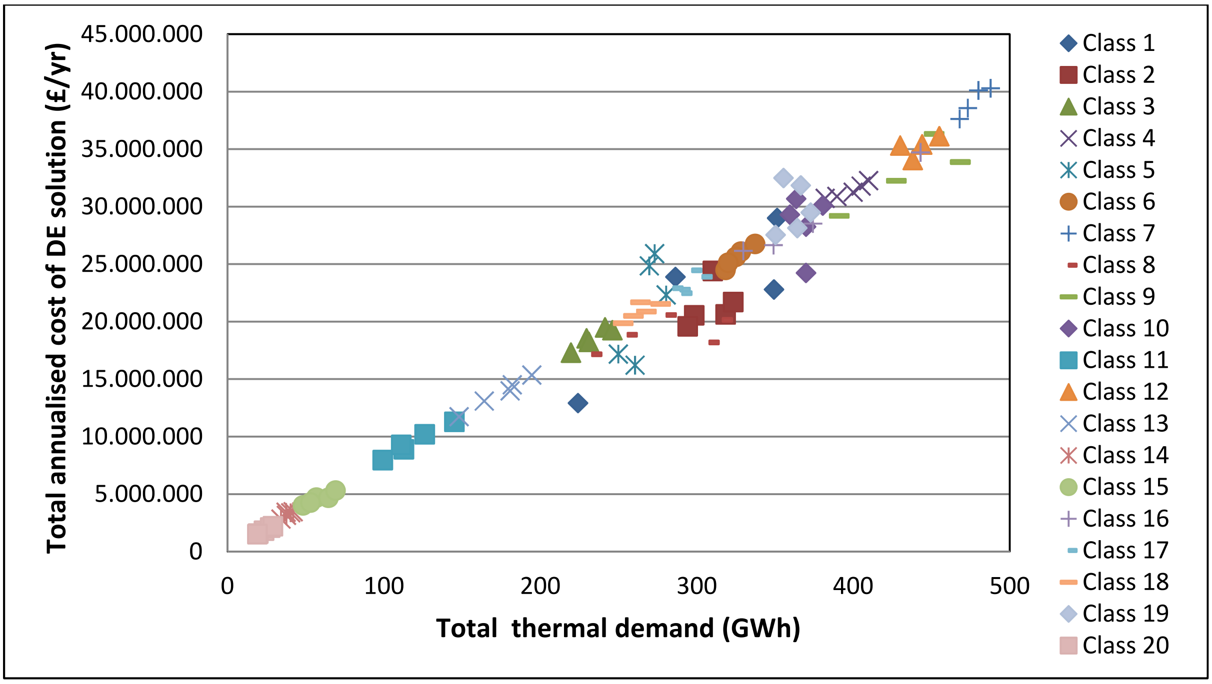


Figure 5-12: Total annualised cost of DE solution (DE centre and DHN) for five zones in each class

Figure 5-13 presents CO_2 emissions from fuel burnt to satisfy the demand of the zone. As the fuel that is selected for all DE solutions is natural gas, it is not surprising that very similar values, per unit of heat delivered, are obtained for all zones in each class – see also Figure H-6.

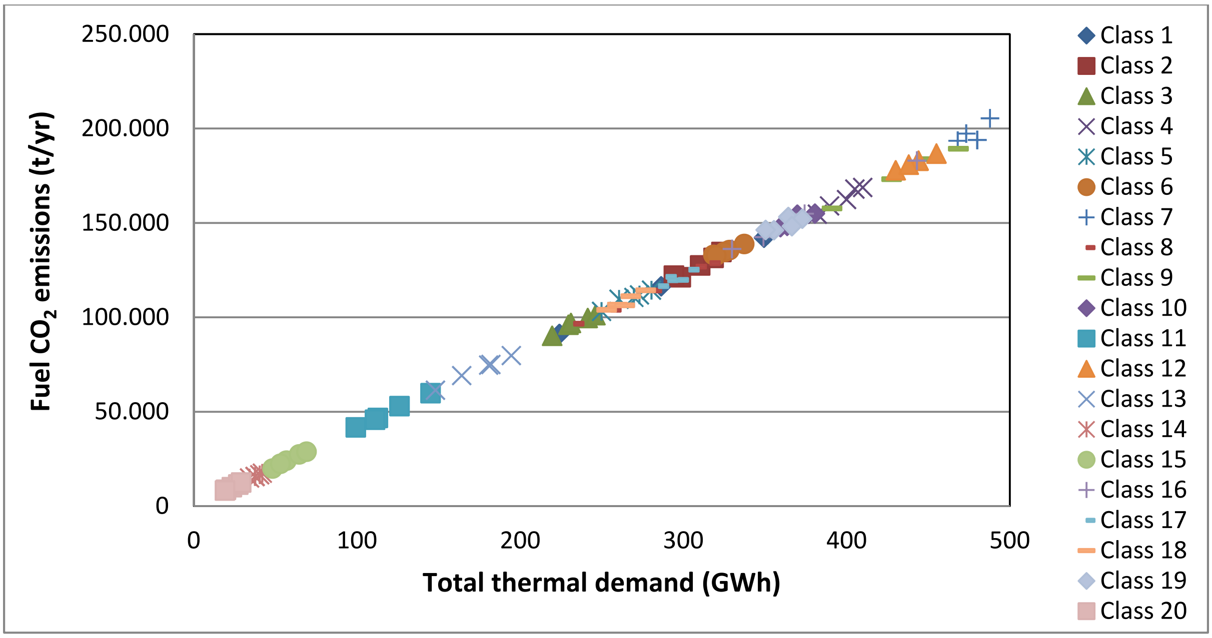


Figure 5-13: Fuel CO_2 emissions

Figure 5-14 presents avoided CO2 emissions. Where the average amount of electricity generated per unit of heat delivered is similar and the types of technologies applied for electricity generation are similar, it would be expected that the avoided emissions (per unit of heat delivered) would be similar – see also Figure H-7.

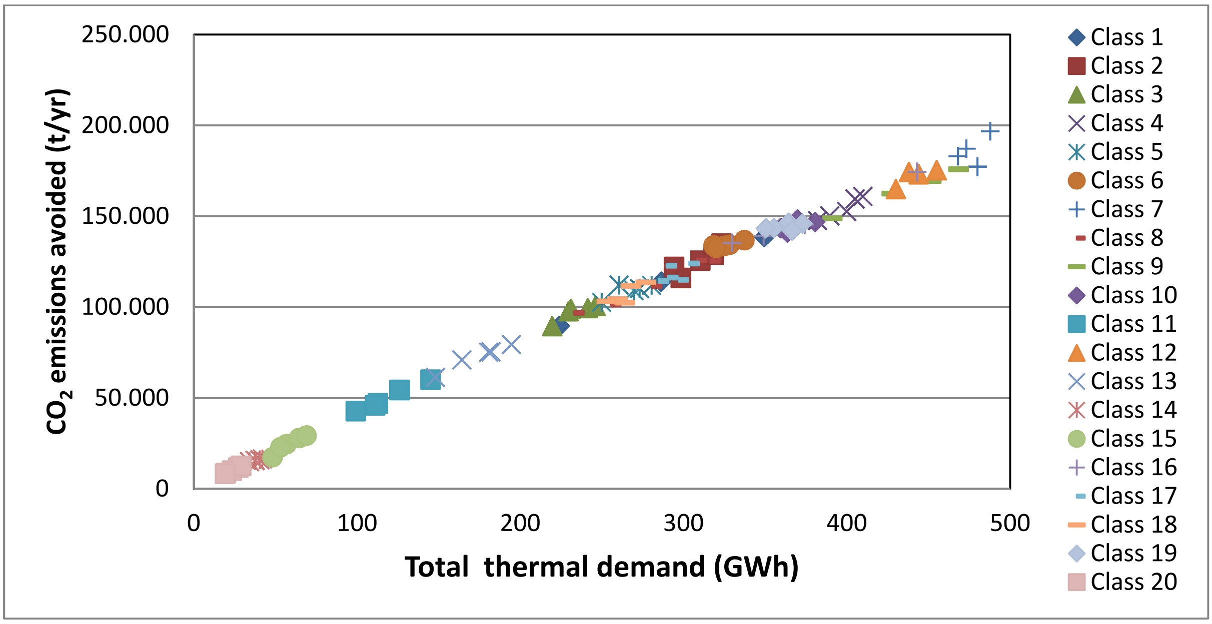


Figure 5-14: CO_2 emissions avoided for five zones in each class

Figure 5-15 presents the 'global' emissions (fuel emissions less avoided emissions). Note that the amount of CO2 generated is very similar to the amount of CO2 avoided, so the net amount of CO2 emitted by an energy centre is typically very low. The apparent scatter is a result of the small differences between fuel CO2 emissions and avoided CO2 emissions. The global CO2 emissions are around 0.03 t/MWh, as illustrated in Figure H-8.

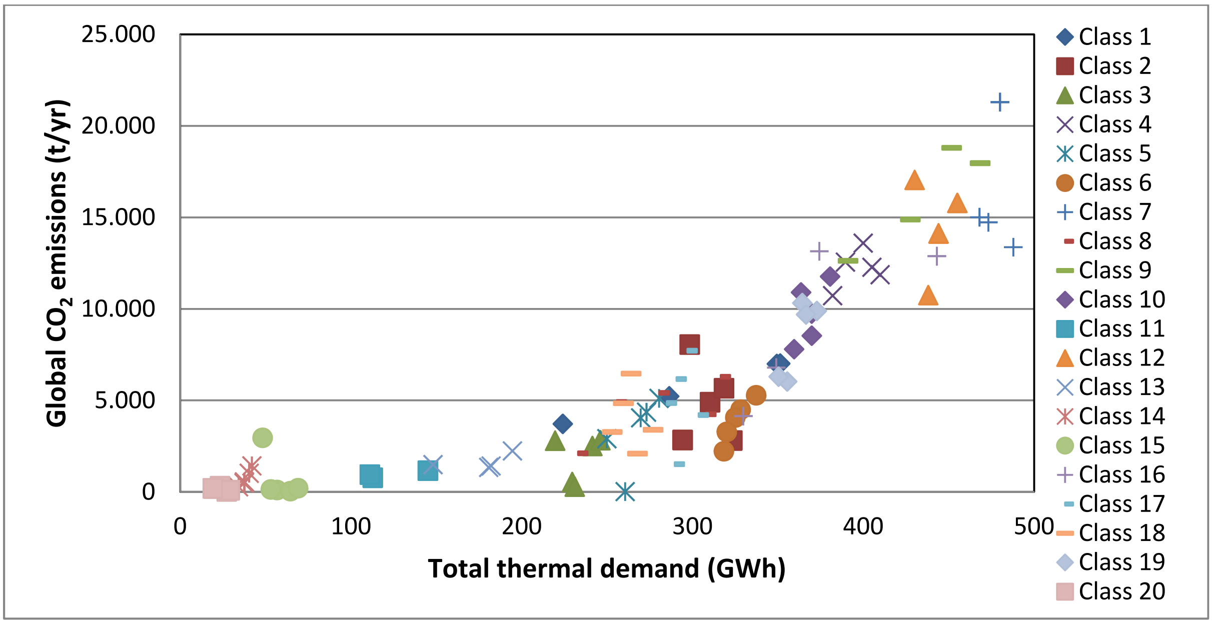


Figure 5-15: Global CO_2 emissions for five zones in each class

Figure 5-16 represents the proportion of the heat demand that is met by CHP units (as opposed to heat-only units such as boilers). Clearly, in all zones, the optimised operating schedule and optimised design makes extensive use of CHP units. CHP units provide economic solutions – in spite of their relatively high capital costs, they have advantages over boilers because of the significant revenue from electricity generation.

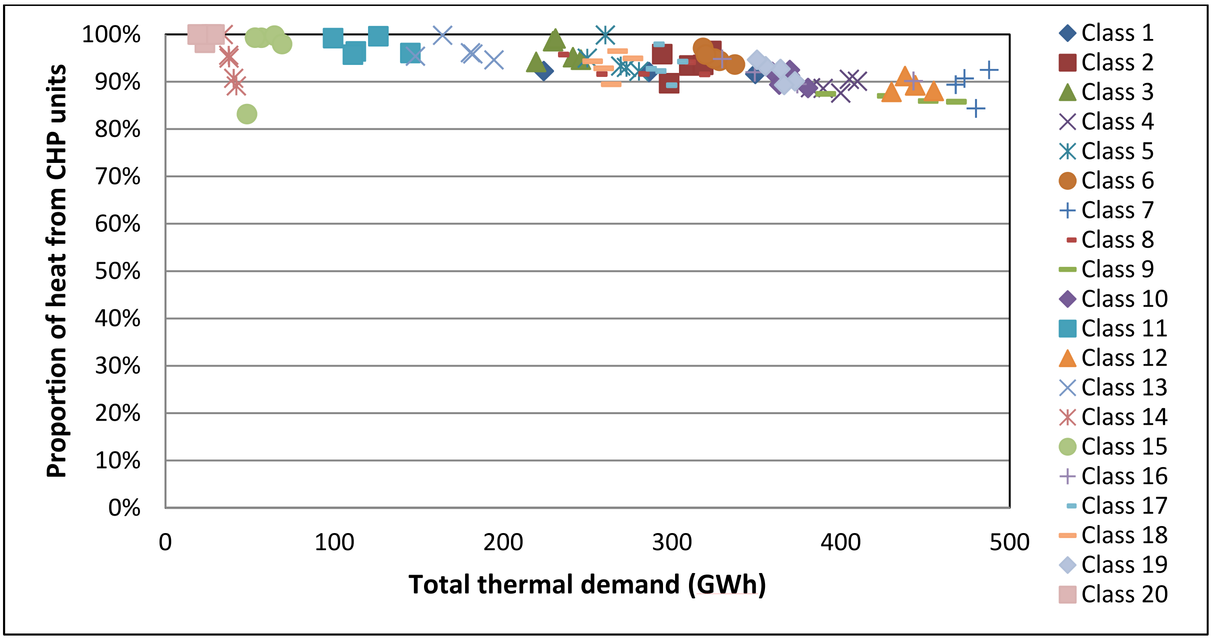


Figure 5-16: Proportion of heat delivered that is produced by CHP units

5.5.4 Quantitative assessment of homogeneity: design results

Section 5.5.3 represents design results for zones in each class graphically. Homogeneity is assessed quantitatively by taking the maximum difference between the performance of the characteristic zone and that of the four quartile zones. The performance is expressed per unit of heat delivered, as this value is relevant when extrapolating results for the class based on the performance of the characteristic zone. For example:

$$\begin{aligned} &performance\ for\ any\ class\ i \\ &\qquad = (performance\ per\ unit\ of\ heat\ delivered\ for\ CZi) \\ &\qquad * (total\ thermal\ demand\ for\ class\ i) \end{aligned}$$

Equation 5-1: Performance of classes

In Appendix H —, graphs present the performance per unit of heat delivered basis of five zones in each class, as presented in Table 5-5. In Table 5-7 the homogeneity of each class is quantified; differences of more than 20% are highlighted, indicating less homogeneous classes.

Class no.	Total annualised cost of DE centre per unit heat delivered (£/kWh)	DHN annualised capital cost per unit heat delivered (£/kWh)	Total annualised cost of DE solution per unit heat delivered (£/kWh)	Annualised capital cost of DE centre per unit heat delivered (£/kWh)	Fuel CO_2 emission per unit heat delivered (t/MWh)	CO_2 emissions avoided per unit of heat delivered (t/MWh)	Proportion of heat produced by CHP units (%)
1	4%	**46%**	**28%**	5%	1%	1%	1%
2	5%	**26%**	18%	5%	1%	4%	4%
3	3%	5%	3%	10%	2%	4%	4%
4	2%	5%	3%	8%	2%	3%	3%
5	6%	**43%**	**32%**	3%	3%	6%	7%
6	2%	4%	3%	2%	1%	3%	3%
7	3%	4%	4%	17%	4%	9%	10%
8	2%	**30%**	19%	3%	2%	5%	4%
9	2%	16%	10%	4%	1%	2%	2%
10	5%	**30%**	**22%**	4%	2%	3%	3%
11	11%	9%	6%	12%	2%	3%	3%
12	4%	6%	3%	9%	0%	2%	2%
13	3%	6%	3%	6%	2%	4%	4%
14	2%	14%	9%	11%	4%	6%	6%
15	8%	16%	12%	18%	3%	17%	16%
16	1%	7%	4%	11%	1%	6%	6%
17	12%	5%	6%	13%	4%	9%	10%
18	5%	6%	5%	10%	3%	7%	7%
19	5%	**27%**	18%	3%	4%	3%	4%
20	9%	1%	3%	19%	3%	3%	2%

Table 5-7: Quantitative analysis of scatter in classes (Values above 20% are highlighted)

5.6 Conclusions

In conclusion, these homogeneity studies indicate that the 20 classes generated are indeed significantly homogeneous. It is also clear that the classes are not all similar to each other – i.e. there is a satisfactory distinctiveness between the classes. The analysis presented in this report builds confidence in the approach used to classify zones in WP2. The results also support the use of the proposed approach, to carry out design and analysis for the characteristic zone in each of a limited number of classes and extrapolate those results to all zones in the class in order to assess the potential for macro-scale distributed energy in GB. The results represented above confirm that the performance of the DE centre, in terms of cost and CO2 emissions per unit of heat delivered, is similar for zones within a class.

However, it is clear that the model being used to estimate the cost of the district heating network is overly simplistic; it is suspected that its simplicity is introducing spurious scatter. A more sophisticated and realistic model for estimating the DHN cost is under development in the Macro DE project.

Chapter 6 — Methodology to assess industrial waste heat and its potential for district heating

6.1 Introduction

The broad objective of the MacroDE project is to assess the opportunity for providing low carbon distributed energy (DE) solutions to aggregated energy demand zones across the UK. One way to increase DE potential is to make use of heat sources, for example coming from industrial processes that are currently being lost. Consequently, the work described in this section is aimed at assessing the industrial waste heat[32] resources and their locations across GB.

The report includes the description of the methodology developed to select industrial sectors and evaluate waste heat potential as well as the confidence level in the results. The deliverable also includes an excel table providing results of waste heat potential per industrial sector and per site location (postcode), in terms of quality (temperature) and quantity (energy). A map indicating location and size of heat source is also provided.

6.2 Description of the methodology

6.2.1 A methodology in 4 steps

The approach adopted to calculate the industrial waste heat (or industrial excess heat which will be used synonymously) is composed of 4 steps. The procedure used in step 1 is largely described in (McKenna 2009), but step 2, 3 and 4 will be described in this document.

- **Step 1: Determination of energy consumption per industrial site**
  Since, there are no public data of energy consumption per industrial site, a procedure based on the use of site-specific data contained in the EU-ETS National Allocation Plan (DEFRA 2010), which provide annual CO2 emissions per committed industrial site, was adopted.

- **Step 2: Determination of waste heat ratios**
  The work determines for each industrial sector the ratio of waste heat to input fuel energy, in other words, the fraction of the primary energy consumed by a site that is lost from the site as waste heat. It is based on previous analysis made by University of Bath (McKenna 2009), complemented by EDF-R&D feedback from audits. When several different industrial processes are existing within a given sector (i.e. blast furnace or basic oxygen furnace for the iron and steel sector) several ratios can be provided. The ratios are expressed with "lower" and "upper" bands, to take into account the large discrepancy between industrial sites.

- **Step 3: Determination of waste heat potential per industrial site**
  Application of waste heat ratio (determined in step 2) to energy consumption (determined in step 1) provide the waste heat potential per industrial sector (UK-SIC codes).

[32] "waste heat" refers here to "excess of heat", i.e. heat that should be valorised out of the industrial site (on-site use or recovery is out of scope for this study).

- **Step 4: Visualisation**
  The locations of each industrial site with a potential for usable waste heat (and their sector reference) have been marked on a UK map. A picture of this map is included in this report and added in a higher resolution to this deliverable.

6.2.2 Data sources

Due to the large number of industrial sectors and their variety in terms of processes and energy consumption, and also because of the short period of the study, it was not possible to make measurement campaigns or surveys. Literature results were used as much as possible, but the five main references on the subject are reported below.

For **step 1**, total energy consumption per sector were obtained with the following four main sources:

- Digest of UK Energy Statistics (DUKES) (DECC 2009)
- Energy Consumption in UK (ECUK) (DECC 2010)
- EU Emissions Trading Scheme database (DEFRA 2010)
- McKenna, R.C., "Industrial Energy Efficiency – Interdisciplinary perspectives on the thermodynamic, technical and economical constraints (McKenna 2009)

For **step 2**, ratios of waste heat potentials were obtained using the following two main sources:

- McKenna, R.C., "Industrial Energy Efficiency – Interdisciplinary perspectives on the thermodynamic, technical and economical constraints (McKenna 2009)
- Waste Heat Recovery for Industrial Facilities, Electric Power Research Institute, Report No. 1020134 (EPRI 2010).

6.2.3 Assumptions

Sector coverage

A report by Element Energy (ElementEnergy 2010), for the UK Climate Change Committee (CCC) estimated direct emissions (i.e. CO2 emissions arising from combustion of fuels on site) from industrial sources at 125 MtCO2 in 2008, which represented about 23% of total UK emissions (531 MtCO2). It estimates that about 71% of industrial emissions (89 MtCO2) are already covered by the EU Emissions Trading Scheme (EU ETS). The 36 MtCO2 emitted from non-EU ETS activities represents a large number of typically small and diverse sources unsuited for capture, predominantly in the food and drink, chemicals, rubber and plastics, engineering and textiles sectors.

Figure 6-1 shows industry emissions for 2008 broken down by sector type, with emissions from industrial CHP facilities split out as a separate source. It can be seen that just four sectors (iron and steel, refineries, chemicals production, cement and food and drink) account for about 60% of all emissions. Iron and steel was the largest contributor with over 20 MtCO2 (primarily arising from fuel combustion in blast furnaces in large integrated iron and steel plants). These sectors are followed by fuel combustion in (predominantly gas-fired) CHP facilities and process and fuel emissions from cement kilns. Emissions from the remaining were distributed across a wide range of sectors and processes including food and drink, textiles, lime, pulp and paper, and other sectors. The category indicated as "other" accounted for around 20 MtCO2 in 2008 and represents a large number of small and medium-sized industrial and manufacturing sites inside and outside of the EU ETS including e.g. vehicle and machinery production, electrical engineering, mining and quarrying, gypsum and plasterboard production and tires manufacture.

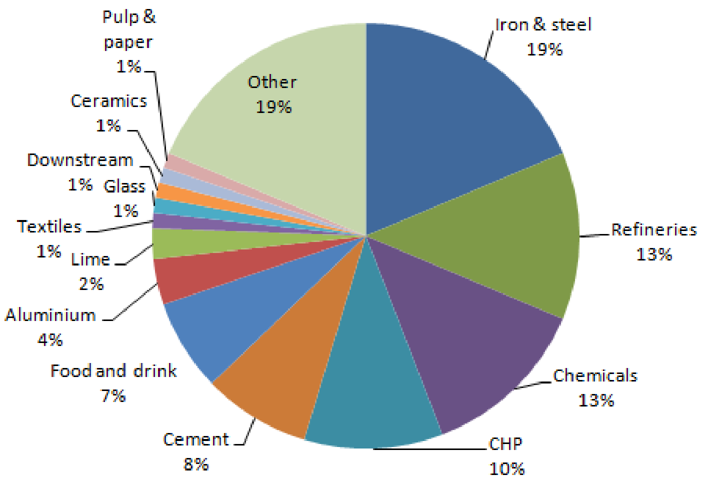


Figure 6-1 : Industry direct CO2 emissions in UK (2008)

Quality of heat

Although the quantity of waste heat available is an important parameter, it is not alone an effective measure of waste heat recovery opportunity. Indeed, it is important to specify the waste heat quality, which means the usefulness of the waste heat, as determined by its temperature.

Waste heat streams are often grouped into 3 temperature categories to distinguish how these streams might be used, with different energy recovery technologies: low (< 100°C), medium (100-500°C), and high (> 500°C) (see examples in Table 6-1).

At high temperature waste heat streams can be used for a wide range of industrial plant needs, including the production of process steam with electric power (i.e. CHP bottoming cycles), combustion air preheating, and many low temperature applications (e.g. space heating or water heating) (see Figure 6-2).

In the framework of our study, since we want to obtain value of possible supply of heat outside of the plant (i.e. with Heat networks which needs heat around 80°C), it seems reasonable to consider heat with a minimum of 100°C. Below this temperature, the waste heat may still have a value but heat pumps will need to be used to recover waste heat.

In addition, most publications like DOE (DoE 2008), provide a reference temperature of 100 – 130 °C for an efficient utilisation of heat outside of the industrial process.

Consequently, we can consider **100°C as reference temperature for heat recovery**.

Temperature Classification	Waste Heat Source	Use for Recovered Energy	Technologies

High
 	
(> 500 °C)
	• Furnaces
	  • Steel heating
	  • Basic oxygen
	  • Glass melting
	• Coke ovens
	• Fume incinerators
		• Combustion air preheat
		• Process steam
		• Power generation
		• Furnace load preheating
		• Medium or low temperature process needs
			• Passive heat exchangers
			  • Recuperators, regenerators
			  • Air preheaters
			  • Regenerative/recuperative burners
			  • Waste heat boilers
			• CHP – steam driven

Medium
 	
(100 – 500 °C)
	• Combustion exhaust streams (Steam boiler , Gas turbine, IC engine)
	• Heat treating furnaces
	• Ovens (Drying, Baking)
	• Cement kilns
		• Combustion air preheat
		• Process steam
		• Power generation
		• Feedwater preheating
		• Low temperature process needs
			• Passive heat exchangers
			• CHP
			  • Steam cycle
			  • Organic Rankine cycle (ORC)
			  • Kalina cycle?

Low
 	
(< 100 °C)
	• Combustion products from recovery systems
	  • Gas fired boilers
	  • Ethylene furnaces
	• Steam condensate
	• Cooling Water
	  • Furnace
	  • Air compressors
	  • Refrigeration condensers
	• Ovens (Drying, Baking)
	• Hot process liquids or solids
		• Space heating
		• Domestic water heating
		• Low temperature process needs
			• Heat pump (increase temperature to useful range)
			• ORC
			• Kalina cycle?

Table 6-1 : Example Waste Heat Streams Classified by Temperature
Source: Adapted from DOE (DoE 2008).

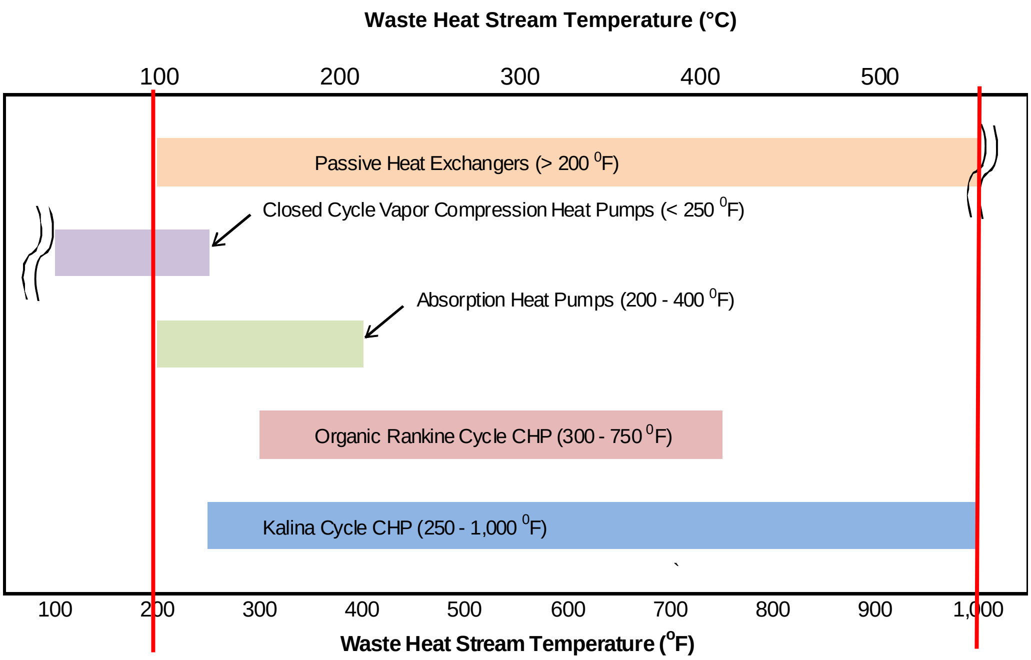


Figure 6-2 : Potential waste heat recovery technologies depending on heat quality (EPRI 2010) (red lines indicate the low and high boundaries of this study)

Quantity of heat: potential coverage

The quantity of waste heat contained in a waste stream is a function of both the temperature and the mass flow rate of the stream:

$$E = m \times \Delta H(T)$$

Equation 6-1: Calculation of available energy as waste heat

Where *E* is the waste heat (kW); *m* is the waste stream mass flow rate (kg/s); and Δ*H(T)* is the waste stream specific enthalpy (kJ/kg) as a function of temperature, relative to a reference state.

Enthalpy is not an absolute term, but it must be measured against a reference state. In this report, the enthalpy of waste heat streams is calculated at room temperature and atmospheric pressure ambient.

According to EPRI data (see (EPRI 2010)), the top 5 waste heat sectors (petroleum and coal products, chemical manufacturing, paper, primary metals and nonmetallic minerals) account for almost 90% of all waste heat potential (see Table 6-2).

	< 150°C	**150 < < 230°C**	**230 < < 650°C**	**650°C <**
Food & drink	1	3	3	0
Wood	8	8	1	0
Pulp & paper	**11**	**17**	**2**	**0**
Petroleum	**12**	**18**	**72**	**13**
Chemicals	**29**	**15**	**12**	**52**
Non-metallic minerals	**3**	**8**	**2**	**0**
Iron & steel	**23**	**30**	**5**	**4**
Others (13 sectors)	13	0	2	30
	100	100	100	100

Table 6-2 : Waste heat potential (% of total waste heat at a given heat quality) per sector
Source: Adapted from DOE (DoE 2008).

Those sectors are all in the scope of the study, with the exception of oil refineries (or petroleum as

Table 6-2) because these are considered to be too complex for inclusion. Indeed, basic processes used in petroleum refineries include distillation processes (fractionation), thermal cracking processes, catalytic processes, and treatment processes. Although these processes use large amounts of energy, modern refineries use heat produced in exothermic reactions for heating other processes, resulting in integrated heat recovery systems for process use. Many exhaust streams still contain high-quality waste heat that could be recovered for power production, but need to be evaluated on a case by case study.

Figure 6-3 illustrates that the majority of heat recovery is estimated to be within the temperature band from 100°C to 500°C.

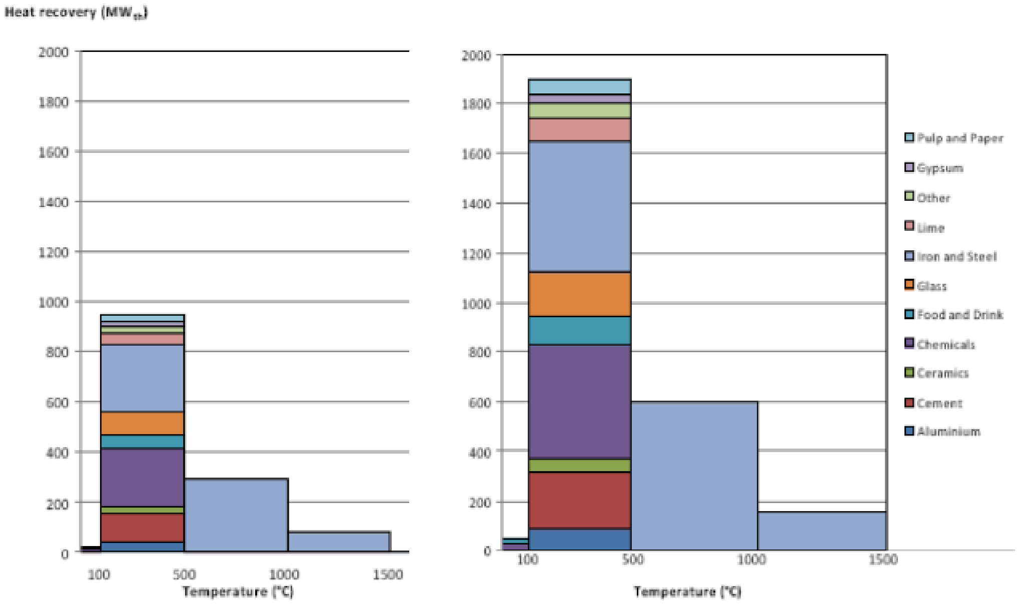


Figure 6-3 : Low (left) and high (right) heat recovery potential by sector against temperature (see (McKenna 2009)

6.3 Energy consumption per industrial site

In total, more than 260 industrial sites are considered, each site being characterized by:

- Company name
- Postal address (including postal code)
- UK-SIC code (2 digits), which refers to the sector reference
- Annual Energy consumption (electricity and per fuel type : coal, gas, oil, coke) in MWh or PJ (10^{15} J)

The following sectors are included in the report :

- Cement
- Lime
- Glass
- Iron & steel
- Aluminium
- Pulp & Paper
- Chemicals
- Ceramics
- Food & drink (breweries, malteries, sugar cane, sugar beets)

6.4 Calculation of waste heat ratios per sector

The same approach was taken for Cement, Glass, Pulp & Paper, Ceramics and Food & drink, while for aluminium, iron and steel, chemicals, and lime specific methodologies were adopted either due to their heterogeneous nature and/or because parts thereof have not been included in the National Allocation Plan (DEFRA 2010). For these last four sectors the method employed involved information about production capacities for individual sites and products, in conjunction with specific energy consumptions for these processes[33].

6.4.1 Cement

The cement sector is fully included within the EU-ETS.

Four routes are currently employed in the UK for the manufacture of cement, namely wet, semi-wet, semi-dry and dry. The recent tendency has been towards the dry process because it is less energy intensive, but the process employed depends largely on the nature of the raw materials. About 55% of industry capacity is based on the dry process, with another 30% from the wet process. The remainder is either semi-dry or semi-wet.

The common process of cement manufacture involves calcining of calcium carbonate (from limestone) at a temperature of around 1000°C to produce calcium oxide. The clinkering process then occurs at around 1500°C, when the calcium oxide reacts with silica, alumina and iron oxide to form the silicates, aluminates and ferrites of calcium which comprise clinker. The clinker is then grounded and blended with gypsum and other additives to produce saleable cement. In principle the wet and dry processes are identical, but the degree of moisture content in the wet materials can necessitate additional technology such as chains within the kiln to break up lumps, as well as additional energy input.

Most cement plants have some kind of preheater and/or calciner for the raw materials, which use heat from the back end of the kiln. A preheater simply raises the temperature of the raw materials prior to them entering the kiln, whereas a precalciner actually starts and carries out the majority of the calcination process before the kiln. Hence there are usually two exhausts from a cement plant: from the precalciner or preheater at a temperature of around 200-300°C and from the cooler at around 300-400°C (EIPPCB 2010). In some cases the useful heat in either or both of these streams may already be utilised, but it is difficult to estimate the utilisation.

The analysed sources suggest that final exhaust temperatures in efficient plants can be as low as 150°C (see (McKenna 2009)) but can go up to 300°C in older plants (see (EPRI 2010)). The lower figure (i.e. T=150°C) has been considered in this study, as confirmed by EDF-R&D experts.

The range for the exhaust heat as a fraction of the total heat input was taken as 10% to 20% (see (McKenna 2009; EPRI 2010), confirmed by EDF-R&D experts). These data are reported in Table 6-3.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Cement	150	10	20

Table 6-3 : Exhaust temperature and heat recovery potentials for cement industry

[33] Specific energy consumption data was mostly obtained from the relevant sector BREFs (Bat REFerence document – BAT : Best Available Technologies) and EU ETS Benchmarking studies – See EIPPCB (2010). Final Draft BAT Guidance Note on Best Available Techniques for the Initial Melting and Production of Iron & Steel Sector, European Integrated Pollution Prevention and Control Bureau (EIPPCB).

6.4.2 Lime

The lime sector is covered by the EU-ETS as far as the production capacity exceeds 50 tons per day.

The lime sector has been further broken down according to the particular type of technology employed at a given site. All British Sugars plants use a so-called mixed feed shaft kilns(Entec 2006).

Lime production involves calcining calcium and/or magnesium carbonates in the temperature range of 900-1500°C but sometimes higher (EIPPCB 2010). The reaction forms the respective oxide (i.e. CaO or MgO) and liberates carbon dioxide.

Given the similarity of the process to that of cement manufacture, the assumptions are the same as for the cement sector, as summarized in Table 6-4.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Lime	300	10	20

Table 6-4 : Exhaust temperature and heat recovery potentials for lime industry

6.4.3 Glass

All large scale glass manufacturing plants with a capacity over 20t/day are included within the EU-ETS. This generally covers all large-scale producers of glass but excludes artisans and manufacturers of specialty glass products. In total, 23 glass furnaces installations have been identified in UK.

The glass industry uses raw material melting furnaces, annealing ovens, and tempering furnaces, all operated at high temperatures. Modern glass factories use regenerative furnaces to maintain high energy efficiency. In addition, electric boosting is used increasingly on furnaces to improve efficiency and yield, and oxy-fuel firing[34] reduces energy usage and increases efficiency.

For regenerative furnaces, the exhaust gases in the stack are at a temperature between 450°C (DoE 2008)and 550°C (see (McKenna 2009) and confirmed by EDF-R&D experts). For older plants, using recuperative furnaces, the exhaust temperature ranges up to 1000°C. In order to be conservative a lower value of 500°C will be considered in the study.

Since energy in exhaust gases represents around 30% of the total heat input, a range of 10-20% can be adopted for the fraction of heat lost in the exhaust. This is confirmed by an averaged value of 15% provided by EPRI (see (EPRI 2010)). As an example, an EDF audit in France provided a potential waste heat recovery of 8 MW_{th} over 47 MW input, which count for 17% as confirmed by EDF-R&D experts.

Data for the glass sector are summarized in Table 6-5.

[34] Oxy-fuel combustion is the process of burning a fuel using pure oxygen instead of air as the primary oxidant. Since the nitrogen component of air is not heated, fuel consumption is reduced, and higher flame temperatures are possible

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Glass	500	10	20

Table 6-5 : Exhaust temperature and heat recovery potentials for glass industry

6.4.4 Iron and Steel

Iron and steel plants producing primary or secondary steel at a rate of more than 2.5 tonnes per hour are included in the EU ETS. This includes 14 sites in the UK, three of which are large integrated steelworks, seven are electric arc furnaces (EAFs) producing secondary steel, and four are other processing (coating, tin plating and strip mills) facilities owned and operated by Tata (former Corus) (see (Entec 2006)).

Steel production in the UK is concentrated in the blast furnace/basic oxygen furnace route (for primary steel) and electric arc furnace route (for secondary steel).

The three integrated iron and steel works currently in operation in the UK are Teeside, Scunthorpe and Port Talbot respectively. Details of these sites along with the production capacities used in this study are presented in (McKenna 2009). But the crux of any integrated plant is the blast furnace, which is by far the largest energy consumer (see (EIPPCB 2010)).

Steel mills have a number of high temperature heat recovery opportunities. In integrated mills, waste heat can be recovered from coke ovens, blast furnaces for iron production, and basic oxygen furnaces for steel production. There are also opportunities to recover waste heat from the electric arc furnace in steel "mini-mills" that produce steel largely from recycled scrap.

This section examines waste heat recovery opportunities from coke ovens, blast furnaces, basic oxygen furnaces that are major energy consuming parts of integrated steel production, and at electric arc furnaces that are the major energy consuming process at steel mini-mills.

Coke ovens

Coke ovens produce coke from coal, for use in the blast furnace as a reducing agent and fuel.

Within the oven coal is heated for several hours or days to produce coke through pyrolysis. The coal is loaded into the oven and heated by burning gas through flues in the walls. The vast majority of coke production is in slot ovens, which evolved to collect the by-products and manufacture coke oven gas. The temperature required within the coke ovens is around 1100°C.

According to (ElementEnergy 2010) and (Bisio and Rubatto 2000),waste heat could be recovered from:

- the incandescent coke at 1100°C accounting for 43-60% of the thermal energy output
- the hot gas exiting the coke ovens at 650-800°C (known as Coke Oven Gas : COG);
- the waste combustion gas at around 200°C after some heat has already been recovered

Sintering

Sinter plants produce the fine powder of iron ore for injection into the blast furnace.

The process involves heating the blended raw materials (including fine ores, additives, recycled materials from downstream operations, etc.) on a travelling grate to temperatures in the region of 1400°C. Cooling occurs at the end of the grate and may be integrated or, more common, separate to the strand. Separate coolers layer the calcined sinter up to about 1m thickness in a large rotating structure and cool with large volumes of air, which are forced upwards through the layer. Some heat content of the air may then be recovered in a waste heat boiler, or used to preheat the raw materials. The assumption for this study is that the heat from the sinter cooler is not presently recovered, and that this presents a considerable opportunity for heat recovery. Therefore a heat temperature of 1350°C is considered as well as a high recovery potential (50-100%) (Table 6-6).

Blast Furnace

The blast furnace (BF) is the vessel within which iron ore is reduced by coke at high temperatures to yield pig iron. The furnace is a tall structure with a temperature profile along its vertical height. It is charged from the top with burden, including sinter, coke, flux agents (lime), and blast furnace gas (BFG) is drawn off for use as an on-site fuel. A hot air blast is blown in through tubes known as tuyeres lower down in the belly of the furnace. This reacts with the reducing agents to form mainly carbon monoxide, which in turn reacts with the iron oxide to form iron. Molten iron and slag are tapped off from the hearth at the base of the furnace. The temperature within the blast furnace ranges from 1500°C at the top to in excess of 2000°C in the belly, and the molten iron leaves the furnace at around 1500° C.

Overall, blast furnaces are very efficient with losses representing less than 10% of the energy input (see (IEA 2007)). Areas for potential heat recovery include the Blast Furnace Gas (BFG), the molten iron and the slag. The latter is ruled out here because the technology is only at the prototype stage and development has been halted since the 1980s. The sensible heat of the molten iron cannot feasibly be utilised as the molten metal is transferred directly to the basic oxygen furnace. Any heat recovered from this stream would have to be transferred back into the metal before the BOF.

Hence the only technical heat recovery potential for blast furnace streams is assumed to be the BFG. It is estimated that, using a top pressure recovery turbine, primary energy savings of 0.3GJ/t carbon steel can be achieved (see (Worrell, Price et al. 2001)). A value of 150°C was also provided by Corus as lower temperature of BFG and 250°C is provided by DOE (see (DoE 2008)).

Another opportunity for waste heat recovery is from the combustion exhaust gases leaving hot blast stoves. The gases are at temperatures of approximately 250°C (see (DoE 2008)). The blast stove exhaust gas is relatively clean and is more compatible with heat recovery devices, making heat recovery from blast stoves a more common practice. The heat can be used to preheat combustion air and/or fuel gas. A conservative value of 150°C is assumed (see Table 6-6).

Basic Oxygen Furnace

The basic oxygen furnace (BOF) converts pig iron into steel by adding oxygen to remove the carbon, as well as small amounts of silicon, manganese and phosphorous. Molten iron and steel scrap are charged into the furnace and pure oxygen is blown in through a liquid-cooled lance. The most common technology is the Linz-Donowitz (LD) converter, which is a pear-shaped vessel into which the lance is lowered.

The temperature inside the furnace is around 1700°C, and after secondary metallurgical processes (such as homogenising), the molten steel is transferred to a casting ladle at around 1600-1800°C. The gases produced during oxygen blowing (BOS gas) exit the furnace at the same temperature as the molten steel, and are therefore a suitable target for heat recovery.

A potential method for doing this is by using a heat recovery boiler to raise steam (Worrell et al., 2001). The temperature of the heat to recover is considered to be 1600°C (see Table 6-6).

Continuous casting

Continuous casting is the state of the art method for casting steel. It is a continuous process that replaced its predecessor, batch-wise casting in moulds before reheating for rolling. The molten steel falls from the casting ladle or tundish at first under gravity and then supported by rollers from the casting ladle, gradually reaching the horizontal. Thus a continuous strand of material is formed, which is then cut with a torch cutter. The main potential for heat recovery at this stage is from the solidified steel when it is at a temperature of around 800°C, with a radiant heat recovery boiler.

Electric Arc Furnaces

Secondary steel is produced through electric arc furnaces (EAFs) in the UK. Within the furnace, which is charged with scrap before being electrically heated in a batch-wise process, the temperatures reach around 1600°C. The off-gas is typically used to preheat the scrap before melting. Recent improvements of EAFs made them reach almost the theoretical value of specific energy consumption. Hence the exhaust gases from EAFs are not considered a likely candidate for heat recovery, and the only potential for this study is assumed to be through utilising some of the heat in the steel in the continuous casting and rolling stages, as described above.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Coke oven gas	1000	10	25
Coke oven waste gas	200	50	100
Sintering	1350	50	100
Blast furnace	150	15	30
Basic Oxygen Furnace	1600	10	20
Electric Arc Furnace	200/1200	10	20
Continuous casting	980	25	50
Hot rolling	900	30	60

Table 6-6 : Exhaust temperature and heat recovery potentials for iron & steel industry (% refer to waste energy / primary energy input of the site)

6.4.5 Aluminium

The aluminium sector is considered in the EU-ETS only through electricity generation activities. This includes the 420MW coal-fired power station at Lynemouth, which is owned and operated by Alcan, providing electricity to its Lynemouth smelter.

Hence the three aluminium smelters in the UK have been included as additional (the operational data for these sites has been gathered from relevant company publications). These are the two Alcan smelters at Lynemouth and Lochaber, and the Anglesey Aluminium smelter at Holyhead.

Aluminium smelting within the UK only uses the prebaked anodes route rather than the Soderberg route which uses anode paste. Aluminium manufacture involves the electrolysis of alumina (aluminium oxide), obtained from bauxite, whilst dissolved in a bath of sodium aluminium fluoride (cryolite) at a temperature of around 1000°C. The main energy use on site is in the form of electricity for the electrolysis, but there are also other significant on-site energy demands for ancillary activities such as casting. The exhaust gases from the process are drawn off the reduction cells, filtered to conform to environmental legislation and then released to the atmosphere at a temperature of around 100°C (see (NineSigma 2007)). Overall around half of the input energy is lost as heat, and 30% of this is in the enthalpy of the off-gas (see (NineSigma 2007)).

The estimated heat recovery potential for the sector has been determined based on this assumption that 15% of the final energy input is lost in the exhaust gas at 100°C. Hence the exhaust fraction range has been set at 5-10%.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Aluminium	100	5	10

Table 6-7 : Exhaust temperature and heat recovery potentials for Aluminium industry

6.4.6 Pulp and paper

Of the pulp and paper sector the EU ETS covers all pulping of timber or other fibrous materials and the production facilities for paper and board with capacities exceeding 20 tonnes per day. In total 64 installations are included in the Phase II National Allocation Plan, with 19 of these being mainly or wholly CHP units. The remainder of the sites produce paper and tissues of various qualities, with the energy intensity of production being roughly proportional to the paper quality.

For non-CHP related installations the heat load is assumed to be derived from boilers and steam systems. The scope for improvements in energy efficiency within this sector is estimated to be significant with future technologies, and with new drying methods the net specific heat consumption of paper production could be reduced to almost zero (see (Beer 1998)). With current technology, however, the scope for improvement through heat recovery is estimated around 10%, through improvements in drum and Yankee dryers (see (Energetics and E3M 2004)).

There are several suggested measures for improving overall efficiency and recovering wasted heat that, considering their applicability to the industry as a whole and the savings they each stand to achieve, could together achieve around 10%. Hence, as for several other sectors, the potential range for heat recovery is set at 5-10% for pulp and paper.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Pulp and paper	150	5	10

Table 6-8 : Exhaust temperature and heat recovery potentials for pulp & paper industry

6.4.7 Food and drink

The food and drink sector is only covered by the EU ETS in terms of thermal energy activities.

The food and drink sector has been broken down into homogeneous subsectors as far as possible, including breweries, maltings, distilleries, and sugar manufacturing plants – through both the beet and cane routes. Nevertheless, along with the chemicals sector, food and drink is one of the most heterogeneous sectors in terms of energy activities. A precise breakdown into independent sectors has therefore not been possible and the results are therefore expected to be less robust than for other sectors. The assumed heat recovery potentials are shown in Table 6-9.

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Breweries	150	5	10
Distilleries	80	5	10
Maltings	40	5	10
Sugar beet	200	5	10
Sugar cane	150	5	10

Table 6-9 : Exhaust temperature and heat recovery potentials for food and drink industries

6.4.8 Chemicals

The chemicals sector is only covered by the EU ETS in terms of its thermal operations, and then only for installations rated at 20MWth and over. The sector is highly diverse, with 19 separate subsectors having very different market drivers and priorities. In total, there are around 3500 companies, represented by some 20 trade associations. This diversity means that it has not been possible to model the sector’s activities accurately. This section describes the key processes involved in the manufacture of several key chemicals, and the methodology employed to estimate heat loads and recovery potentials. As outlined above this work attempts to identify the technical potential for process improvement. Results are summarized in Table 6-10.

Ammonia

Ammonia is the source of nearly all synthetic nitrogen fertilisers produced in the world. It is manufactured by combining nitrogen and hydrogen in the Haber process. Globally over 80% of ammonia is produced through the steam reforming of hydrocarbon feedstocks (natural gas, naptha, LPG, refinery gas). The production route in the UK is exclusively through the steam reforming of hydrocarbon feedstocks. The synthesis of ammonia, in which nitrogen and hydrogen are reacted over an iron catalyst, is actually exothermic. The energy-intensive part of ammonia production is in the manufacture of hydrogen, which typically occurs in a two-stage reforming process at temperatures up to around 1000°C.

There is a large degree of energy integration in modern ammonia plants, and much of the exhaust heat from the reformers and ammonia synthesis plants is recovered (see (EIPPCB 2007)). Notwithstanding this high degree of energy integration, it is estimated that 10% of waste heat from the reformers is recoverable (see (Energetics and E3M 2004)).

A value of 10% is therefore employed as an upper limit for the technical potential, with the lower estimate at 5%. The temperature of the exhaust from the secondary reformer after heat has been recovered in a waste heat boiler is around 350 °C (see (EIPPCB 2007).

Chlorine

Chlorine and sodium hydroxide are both used extensively within the chemicals industry, chlorine mainly to produce chlorinated compounds (e.g. vinyl chloride) and sodium hydroxide to produce various organic and inorganic chemical compounds with a variety of applications (including soaps and detergents).

The two are manufactured through the electrolysis of brine (sodium chloride solution), which is carried out in three types of cell: mercury, membrane and diaphragm. Only the first two of these technologies are employed in the UK, with the vast majority of UK capacity employing the most efficient of the three, membrane technology. The membrane process is more efficient because it uses less electricity, some of which is substituted for steam.

The temperature within the cells is around 70°C and 85°C for the mercury and membrane processes respectively (see (Brown, Hamel et al. 1985)). Hence the scope for heat recovery per se is limited or non-existent, but improvements in overall efficiency can be made by optimising the way in which heat and electricity are generated and supplied to this and adjacent processes.

Ethylene

Ethylene is the main raw material for the petrochemicals industry. It is manufactured through thermal or catalytic cracking of hydrocarbon feedstocks. Ethylene manufacture in the UK is carried out at four sites, where hydrocarbon feedstocks are thermally cracked. There is no ethylene production in the UK by catalytic cracking.

The basic process involves preheating the hydrocarbon feedstocks to around 650°C in the preheater section before mixing with steam and cracking at around 850°C (see (Worrell, Phylipsen et al. 2000)). The gas mixture is then rapidly quenched to 400°C to stop the reaction, producing high pressure steam, before water is injected to further lower the temperature. The liquid is then extracted as the gaseous fraction is fed to a fractional distillation column.

The SEC for steam cracking of ethane feedstocks is around 15-25GJ/t excluding feedstock energy. Steam crackers require large amounts of energy at a high temperature to promote disassociation of the chemical feedstock, but also employ cryogenic separation processes to purify and separate the products. Hence crackers have a large degree of energy integration, which is typically achieved by recovering as much as possible of the heat from the front end to use as work for separation. The complexity of steam crackers and the proprietary nature of the technology means that one temperature for the exhaust was not obtainable. It was therefore assumed that the exhaust temperature lies in the range 100-500°C. The estimated heat recovery from steam crackers is around 10% (see (Energetics and E3M 2004)). Hence this is employed as an upper limit to the technical potential with 5% as the lower bound.

Other chemicals

The three chemicals ammonia, chlorine and ethylene together amount to around 23-34% of total chemicals sector energy consumption in 2006. If we consider also titanium dioxide and soda ash production, about 40% of the chemical sector's total energy consumption in taken into account.

The potential for heat recovery from these two processes is thought to be 5-10% (see (Energetics and E3M 2004).

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Ammonia	350	5	10
Ethylene	100-150	5	10
Other major chemicals		5	10

Table 6-10 : Exhaust temperature and heat recovery potentials for main chemicals

6.4.9 Ceramics

The ceramics subsector is covered by the EU ETS. The vast majority of ceramics installations covered by the scheme are brick manufacturers. All ceramics installations included in the analysis are therefore treated as brick manufacturing sites, which are assumed to be using tunnel kilns, the state of the art for large scale brick manufacture.

The tunnel kiln has three main zones along its production line with a corresponding temperature profile: the preheat, firing and cooling zones (see (Entec 2006)). The temperature within the firing zone typically reaches around 1000°C, although this can be higher for roof tiles and facing bricks, and a uniform temperature profile is required across the wares. As the bricks are drawn through the kiln on cars, air is drawn through the kiln in the opposite direction. Most manufacturers take waste heat from the kiln and pass it to the dryer, but around 35-40% of the total heat input is exhausted up the chimney at a temperature of around 150°C. For this work, in order to reflect the technical constraints on heat recovery as well as the fact that not all sensible heat can be recovered, limits of 10% and 20% have been selected for the fraction of waste heat (see Table 6-11).

Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Ceramics	150°C	10	20

Table 6-11 : Exhaust temperature and heat recovery potentials for ceramics

6.5 Waste heat quantity and quality

6.5.1 Summary of data

Sector / Process	Exhaust Temperature (°C)	Low exhaust fraction (%)	High exhaust fraction (%)
Cement	150 – 300	10	20
Lime	150 – 300	10	20
Glass	500	10	20
Iron & steel Coke oven gas	1000-1100	10	25
Iron & steel Coke oven waste gas	200	50	100
Iron & steel Sintering	1350	50	100
Iron & steel Blast furnace	200	15	30
Iron & steel Basic Oxygen Furnace	1600	10	20
Iron & steel Electric Arc Furnace	1200	10	20
Iron & steel Continuous casting	980	25	50
Iron & steel Hot rolling	900	30	60
Aluminium	100	10	20
Pulp and paper	150	10	20
Food and drink Breweries	150	5	10

Food and drink Distilleries	80	5	10
Food and drink Maltings	40	5	10
Food and drink Sugar beet	200	5	10
Food and drink Sugar cane	150	5	10
Ammonia	350	5	10
Ethylene	150	5	10
Chlorine	100-150	5	10
Ceramics	150	10	20

Table 6-12: Summary of the waste heat potential of different industrial sectors

6.5.2 Validation of sector coverage

In order to estimate the share of the total industrial waste heat in the UK covered in this analysis, the results have been compared to the ECUK dataset (see (DECC 2010)), the most highly disaggregated data available on industrial energy use in the UK.

The estimated total energy use for each industrial sector is shown in Table 6-13. Also shown in Table 6-13 are the sector totals in primary energy from the model against the second period results of the Climate Change Agreements (CCAs), which ran from 2002-2004. The conversion to primary equivalents within the model has been carried out by multiplying electricity use by 2.6, which is the reciprocal of the overall grid efficiency of 38% used by DEFRA in converting electricity into primary energy terms. The CCA results could not be summed for other industry sectors because of differences in coverage and problems with allocation of activities.

Most sectors show a good agreement with the ECUK data. This is especially the case with energy-intensive sectors (i.e. Ceramics, lime, pulp and paper...). In a few specific cases the correspondence with ECUK is weaker (iron and steel), due to some allocation difficulties (i.e. processing of raw materials like coke is included in Iron and Steel in CCAS, not in our study). The primary energy results, however, show a very good correspondence with the CCA results with a coverage of about 80%.

	Primary energy (PJ)		Final energy (PJ)	
	CCAS	Our study	ECUK	Our study
Aluminium	63	56	25	26
Cement	56	53	61	44
Ceramics	23	24	23	19
Chemicals	279	230	190	203
Food and drink	127	104	143	83
Glass	38	31	55	29
Iron and steel	308	294	67	251
Lime	10	17	14	16
Pulp and paper	98	98	109	78
Other industry	~200	49	450	39
Total industry	**1200**	**956**	**1137**	**787**
Share (%)		**80**		**69**

Table 6-13 : Comparison CCAS and EU-ETS (our study) annual energy consumption coverage per industrial sector (final energy means end-use energy including electricity)

6.5.3 Waste heat potential

Results are gathered in a table (provided as attached file) summarising energy consumption and waste heat potential for each site in the database developed from the EU ETS covered sites. An extract is provided in Table 6-14.

						Energy Consumption (PJ)			Heat recovery				
Installation Address	Postcode - 7 digit	EU ETS Sector	SIC code - 3 digit	Subsector	SIC code - 4 digit	Fuel	Electricity	Total	Exhaust T°	Low heat recovery fraction	High heat recovery fraction	Waste heat (low) MW	Waste heat (high) MW
						PJ/yr	PJ/yr	PJ/yr	°C	%	%	MWth	MWth
Naylor Drainage Cawthorne Site South Yorkshire S75 4AD	S75 4AD	Ceramics	26.2	Ceramics _bricks	26.26	0.18	0.00	0.18	150	5%	10%	0.3	0.6
Carlton Main Brickworks Ltd - Grimethorpe Clay Burn Road South Yorkshire S72 7BG	S72 7BG	Ceramics	26.2	Ceramics _bricks	26.26	0.18	0.00	0.18	150	5%	10%	0.3	0.6

Table 6-14 : Example of waste heat potential table

6.5.4 Validation of waste heat coverage

When applying the ratio waste energy/final energy consumption to the total energy consumption per sector as shown in Table 6-13, the share of industrial waste heat treated in this study, as regard to total potential industrial waste heat across UK, can be obtained. Results are illustrated in Table 6-15. The study covers about 70% of total waste heat.

	Energy consumption (PJ)		Waste Heat	Waste heat potential (MWth)	
Industry sector	Our study	ECUK	ratio	Our study	ECUK
Aluminium	26	25	0,10	87	83
Cement	44	61	0,20	264	366
Ceramics	19	23	0,10	63	76
Chemicals	203	190	0,10	678	634
Glass	29	55	0,20	195	369
Iron and steel	251	67	0,13	1074	287
Lime	16	14	0,20	96	84
Pulp and paper	18	109	0,10	51	307
Pulp and paper - CHP	60		0,00	0	0
Food and drink	42	143	0,10	133	453
Food and drink - CHP	42		0,00	0	0
Other non minerals	3	450	0,20	20	1268
Others	27		0,08	61	
Gypsum	9		0,10	32	
'otal Energy consumptio	789	1137	Total Waste heat	2754	3930
			Waste heat covered in study (%)		70

Table 6-15 : Comparison DUKES (see (DECC 2009)) and EU-ETS (our study) on waste heat coverage per industrial sector

6.6 Visualisation of waste heat

All potential industrial sites in the UK are have geographical co-ordinates and the recoverable waste heat (both low and high scenario) are visualised using Geographic Information System. Afterwards, waste heat arising from the individual sites is aggregated at each MLSOA/IGZ, and thus provides a spatial representation of combined recoverable waste heat.

6.6.1 Methodology

Several steps are involved to visualise the recoverable waste heat. At first all the potential industrial sites in the UK are geo-localised. Afterwards, a geo database on the energy consumption and waste heat potential (similar to Table 6-15) is prepared. In the next step, those industrial sites are visualised by their exact geographic location. Then it has been possible to create maps according to the exhaust temperature or the recoverable heat at low and high recovery scenario.

In the final step, it is possible to identify which sites belong to which MLSOA/IGZ. Therefore, GIS operation (e.g. spatial join, dissolve operation, etc.) has been carried out to calculate aggregated waste heat for the industrial sites that belong to particular MLSOA/IGZs.

For this purpose, considering the reference temperature (100°C, as discussed in section 6.2.3) for heat recovery, characteristics of industrial sites (some sites e.g. food and drink, chemical are left out), as well as geographic location 245 sites are identified as potential source of waste heat. These sites are located across 201 MLSOAs/IGZs in the UK.

6.6.2 Results and maps

In the GB, total waste heat recovery potential varies between 1015 MWth (low scenario) and 2024 MWth (high scenario),. The geographic distribution of recoverable heat potential at various exhaust temperature and at low and high scenarios is shown in Figure 6-4 and Figure 6-5. The waste heat is aggregated at the MLSOAs. The maximum potential (at high scenario) exists in the MLSOA Redcar and Cleveland 003 in North East region and in MLSOA Neath Port Talbot 019 in Wales region, 360 and 339 MWth respectively. Both of these are sites with major steelworks. Figure 6-6 shows the top 20 MLSOA/IGZs where maximum waste heat recovery potential at both high and low scenario exists.

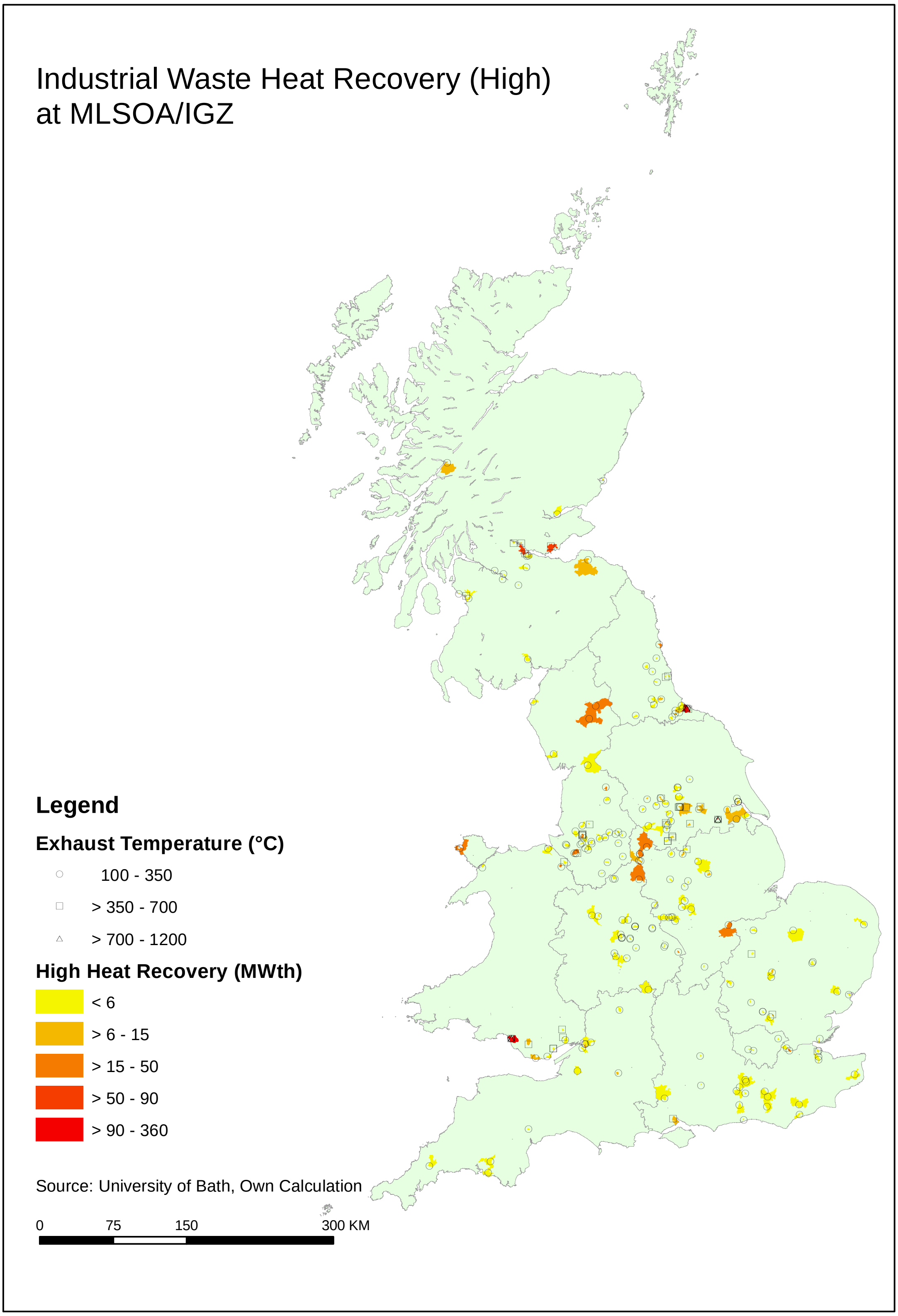


Figure 6-4: Maps presenting high heat recovery scenarios and corresponding waste heat potential across GB

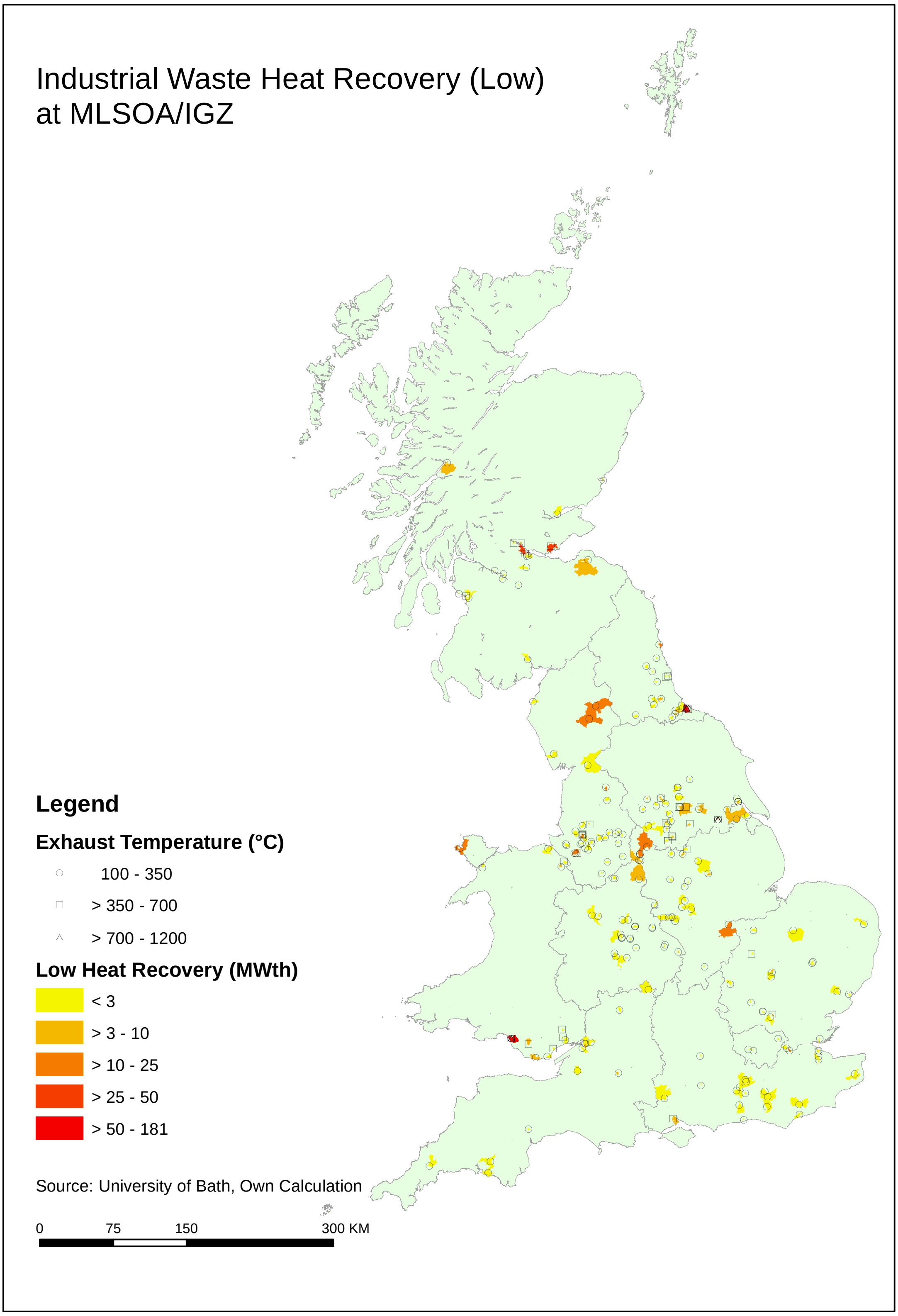


Figure 6-5: Maps presenting low heat recovery scenarios and corresponding waste heat potential across GB

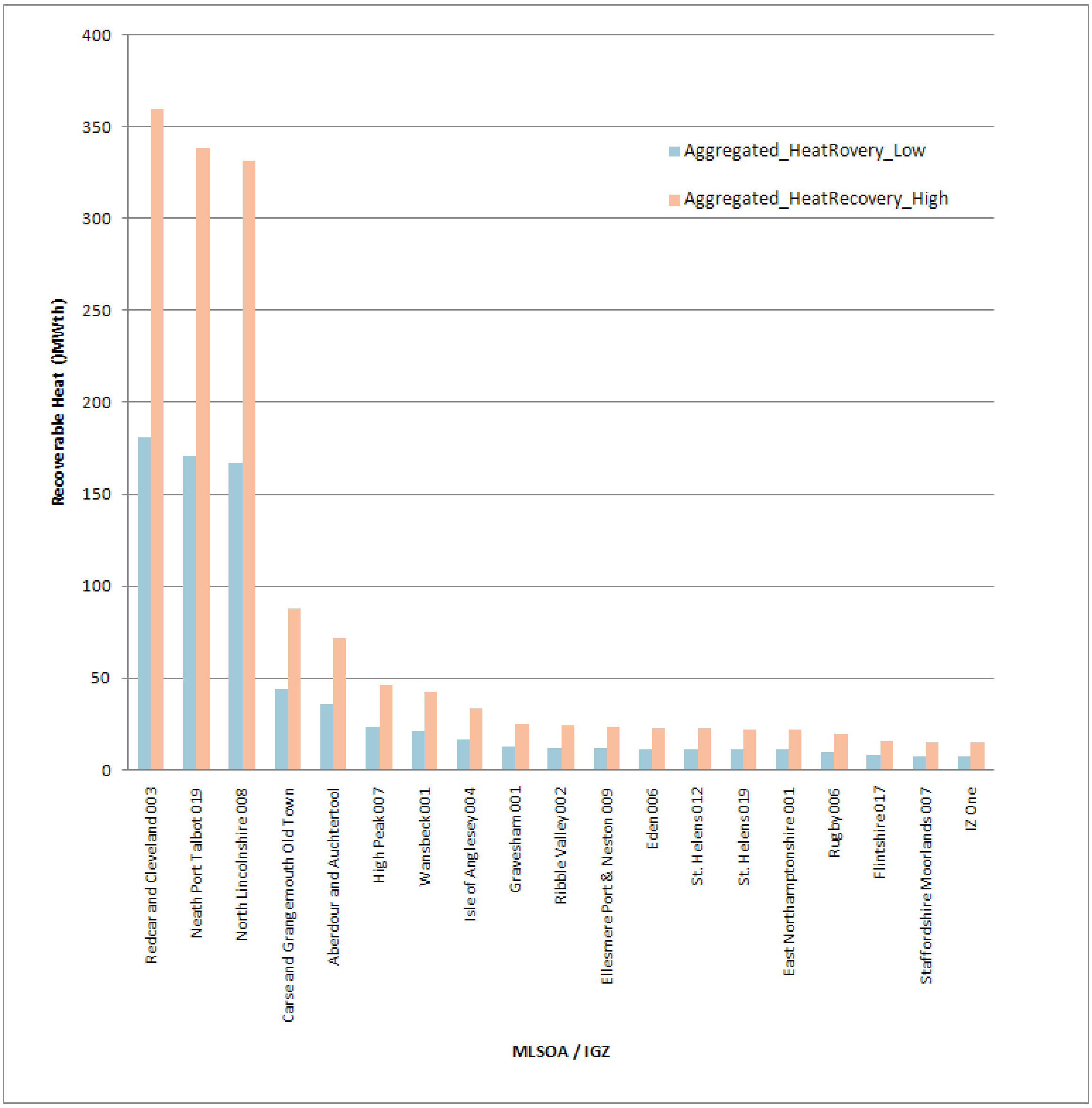


Figure 6-6: Aggregated recoverable heat (low and high scenario) at the top 20 MLSOA/IGZs in the GB

6.7 Conclusion and Recommendation

Excess heat from industrial sites is a valuable heat source for DE centres. This study shows that waste heat up to 2754 MWth can be used to supply the demand of district heating. Although this is only 1.1 % of the total heat demand within all DE zones, the analysis needs to be made on a case by case basis. Especially in the industrial dominated areas of GB (e.g. in Redcar and Cleveland, Neath Port Talbot or North Lincolnshire, see Figure 6-6) the potential for using waste heat and thus saving CO2 emissions can be considerably profitable for those areas concerned.

To ensure the availability of waste heat and to analyse the seasonal effects of industrial heat supply and domestic and tertiary heat demand further studies are necessary. As this study has been based on an annual heat demand figure it would be interesting to investigate the effects of temporal waste heat curves in comparison to the load curves (see Chapter 3 — Temporal Demand Calculation) within one DE zone.

Chapter 7 — Final Conclusion

Work Package 2 of the ETI Macro DE project investigates the energy demand of all MLSOAs in GB, creates macro-scale DE zones of up to 500 GWh where district heating is viable and clusters them into 20 classes. These classes have been tested for their homogeneity within a class and the heterogeneity between classes. Additionally, the temporal demand for the residential and tertiary sectors is calculated and a study about the potential of waste heat in GB conducted.

The total GB demand has been calculated for the residential and the tertiary sector, both electrical and heat demand. Indicators data, benchmarks and profiles have been collected on different levels, down to postcodes, to ensure the requested granularity. It can be concluded that the knowledge about the tertiary sector is lower than the residential sector. The reasons for this could be the diversity of the tertiary sectors, the lower heat demand, but most likely the fact that historically the tertiary sector has been of less interest. In conclusion, the total calculated GB heat demand of these two sectors is 458,392 GWh per year and the total electrical demand 198,747 GWh per year.

Of this demand nearly 50 % has been assessed as viable for district heating. These 4660 MLSOAs out of 8429 have been clustered into 948 DE zones. Although more than 50 % of the MLSOAs are included within the DE zones, they cover only about 4 % of the area of GB, which leads to the conclusion that macro-scale district heating can only be established in areas of high residential and tertiary energy demand, thus cities or other densely populated areas. A limitation of this study has been the size of MLSOAs, especially in rural areas where a lot green space like forests or fields are included in the MLSOA area. For these MLSOAs a further study, e.g. on the base of LLSOA or even smaller is recommended.

During the classification process the 948 created DE zones have been grouped into 20 classes. Out of each class, one DE zones has been chosen as representative for all DE zones of this class. These 20 Characteristic Zones will be used for further analysis within the ETI Macro DE project. The Characteristic Zones have to be analysed independently of their actual locations, because they are typical for their class in terms of energy demand, but not in terms of area or region.

The 20 classes have been tested on their distinction between each other and the homogeneity within each class. The first task – the distinction – can be shown by the difference of annual heat demands, energy densities or further parameters.

The second part of work package 2 was a study about the availability of waste heat in GB. It was shown that industrial sites can be an additional heat source for DE zones as the heat is already produced and is currently being lost. There is the need for further analysis to assess the temporal availability of the waste heat source and how surplus heat[35] can be damped, but this should be done on a site by site basis.

[35]Here, surplus heat means the exceeding of the heat demand by the waste heat supply.

References

BDEW (1999). Repräsentative VDEW-Lastprofile. VDEW Materialien. Frankfurt, German Association of Energy and Water Industries.

BDEW (2000). Anwendung der Repräsentativen VDEW-Lastprofile step - by - step. VDEW Materialien. Frankfurt, German Association of Energy and Water Industries.

BDEW (2006). Praxisinformation P2006/8 Gastransport/Betriebswirtschaft - Anwendung von Standardlastprofilen zur Belieferung nicht-leistungsgemessener Kunden. Berlin, German Association of Energy and Water Industries.

Beer, J. G. D. (1998). Potential for Industrial Energy Efficiency Improvement in the Long Term. Utrecht, University of Utrecht. **PhD Thesis**.

Bisio, G. and G. Rubatto (2000). "Energy saving and some environment improvements in coke-oven plants." Energy **25**(3): 247-265.

Brown, H. L., B. B. Hamel, et al. (1985). Energy Analysis of 108 Industrial Processes. Philadelphia, The Fairmont Press, Inc.

CIBSE (2008). Energy Benchmarks, TM46: 2008, The Chartered Institute of Building Services Engineers.

Creditsafe (2010). License Agreement to use Creditsafe data, Creditsafe Business Solution Limited.

DECC (2008). MLSOA electricity and gas: 2007, Department of Energy and Climate Change.

DECC (2009). Digest of United Kingdom Energy Statistics 2009. I. MacLeay, K. Harris and A. Annut. http://www.decc.gov.uk/assets/decc/Statistics/publications/dukes/1_20100208131106_e_@@_dukes09.pdf.

DECC (2010). Energy Consumption in UK (ECUK), industrial data table, Department of Energy and Climate Change.

DEFRA (2010). EU Emissions Trading Scheme: Approved Phase II, National Allocation Plan 2008 - 2012. London, Department of Environment, Food and Rural Affairs.

DoE (2008). Waste Heat Recovery: Technology and Opportunities in U.S. Industry, U.S. Department of Energy, Industrial Technologies Program.

EIPPCB (2007). Reference Document on Best Available Techniques for the Manufacture of Large Volume Inorganic Chemicals - Ammonia, Acids and Fertilisers, European Integrated Pollution Prevention and Control Bureau (EIPPCB).

EIPPCB (2010). Final Draft BAT Guidance Note on Best Available Techniques for the Initial Melting and Production of Iron & Steel Sector, European Integrated Pollution Prevention and Control Bureau (EIPPCB).

EIPPCB (2010). Reference Document on Best Available Techniques in the Cement and Lime Manufacturing, European Integrated Pollution Prevention and Control Bureau (EIPPCB).

ElementEnergy (2010). Potential for the Application of CCS to UK Industry and Natural Gas Power Generation for Committee on Climate Change. Cambridge, Element Energy Limited. **Final Report Issue 3**.

Energetics and E3M (2004). Energy Use, Loss and Opportunities Analysis: U.S. Manufacturing and Mining. Washington, U.S. Department of Energy.

Entec (2006). DTI EU Emissions Trading Scheme Phase II - Review of New Entrants' Benchmarks - Integrated Steelworks. Northwich, Entec UK Limited. **www.bis.gov.uk/files/file33265.pdf**.

Entec (2006). DTI EU Emissions Trading Scheme Phase II - Review of New Entrants' Benchmarks - Lime. Northwich, Entec UK Limited. **www.bis.gov.uk/files/file33263.pdf**.

EPRI (2010). Waste Heat Recovery for Industrial Facilities, Technical Update, Electric Power Research Institute. **1020134**.

IEA (2007). Tracking Industrial Energy Efficiency and CO2 Emissions. Paris, International Energy Agency. **www.iea.org/textbase/nppdf/free/2007/tracking_emissions.pdf**.

McKenna, R. C. (2009). Industrial Energy Efficiency - Interdisciplinary perspectives on the thermodynamic, technical and economical constraints. Bath, University of Bath. **PhD thesis**.

McKoen, K., A. Koch, et al. (2011). Development of a methodology to calculate energy demand. Karlsruhe, Paris, EIFER. **2.0.1**.
Meidl, P. and K. McKoen (2010). Discussion Memo: Assumptions and Initial Results for the Clustering Approach within WP 2.1. Karlsruhe, EIFER.
Murshed, S. M., K. McKoen, et al. (2010). Comparison of Geoplan and Marketsafe data. Karlsruhe, EIFER.
NineSigma (2007). Request for Proposal: Waste Heat Recovery from Aluminium Smelting Off-Gas. Cleveland, Ohio, NineSigma, Inc.
ONS (2009). Statistical Bulletin: Labour market statistics, Office of National Statistics.
WMBW (2007). Nahwärmekonzepte - Kraft-Wärme-Kopplung und erneuerbare Energien. Stuttgart, Wirtschaftsministerium Baden-Württemberg.
Woods, P. (2011). Discussion paper on scenarios and sensitivities for WP5. London, AECOM. **3a**.
Worrell, E., D. Phylipsen, et al. (2000). Energy use and energy intensity of the U.S. chemical industry.
Worrell, E., L. Price, et al. (2001). "Energy efficiency and carbon dioxide emissions reduction opportunities in the US iron and steel sector." Energy **26**(5): 513-536.

Appendix A — Appendices Description of company records in Marketsafe data

Limited Company Information	
File:	Postcodes_EngWalScot
Description:	Postcodes from the UK
Rows:	2387476
# of distinct MLSOA:	8429
Original file:	POSTCODES table
Rows original file:	2491365
Explanation:	Extra postcodes here are most probably coming from North Ireland
File:	tblACC01_Sort_MaxAc001
Description:	File with company registration number and employee data
extraction process:	we selected the last possible accounting date of the company and the employee numbers associated with it
Rows:	2806228
total # of distinct ORG:	2806228
total # of distinct ORG with 0 employee	2692106
Original file:	ACC01
Rows original file:	10011126
# of distinct ORG original file:	2806228
File:	tblRAD01
Description:	File with company registration number and postcode data
Rows:	7636468
# of distinct ORG:	7082517
# of distinct Postcodes	1000303
# of ORG with no Postcodes	0
Original file:	RAD01

Total rows original file:	~7082517
total distinct ORG original file:	-
File:	tblLCD01
Description:	File with company registration number, registered name and SIC
Rows:	7606408
# of distinct ORG:	7055011
# of distinct SIC	1396
# of ORG with no SIC	3248076
Original file:	LCD01
Total rows original file:	~7055011
total distinct ORG original file:	-
File:	tblRAD01_LCD01_ACC01_PCPRSV_NS
Description:	File with company registration number, postcode data, SIC, employee numbers and postcodes
Rows:	2530617
# of distinct ORG:	2530617
# of distinct NO SPACE Postcodes	553276
# of distinct SIC	1071
total # of ORG with 0 employee	2428982
total # of ORG with no MLSOAIGZ	22102
Non Limited Company Information	
File:	Postcodes_EngWalScot
Description:	Postcodes from the UK
Rows:	2387476
# of distinct MLSOA:	8429
Original file:	POSTCODES table
Rows original file:	2491365

Explanation:	Extra postcodes here are most probably coming from North Ireland
File:	tblNCOD1
Description:	File with company registration number, registered name and postcode information
Rows:	1041745
# of distinct ORG:	1041745
# of distinct Postcodes	487007
# of ORG with null Posctode	10760
Original file:	NCOD1
Total rows original file:	~1041745
total distinct ORG original file:	-
File:	tblNGEN1
Description:	File with company registration number, registered name and postcode information
Rows:	1126503
# of distinct ORG:	1126503
# of distinct SIC	408
# of ORG with null SICN	11866
total # of distinct ORG with 0 employee	52812
total # of distinct ORG with Null employees	51
Original file:	NGEN1
Total rows original file:	~1126503
total distinct ORG original file:	-
File:	F_NLTD_Basic
Description:	File with company registration number, postcode data, SIC
Rows:	1041745
# of distinct ORG:	1041745
# of distinct NO SPACE Postcodes	487006

# of distinct SIC	406
total # of distinct ORG with 0 or null employee	48362

Table A-1: Description of limited and non limited company records in Marketsafe data

Appendix B — List of CIBSE Benchmarks

1. General office
2. High street agency
3. General retail
4. Large non-food shop
5. Small food Store
6. Large food store
7. Restaurant
8. Bar, pub or licensed club
9. Hotel
10. Cultural activities
11. Entertainment halls
12. Swimming pool centre
13. Fitness and health centre
14. Dry sports and leisure facility
15. Covered car park
16. Public buildings with light usage
17. Schools and seasonal public buildings
18. University Campus
19. Clinic
20. Hospital - clinical and research
21. Long term residential
22. General accommodation
23. Emergency services
24. Laboratory or operating theatre
25. Public waiting or circulation
26. Terminal
27. Workshop
28. Storage facility
29. Cold storage

Table B-1: The 29 different tertiary energy benchmark categories of CIBSE (For details see CIBSE, 2008)

Appendix C — Maps of GB heat demand

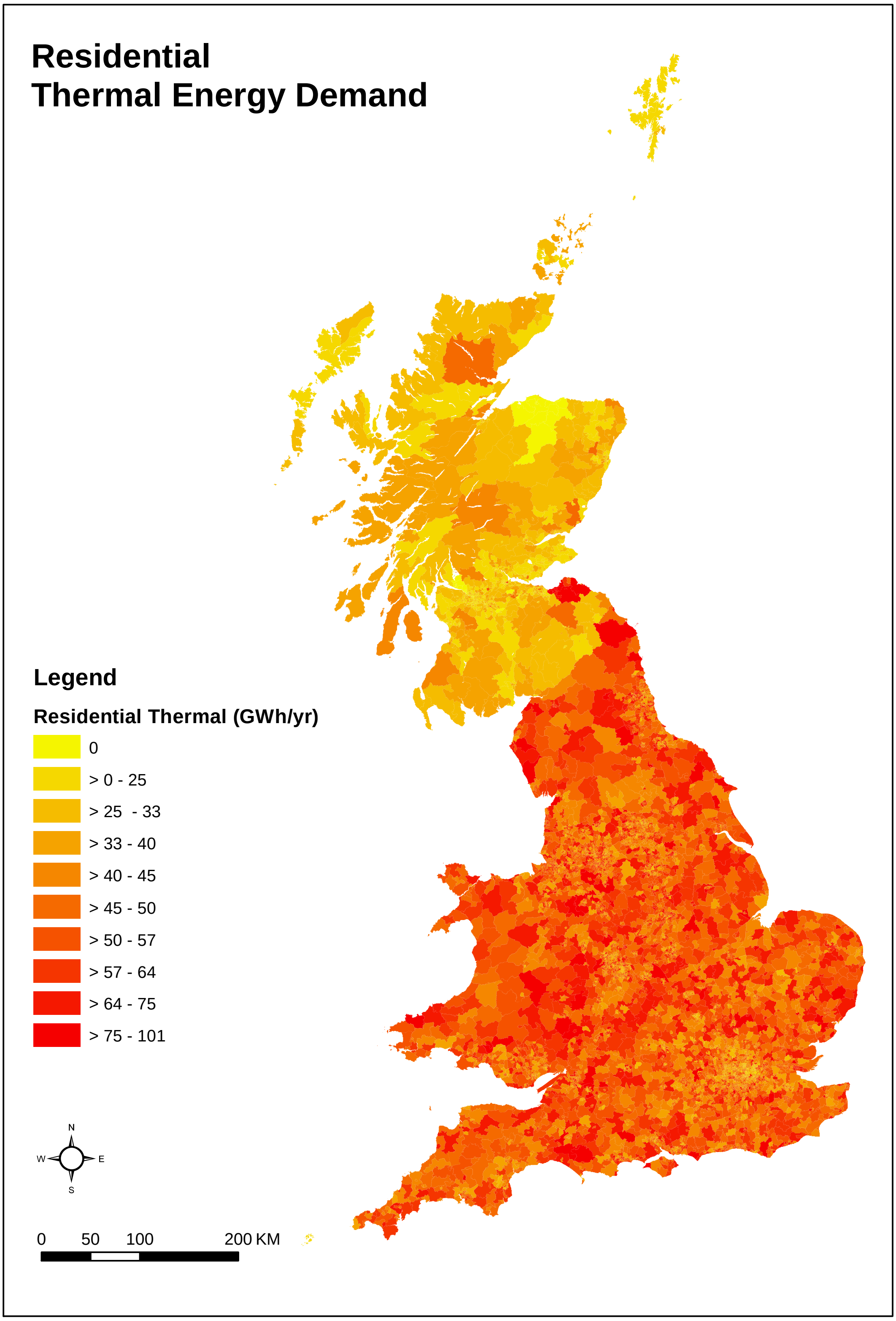


Figure C-1: Residential thermal energy demand in GB

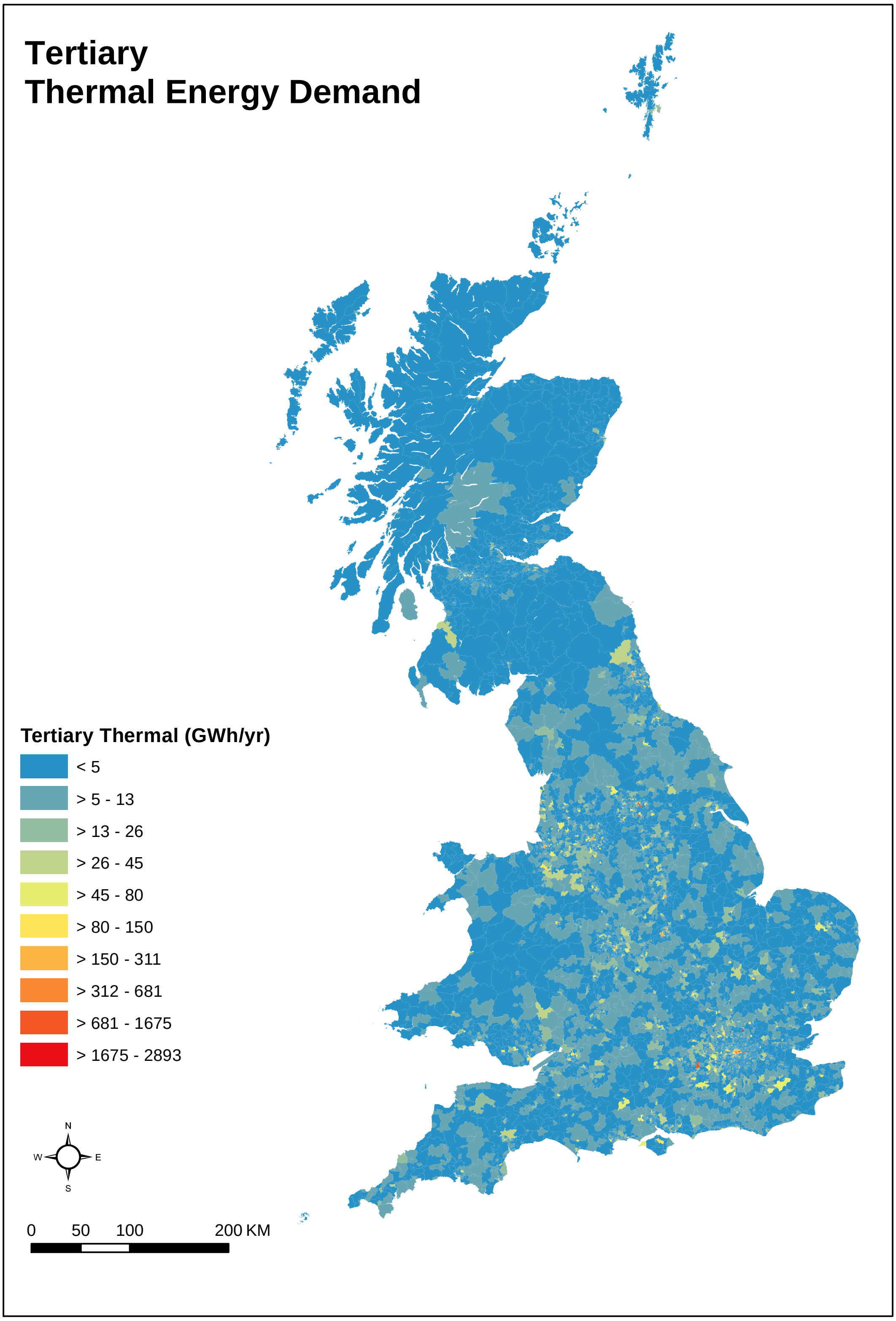


Figure C-2: Tertiary thermal energy demand in GB

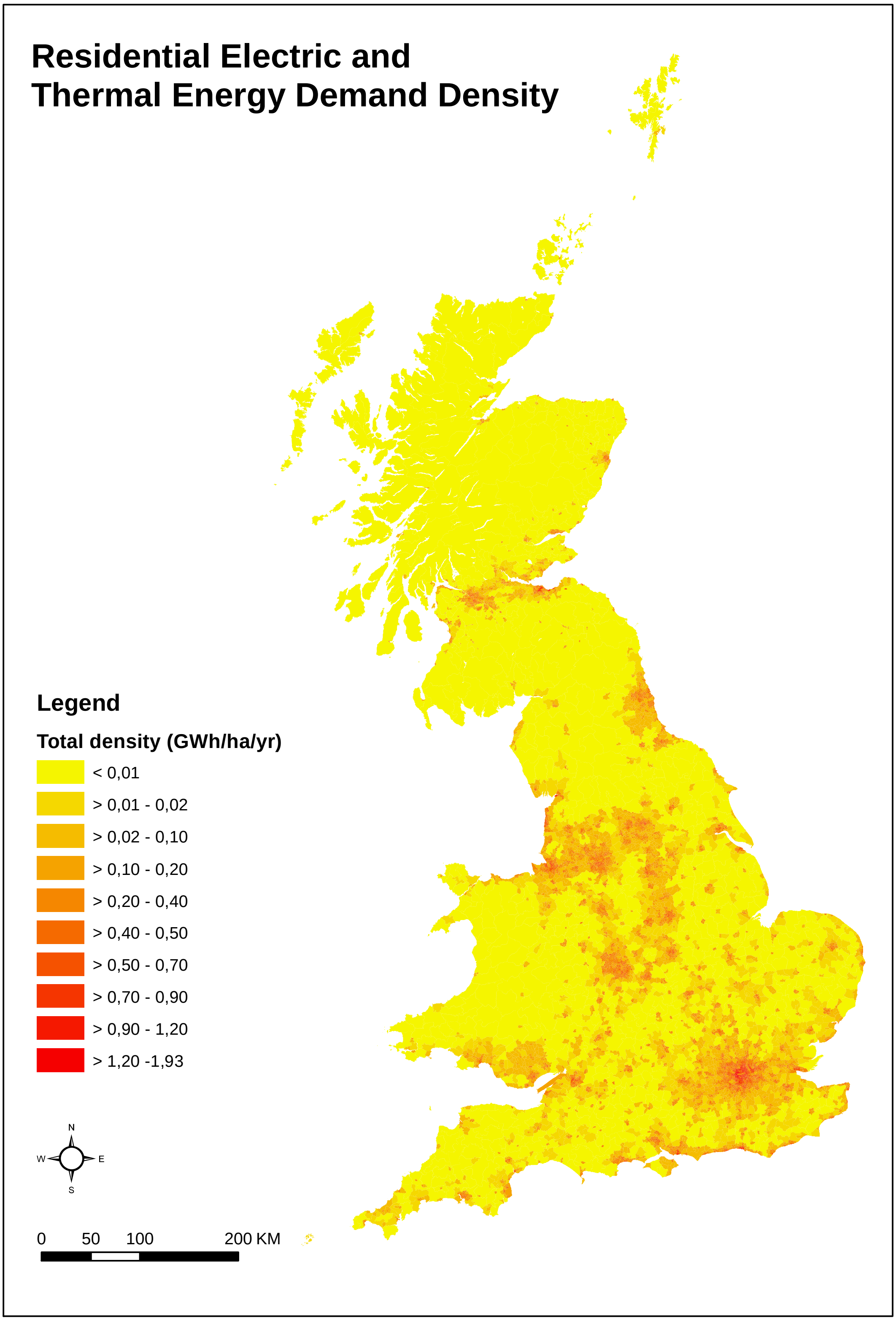


Figure C-3: Residential energy demand density in GB

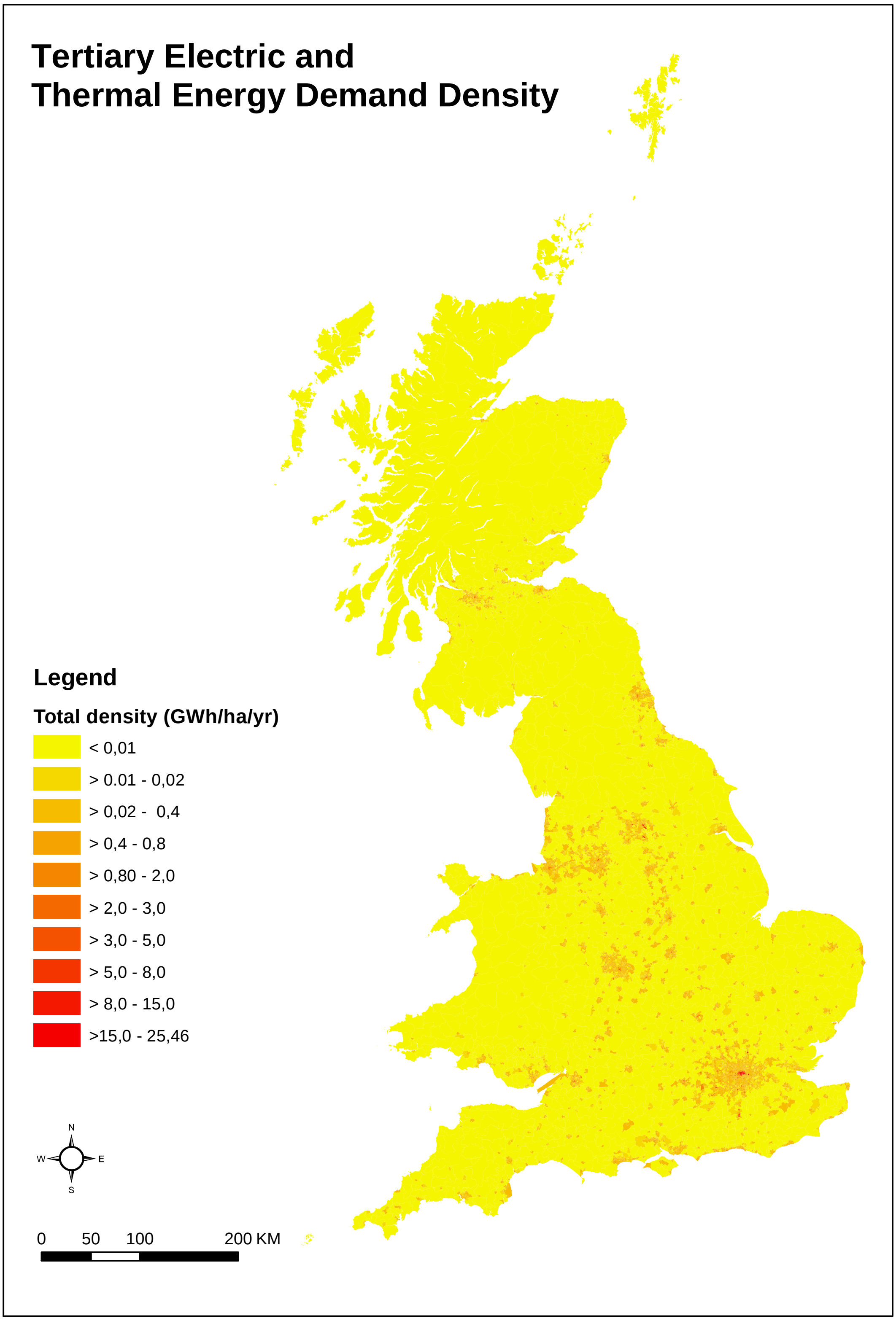


Figure C-4: Tertiary energy demand density in GB

Appendix D — Discussion Memo: Assumptions and Initial Results for the Clustering Approach within WP 2.1

Version from October 29th, 2010

In this memo we will present some initial results and discuss the assumptions taken for the clustering approach to create distributed energy zones for the whole UK within the Macro DE project.

Some of these assumptions are related to statements in the contract, others have come up during our working progress.

The discussion topics are the following:

- Initial Results
- Objective function
- Geopolitical boundaries
- Energy density
- Included number of MLSOA in our approach
- Boundaries of the constraints, i.e. number of zones, lower and upper boundaries for thermal and electrical demand for one zone
- Used input data

D.1 Analysis of Sample Region

Our initial calculations have been done on the GOR Yorkshire and Humber. There are two main arguments why we have chosen this GOR. At first, the local authority of Harrogate including the MLSOA Harrogate 15, on which we did the trial, lies within this region. Secondly, it is showing a mixed structure between very rural areas and dense populated areas, e.g. around Leeds.

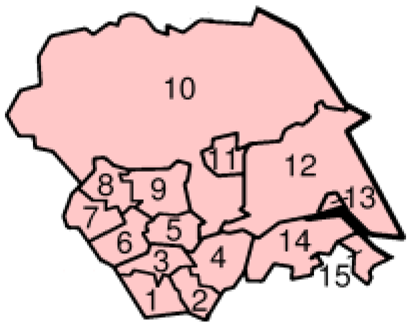


Figure D-1: Local Authorities of Yorkshire and Humber (taken from wikipedia.org)

No.	LA Name	No.	LA Name	No.	LA Name
1	Sheffield	2	Rotherham	3	Barnsley
4	Doncaster	5	Wakefield	6	Kirklees
7	Calderole	8	Bradford	9	Leeds
10	North Yorkshire*	11	York	12	East Riding of Yorkshire
13	Hull	14	North Lincolnshire	15	North East Lincolnshire

Table D-1: Names of Local Authorities of Yorkshire and Humber. The numbers are according to *including the districts Selby, Harrogate, Craven, Hambleton, Ryedale and Scarborough

The local authorities of the GOR are given graphically in Figure D-1 and their names can be found in Table D-1. North Yorkshire is further divided into districts, which will be respected in the calculations.

A map of all MLSOA within Yorkshire and Humber can be seen in Figure D-2. It is showing the energy density of the MLSOA in kWh/hectare (here domestic gas demand in kWh divided by the area in hectare) which will be discussed later on.

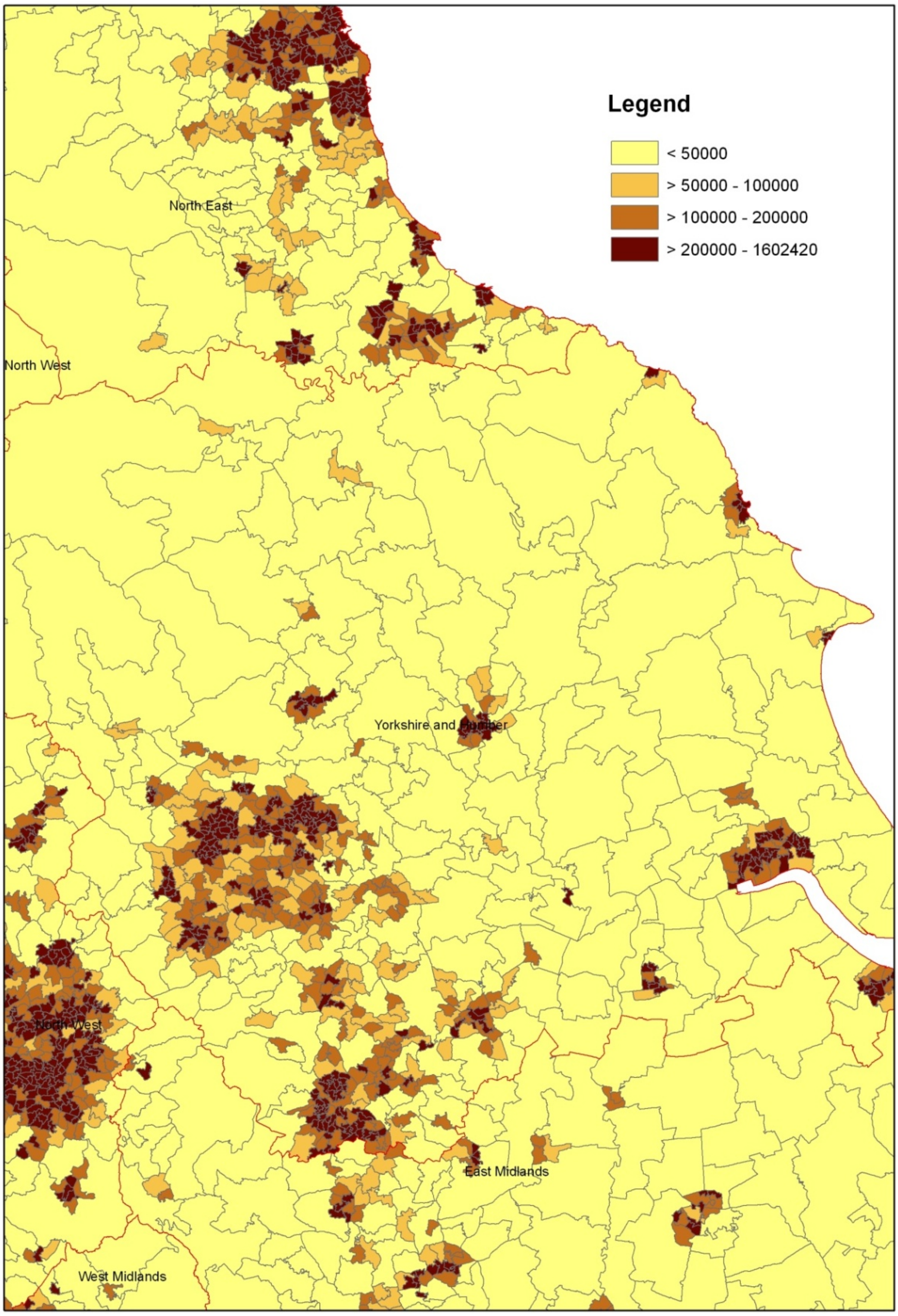


Figure D-2: MLSOAs in Yorkshire and Humber, coloured according to their domestic gas demands using DECC data

D.2 Initial Objective Function:

The initial objective function of the project is a weighted sum of normalized values of an environmental, economical and a distance effect:

- The environmental term represents the benefit in CO2 reduction of including more demand into the DE zones. The term contains a sum of normalized MLSOA demand.
- The economic term represents the revenues minus the costs associated with clustered zones. The revenues are approximated by the demand and the costs by the area per zone. The term contains the normalized MLSOA demand and the normalized MLSOA area.
- The distance term ensures that zones contain contiguous MLSOAs – that all MLSOAs of one zone are directly connected to at least one other assigned MLSOA. It contains the normalized distances between every MLSOA and every possible zone centers (i.e. MLSOAs which have been chosen by the program to be a seed).

These three parts of the objective function are weighted by two factors alpha, beta and the difference of both from one (1 - α - β) to see the influence of all the parameters on the objective function. Alpha and beta lie in the interval between 0 and 1 as does their sum.

$$\max[\alpha * Environmental\ Factor + \beta * Economical\ Factor + (1-\alpha-\beta) * Distance\ Factor]$$
$$= \max\left[\left(\alpha * \sum_{Zones} Normalised\ Zonal\ Energy\ Demand\right) + \left(\beta * \sum_{Zones} Normalised\ Zonal\ Energy\ Density\right) + \left((1-\alpha-\beta) * \sum_{Zones} (Normalised\ and\ Inverted\ Distance\ to\ Zone\ Center)\right)\right], with\ \alpha \le 1,\ \beta \le 1,\ \alpha+\beta \le 1$$

The complete and a bit complex term of the objective function will look like this:

$$\max\left[\left(\alpha * \sum_{Zone}\sum_{MLSOA} \frac{Gas\ Demand(MLSOA)}{\sum_{MLSOA} Gas\ Demand(MLSOA)} * Assign(MLSOA, zone)\right) + \left(\beta * \sum_{Zone}\sum_{MLSOA}\left(\frac{Gas\ Demand(MLSOA)}{\sum_{MLSOA} Gas\ Demand(MLSOA)} - \frac{Area(MLSOA)}{\sum_{MLSOA} Area(MLSOA)}\right) * Assign(MLSOA, zone)\right) + \left((1-\alpha-\beta) * \sum_{Zone}\sum_{MLSOA} \frac{Distance(MLSOA, zone) * Assign(MLSOA, zone)}{\sum_{Zone}\sum_{MLSOA} Distance(MLSOA, zone)}\right)\right]$$

Initial values for α and β have been set to 1/3 to be used as a baseline result with equally weighted parameters. In a different run we set (α = 0.5) and (β = 0.4) which led to slightly better results, but zones were created that were not necessarily continuous.

D.3 Boundaries of constraints

The optimization tool is using constraints which have to be respected for achieving a feasible solution. Thus these constraints have to be clearly defined.

1) Number of zones
   The number of zones has been discussed a lot until today in the project teams. In general, it has been agreed that about 1000 zones will be created out of 5000 MLSOAs, merging on average 5 MLSOA to one zone. In our approach, we will step back from this assumption, because it will be our goal to create the lowest possible number of zones with an economically acceptable energy density (see section about energy density), covering the biggest possible amount of energy demand in GB.

2) Electrical boundaries
   In the memorandum of February 18th, the project consortium agreed that CHP units with a power output between 100 kW_e and 50 MW_e will be considered for the distributed energy zones. This leads to the assumption that a zone should have between 1,000,000 kWh_e and 500,000 MWh_e (see **Error! Reference source not found.**) if the CHP unit is covering about 50 % of the total electrical demand and running for 5000 hours per year at full capacity.

3) Thermal boundaries
   The thermal boundaries have been calculated the same way as the electrical ones. It is assumed that a CHP unit will produce twice as much heat as it produces power (see **Error! Reference source not found.**).

Assumption/Statement	Electricity	Heat
Statement in contract about range of CHP units	100 kW_e to 50 MW_e	-
Heat to electricity ration of a CHP unit is 2:1	100 kW_e to 50 MW_e	200 kW_{th} to 100 MW_{th}
A CHP unit should at least run 5000 hours per year at full capacity	500,000 kWh_e to 250,000 MWh_e → 500 MWh_e to 250 GWh_e	1,000,000 kWh_{th} to 500,000 MWh_{th} → 1,000 MWh_{th} to 500 GWh_{th}
CHP unit satisfies 50 % of zonal demand	1,000,000 kWh_e to 500,000 MWh_e → 1,000 MWh_e to 500 GWh_e	2,000,000 kWh_{th} to 1,000,000 MWh_{th} → 2,000 MWh_{th} to 1,000 GWh_{th}

Table D-2: Statements and Assumption about the sizes of CHP units that might be used in the zones for district heating

D.4 Input Data

The used input data for the optimization tool are mainly taken from DECC. All data are used on a MLSOA level. The domestic gas demand is one of the main inputs, although data for the gas demand of the tertiary and industrial sectors are also available. They are not taken into account, because the industrial demand cannot be separated from the tertiary demand and is not subject to our research as a consumer of district heat. Tertiary demand will be instead calculated by using a data base with information about the location and the number of employees of all companies and a list of conversion rates plus benchmarks indicating the energy consumption per employee for different SIC. The same holds for the electrical demand on MLSOA level.

Other data used from DECC are the area of MLSOAs. Herewith the energy densities of MLSOAs and later on of zones can be calculated.

The center of each MLSOA represents geographic coordinates which are extracted from GPS and used to calculate the distance from other MLSOAs or the zone center, the MLSOA is assigned to. The underlying assumption is to minimize the distance between a zone center and an assigned MLSOA, so that the zones are continuous and all MLSOAs of one zone are directly connected to at least one other assigned MLSOA.

D.5 Energy Density

The energy density is a crucial indicator for the economic viability of a district heating network. A general rule of thumb is that the area should at least have a density of 250 – 300 MWh per hectare per year (see “Nahwärmekonzepte – Kraft-Wärme-Kopplung und erneuerbare Energie”, published by the Ministry of Economics of the federal state Baden-Württemberg).

As these numbers are referring to the built-up area, e.g. cities, while we are using the area of administrative units, the lowest accepted energy density of each zone or MLSOA should be reduced. Therefore, we are currently running different scenarios with a minimum energy density of all MLSOAs of 50.000, 100.000 and 200.000 kWh/hectare. The percentages of MLSOAs in GB below or above these values are given in Table D-3. It is evident that more than 50 % of all MLSOA have an energy density below 200000 kWh/hectare if we consider only domestic demand. If we take the tertiary and industrial demand into account, too, the number increases to 53,3 %. Currently, we are using only the domestic demand, to avoid disturbing the results due to outliers with high industrial demand as discussed in the section “input data” above. Figure 2 is showing a map of Yorkshire and Humber with the classified MLOSA according to their energy densities. It showing dense populated areas, e.g. around Leeds and Sheffield, but most of this region is showing a energy density of less than 50000 kWh/hectare. The impact of these will be discussed further in the section about filtering of MLSOA.

kWh/hectare	0	< 50000	< 100000	< 200000	> 200000
Domestic Demand	1,4 %	26,8 %	36,8 %	58,2 %	41,8 %
Domestic, Tertiary and Industrial Demand	1,3 %	23,8 %	31,9 %	46,7 %	53,3 %

Table D-3: Percentage distribution of the energy density of MLSOAs in GB. Remark: The classes are cumulativ, except the last one.

Figure D-3 and Figure D-4 are showing to scatter plots for the MLSOAs in GB. The horizontal axes are showing the energy demand in kWh, either domestic only or domestic, tertiary and industrial. The vertical axes are showing the energy density in kWh per hectare which is the former mentioned energy demand divided by the area of the MLSOA.

For the domestic demand it can be concluded that the energy demand is following more or less a Gaussian distribution with a mean demand of 50000000 kWh. On the other hand, the density is clearly showing a concentration between the values 0 kWh/hectare and 400000 kWh/hectare, which would lead to the assumption that the mean should be around the threshold value of 200000 kWh/hectare. If we exclude all MLSOA with zero demand and thus a density of 0 from our calculations, the assumption holds. The average energy density of all MLSOA with a domestic gas demand above 0 kWh is 200106,4 kWh/hectare.

If the tertiary and industrial demand is included into the calculations the average demand density is increasing to 278478,5 kWh/hectare. Although the majority of data points seem to be gathered much closer in the lower left corner, it is clearly visible that this effect is caused by a couple of outliers, either having a large demand or a high density. The average demand is increased by less than 50 %, while the maximum value is increased by more than 1300 %!

All in all, it is our recommendation to create zones with minimum energy densities of 200000 kWh/hectare including domestic and tertiary demand.

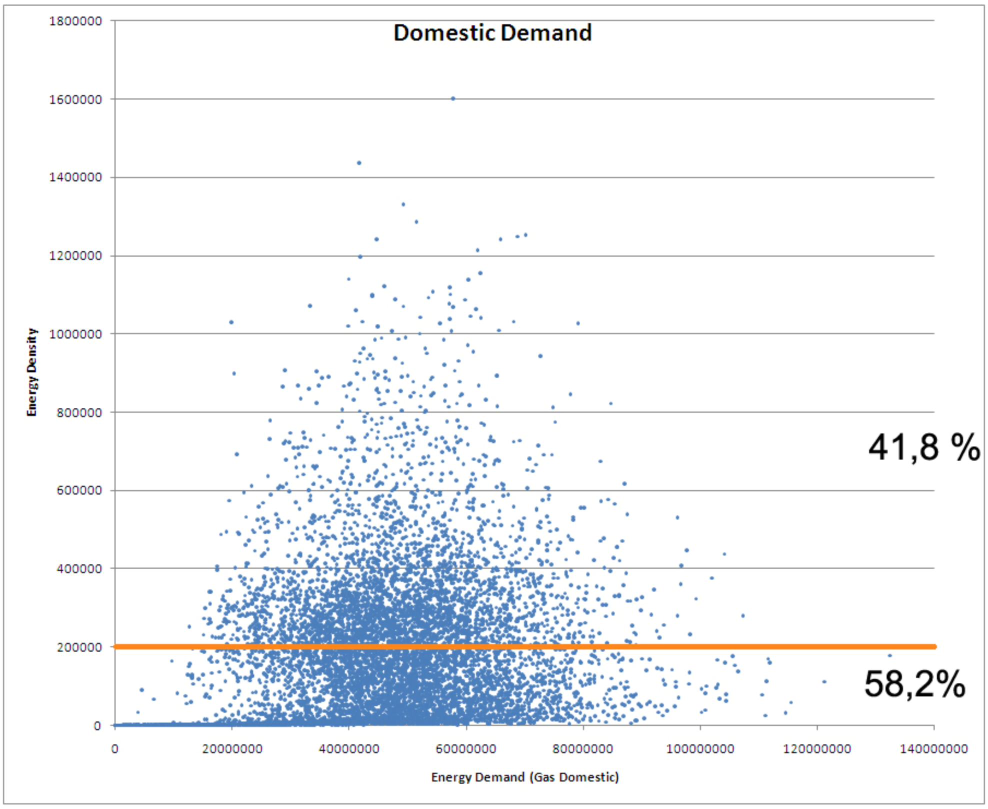


Figure D-3: Scatter plot of MLSOAs in GB showing the domestic gas demand (in kWh) and the corresponding energy density (in kWh/hectare)

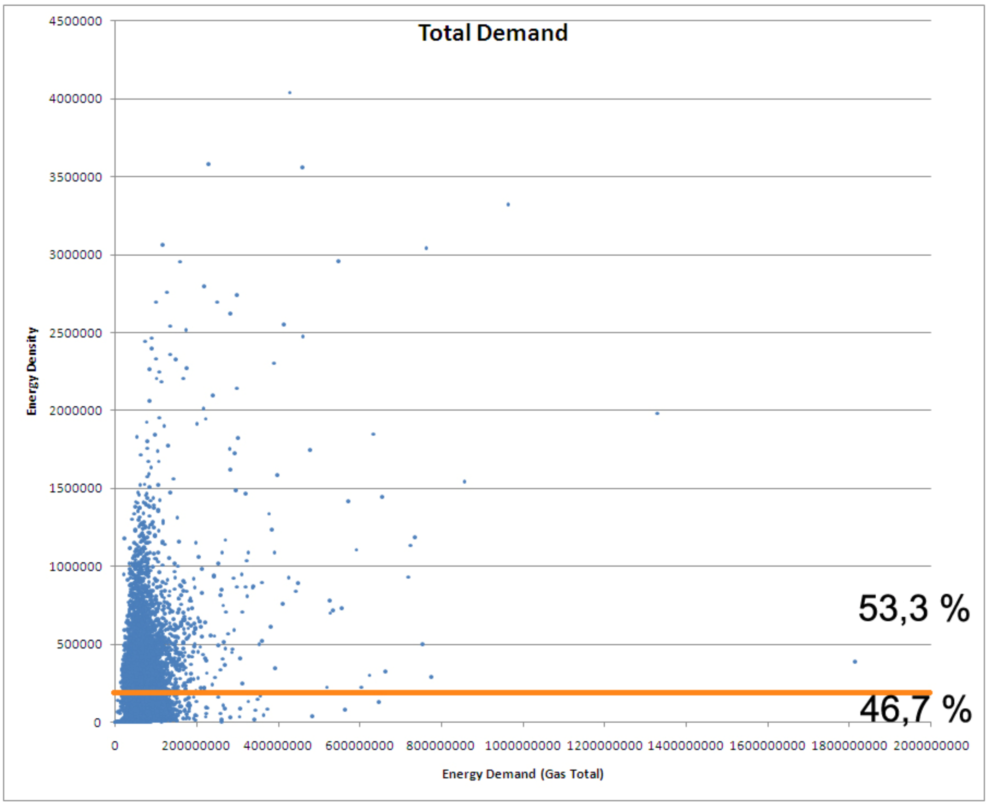


Figure D-4: Scatter plot of MLSOAs in GB showing the total gas demand (in kWh) and the corresponding energy density (in kWh/hectare)

D.6 Geopolitical Boundaries

The geopolitical boundaries have a major impact on the creation of the zones, because they are limiting the number of MLSOAs that are taken into account during one run of the optimization program. Two options are discussed at the moment; the Government Office Region (GOR) and the Local Authority (LA) option (see Figure D-5).

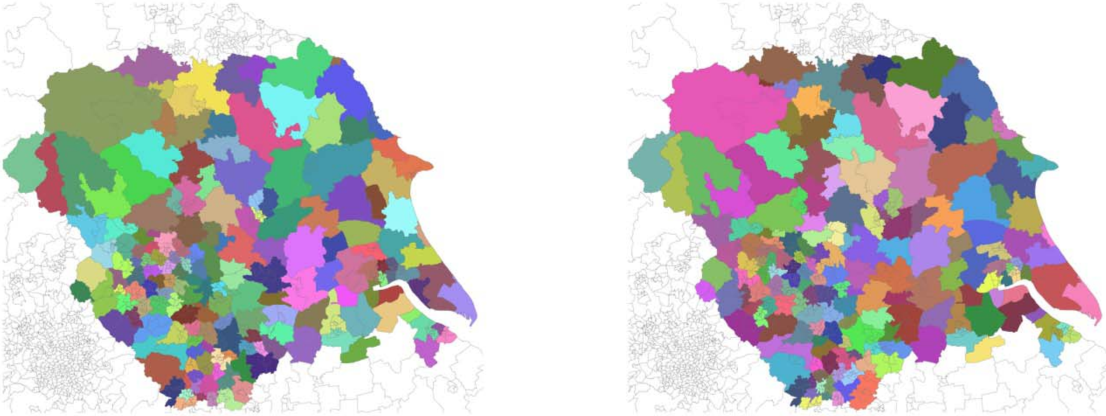
The GOR option is dividing the problem to cluster the whole of the UK into 11 smaller problems, which make it easier to handle the program by reducing the runtime and the amount of data used at the same time significantly. There are 9 GOR in England plus Wales and Scotland. Northern Ireland is excluded from these calculations.

The advantage of this option is a broader view of the problem than at local authority level. The zones are created within one GOR so that small LAs can be attached to other zones, creating, hopefully more efficient ones, or two small LAs can be combined together to one zone.

The disadvantage is the possibility of generating zones that cross LA borders.

The LA option does not allow zones to cross local authority boundaries. In this case LA refers to the Local Authority code given in the DECC data which is a mixture of counties and districts, always showing the smaller unit if existing. This option refers to the assumption that district heating grids will not cross LA borders, as the governments of local authorities will most probably be the driving force for the grid installation. The disadvantages of this approach will be suboptimal solutions, because especially in highly populated areas, e.g. city centers, DE systems which cross LA boundaries could lead to higher efficiencies and higher emission savings.

Although the contract is stating that LA boundaries will be respected for the creation of zone, our recommendation is to use the borders of GORs as fixed borders, allowing zones to cross the borders of LAs. The removal of this constraint leads to better results.



Initial results for Yorkshire and Humber, ignoring LA borders

Initial results for Yorkshire and Humber, respecting LA borders

Figure D-5: Comparison of initial results for Yorkshire and Humber, ignoring (left side) and respecting (right side) the LA boundaries

D.7 Initial Clustered Zone Results

This section presents very initial clustered zone results and our analysis of them.

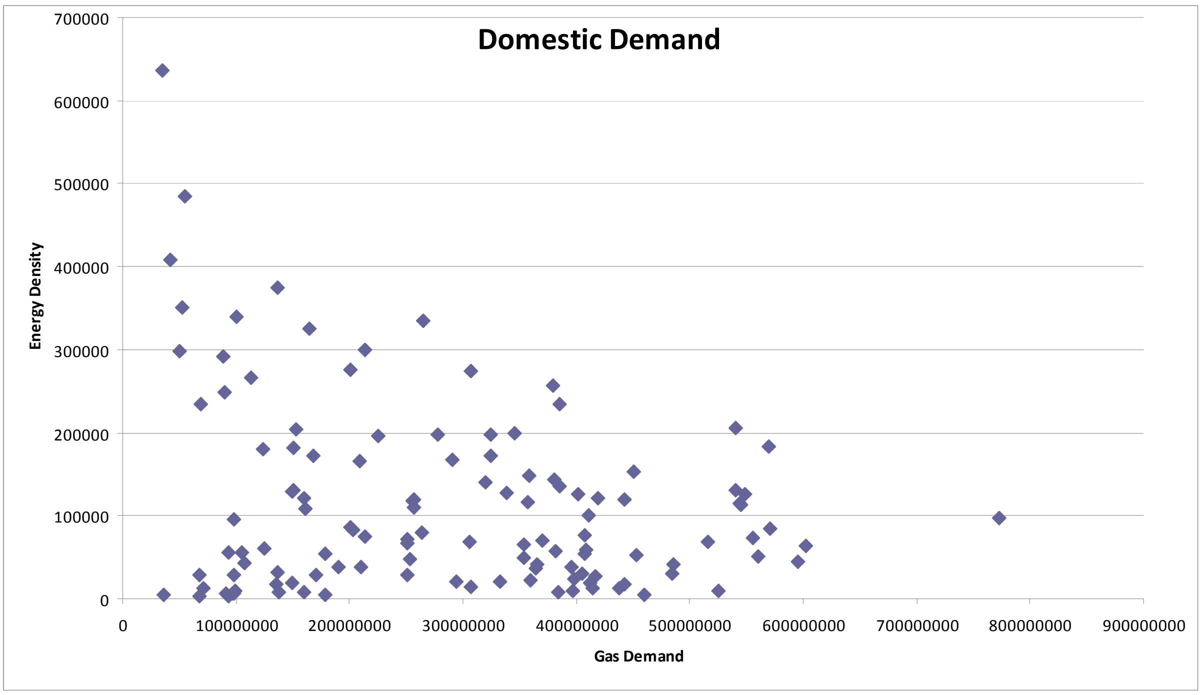


Figure D-6: Distribution of zone density against zone domestic demand

Our initial objective is to create as many economically viable zones as possible and to cover as much of the region demand as possible. As described in the table below, this clustering results gives an average density of 100 MWh/hr and includes all MLSOAs. The zones were created respecting LA boundaries as shown on the right hand side of Figure 5. On average, the number of MLSOAs by zone is 5. Our maximum demand, with a 50 MWe power limit, is 1000 GWh p.a. which is already attained by 1 zone with domestic demand only. On average the zone demand is 300 GWh p.a.

characteristic	average	min	max
# MLSOAs	5	1	13
demand GWh	300 GWh	35	1000
density MWh/hectare	100 MWh/hr	2	600
area km^2	100	1	1000

Table D-4: Zone Statistics

D.8 Current Clustering Investigations

D.8.1 Included Number of MLSOAs

In the UK there are 8429 MLSOAs. At least 113 of them have absolutely no gas consumption according to the DECC data. The same counts for at least 4 local authorities which are representing 23 MLSOAs. These 113 MLSOAs will be excluded from the optimization program, because their demand cannot be estimated. Additionally, it can be assumed that these MLSOAs are very rural and will not be economically viable for a DE solution. The number of excluded MLSOAs can be increased by filtering. One option is the use of the energy density (demand divided by area) which is shown in Table D-5 for four different density classes and two different demand types.

kWh/hectare	0	0 – 50000	50000 – 100000	100000 – 200000	> 200000
Domestic Demand	121	**2136**	847	1800	3525
Domestic, Tertiary and Industrial Demand	113	1895	683	1244	4494

Table D-5: Number of MLSOAs within different classes of energy density (gas density) in kWh/hectare. The second row is showing only the domestic demand, while the third one shows the domestic demand plus tertiary and industrial demand.

Our recommendation is to include all MLSOA with an energy density above 100000 kWh/hectare of the domestic demand. This will lead to an exclusion of 3104 MLSOA, leaving 5325 MLSOAs for the clustering.

Initially, it was agreed that about 5000 MLSOA will be taken into account after filtering unusable MLSOAs.

Additionally, it is possible for the optimization program to eliminate MLSOAs during the clustering process which might not fit into a created zone in its geographical neighbourhood.

Appendix E — List of Zone Parameters

ID	Parameter Abbreviation	Parameter	Unit	Description
1	HeatDom	Heat domestic	kWh	Domestic heat demand per MLSOA/Zone
2	HeatTert	Heat tertiary	kWh	Tertiary heat demand per MLSOA/Zone
3	HeatTot	Heat Total	kWh	Total heat demand per MLSOA/Zone (Domestic + Tertiary heat demand)
4	ElDom	Electricity domestic	kWh	Domestic Electrical demand per MLSOA/Zone
5	ElTert	Electricity tertiary	kWh	Tertiary Electrical demand per MLSOA/Zone
6	ElTot	Electricity total	kWh	Total Electrical demand per MLSOA/Zone (Domestic + Tertiary Electrical demand)
7	DensHeatDom	Domestic Heat Density	kWh/ha	Domestic heat density per MLSOA/Zone
8	DenstHeatTert	Tertiary HeatДенstiy	kWh/ha	Tertiary heat density per MLSOA/Zone
9	DensHeatTot	Total Heat Density	kWh/ha	Total heat density per MLSOA/Zone (Domestic + Tertiary heat density)
10	DensElDom	Domestic Electrical Density	kWh/ha	Domestic Electrical density per MLSOA/Zone
11	DensElTert	Tertiary Electrical Density	kWh/ha	Tertiary Electrical density per MLSOA/Zone
12	DensElTot	Total Electrical Density	kWh/ha	Total Electrical density per MLSOA/Zone (Domestic + Tertiary Electrical density)
13	Population	Population	dimensionless	Number of inhabitants of MLSOA/Zone
14	Area	Area	ha	Area of MLSOA/Zone
15	Households	Number of Households	dimensionless	Number of Households per MLSOA/Zone
16	MLSOANumber	Number of MLSOAs per zone	dimensionless	Number of MLSOAs per Zone
17	BorderLength	Length of common border	km	Length of common border of two neighbouring MLSOAs
18	HeatCoeff	Heat Coefficient	dimensionless	Ratio of tertiary heat demand and domestic heat demand
19	ElCoeff	Electrical coefficient	dimensionless	Ratio of tertiary electrical demand and domestic electrical demand
20	HeatDom/cap	Heat domestic per capita	kWh p.P.	Heat domestic divided by population
21	TotalTertSites	Number of Tertiary sites	dimensionless	Number of tertiary sites per MLSOA/Zone
22	Ter/HH	Tertiary Sties per Households	dimensionless	Ration of tertiary sites and households per MLSOA/Zone
23	Peak W	Peak Winter	kWh	Peak hour of zone in winter time
24	Base W	Base Winter	kWh	Base hour of zone in winter time
25	Peak-Base W	Peak –Base Winter	kWh	Difference between peak hour and base hour of zone in winter time
26	Average W	Average Winter	kWh	Average demand of zone in transition time
27	Peak T	Peak Transition	kWh	Peak hour of zone in transition time
28	Base T	Base Transition	kWh	Base hour of zone in transition time
29	Peak – Base T	Peak –Base Transition	kWh	Difference between peak hour and base hour of zone in transition time
30	Average T	Average Transition	kWh	Average demand of zone in summer time

31	Peak S	Peak Summer	kWh	Peak hour of zone in summer time
32	Base S	Base Summer	kWh	Base hour of zone in summer time
33	Peak – Base S	Peak –Base Summer	kWh	Difference between peak hour and base hour of zone in summer time
34	Average S	Average Summer	kWh	Average demand of zone in winter time
35	Max	Peak	kWh	Annual peak hour of zone
36	Min	Base	kWh	Annual base hour of zone
37	P-B	Peak - Base	kWh	Difference between annual peak hour and annual base hour
38	B/P	Base over peak	dimensionless	Ratio of annual base hour and annual peak hour
39	HeatTot/Max	(Heat total / Peak) Load Factor	dimensionless	Load Factor; Ratio of total heat demand and annual peak hour
40	HT/P-B	Heat total over (Peak – Base)	dimensionless	Ratio of total heat demand and (annual peak hour – annual base hour)
41	TRL	Total road length	Km	Length of all roads in MLSOA/Zone
42	LHD	Line Heat Density	GWh/km or MWh/m	Total Heat Demand per length of roads

Table E-1: List of all Zone parameter and their descriptions

Appendix F — List of Quartile Zones

Out of each class, 4 further zones have been extracted to sent to and analysed by the university of Manchester, if the classes are homogeneous within themselves. The zones represent the first and the third quartiles of the two parameters total heat demand and heat density. The 80 chosen zones and their respective value of the parameter are shown in the Table F-1: to Table F-4: below.

	1st Quartile on Total Heat									
Class Nr.	1	2	3	4	5	6	7	8	9	10
Zone Nr.	87	8	196	378	544	744	262	114	55	659
Heat Total (kWh)	2,86E+08	2,98E+08	2,20E+08	3,90E+08	2,60E+08	3,18E+08	4,73E+08	2,56E+08	4,28E+08	3,59E+08
Class Nr.	11	12	13	14	15	16	17	18	19	20
Zone Nr.	612	687	640	875	622	30	654	478	379	893
Heat Total (kWh)	9,92E+07	4,30E+08	1,64E+08	3,41E+07	4,82E+07	3,30E+08	2,90E+08	2,59E+08	3,55E+08	1,96E+07

Table F-1: Chosen zones for 1st quartile of total heat demand and their heat demand values

	3rd Quartile on Total Heat									
Class Nr.	1	2	3	4	5	6	7	8	9	10
Zone Nr.	31	482	735	679	74	676	900	77	103	207
Heat Total (kWh)	3,51E+08	3,18E+08	2,41E+08	4,10E+08	2,80E+08	3,37E+08	4,88E+08	3,17E+08	4,68E+08	3,70E+08
Class Nr.	11	12	13	14	15	16	17	18	19	20
Zone Nr.	557	421	857	606	511	493	705	632	481	848
Heat Total (kWh)	1,26E+08	4,55E+08	1,95E+08	4,04E+07	6,45E+07	4,43E+08	3,04E+08	2,77E+08	3,73E+08	2,72E+07

Table F-2: Chosen zones for 3rd quartile of total heat demand and their heat demand values

	1st Quartile on Heat Density									
Class Nr.	1	2	3	4	5	6	7	8	9	10
Zone Nr.	31	408	844	779	11	457	127	760	103	941
Total Heat Density (kWh/hectare)	3,22E+06	3,56E+05	2,10E+05	2,20E+05	3,23E+05	2,10E+05	2,30E+05	5,70E+05	5,00E+05	2,99E+05
Class Nr.	11	12	13	14	15	16	17	18	19	20
Zone Nr.	551	313	201	597	933	493	182	377	911	853
Total Heat Density (kWh/hectare)	2,20E+05	2,20E+05	2,10E+05	2,31E+05	2,10E+05	1,04E+06	2,08E+05	2,16E+05	2,13E+05	2,13E+05

Table F-3: Chosen zones for 1st quartile of heat density and their heat demand values

3rd Quartile on Heat Density										
Class Nr.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Zone Nr.	86	213	316	242	502	685	286	98	85	266
Total Heat Density (kWh/hectare)	3,68E+06	4,33E+05	3,20E+05	2,90E+05	3,84E+05	2,80E+05	3,20E+05	7,20E+05	6,10E+05	3,69E+05
Class Nr.	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Zone Nr.	610	395	738	505	392	123	500	51	93	812
Total Heat Density (kWh/hectare)	2,60E+05	2,90E+05	2,40E+05	3,04E+05	2,70E+05	1,42E+06	2,43E+05	2,57E+05	2,60E+05	2,42E+05

Table F-4: Chosen zones for 3rd quartile of heat density and their heat demand values

Appendix G — Demand data and design results for characteristic zones

Class number*	Characteristic zone number*	Annual thermal demand (GWh)	Zone area (ha)	Thermal density (GWh/ha)	Peak thermal demand (MWh)
1	154	349.3	81	4.31	141.8
2	711	310.2	976	0.32	107.9
3	239	229.5	1027	0.22	77.1
4	619	399.9	1749	0.23	139.9
5	70	269.6	821	0.33	103.7
6	509	328.1	1478	0.22	112.5
7	127	480	2097	0.23	189.7
8	145	280.6	467	0.60	106.9
9	29	451.7	533	0.85	159.6
10	138	363.5	1051	0.35	134.3
11	566	112.7	455	0.25	38.2
12	313	444	1993	0.22	153.3
13	686	180.6	806	0.22	59.8
14	535	37.6	149	0.25	13.6
15	734	56.7	255	0.22	19.8
16	148	374.2	256	1.46	144.8
17	23	297	1063	0.28	118.7
18	547	267.7	964	0.28	88.6
19	701	364.2	1757	0.21	121.3
20	909	23.3	108	0.21	7.3

Table G-1: Design inputs for characteristic zones

See Chapter 4 —Classification (p. 87ff) for details.

Class number	Total annualised cost of DE centre (£/yr)	DHN annualised capital cost (£/yr)	Annualised capital cost of DE centre (£/yr)	Total annualised cost of DE solution (£/yr)	Fuel CO2 emissions (t/yr)	Global CO2 emissions (t/yr)
1	9.3E+06	1.3E+07	6.1E+06	2.3E+07	142,000	7,000
2	8.0E+06	1.6E+07	5.5E+06	2.4E+07	127,000	4,900
3	6.2E+06	1.2E+07	4.7E+06	1.9E+07	96,000	550
4	1.0E+07	2.1E+07	6.0E+06	3.1E+07	162,000	13,600
5	7.3E+06	1.8E+07	5.1E+06	2.5E+07	110,000	4,100
6	8.5E+06	1.8E+07	5.9E+06	2.6E+07	136,000	4,500
7	1.3E+07	2.7E+07	7.1E+06	4.0E+07	194,000	21,300
8	7.6E+06	1.3E+07	5.2E+06	2.1E+07	114,000	5,400
9	1.2E+07	2.4E+07	6.7E+06	3.6E+07	184,000	18,800
10	9.5E+06	2.1E+07	6.0E+06	3.1E+07	148,000	10,900
11	2.9E+06	6.0E+06	2.2E+06	8.9E+06	46,700	780
12	1.2E+07	2.3E+07	7.2E+06	3.5E+07	183,000	14,100
13	4.9E+06	9.1E+06	3.6E+06	1.4E+07	74,700	1,300
14	1.1E+06	2.0E+06	8.5E+05	3.2E+06	15,700	550
15	1.6E+06	3.1E+06	1.2E+06	4.7E+06	24,100	108
16	9.9E+06	1.9E+07	6.1E+06	2.9E+07	155,000	13,200
17	8.0E+06	1.6E+07	5.2E+06	2.4E+07	120,000	7,700
18	7.0E+06	1.4E+07	5.1E+06	2.1E+07	111,000	2,100
19	9.1E+06	1.9E+07	5.9E+06	2.8E+07	153,000	10,300
20	5.7E+05	1.3E+06	4.1E+05	1.8E+06	10,100	320

Table G-2: Design results for characteristic zones

Class number	Thermal density / thermal intensity type for the 5 zones analysed in each class				
Zone →	CZ	Q1	Q3	TD1	TD3
1*	HD	HD	HD	HD	HD/HI
2*	LD	HD	HD	HD	HD
3	LD	LD	LD	LD	LD
4	LD	LD	LD	LD	LD
5*	LD	HD	LD	LD	HD
6	LD	LD	LD	LD	LD
7	LD	LD	LD	LD	LD
8*	HD	HD	HD	HD/HI	HD
9	HD	HD	HD	HD	HD
10*	LD	LD	LD	LD	HD
11	LD	LD	LD	LD	LD
12	LD	LD	LD	LD	LD
13	LD	LD	LD	LD	LD
14	LD	LD	LD	LD	LD
15*	LD	LD	HD	LD	LD
16	HD	HD	HD	HD	HD
17	LD	LD	LD	LD	LD
18	LD	LD	LD	LD	LD
19	LD	LD	LD	LD	LD

Table G-3: Thermal demand density type of characteristic and quartile zone in each class
***classes with zones of more than one density / intensity type**

CZ: characteristic zone; Q1, Q3: zones representing first and third quartiles in terms of thermal demand; TD1, TD3: zones representing first and third quartiles in terms of thermal demand density)

Zone types (see Table 5-3)

Low density (LD): thermal demand < 0.35 GWh/hectare

High density (HD): thermal demand > 0.35 GWh/hectare, thermal intensity < 0.04 GWh/meter

High density, high intensity (HD/HI): thermal demand > 0.35 GWh/hectare, thermal intensity > 0.04 GWh/meter

Appendix H — DE solution performance presented on a per unit of heat delivered basis

Figures B.1 and B.2 present the total annualised cost and annualised capital cost of the DE centre per unit of heat delivered, respectively. Table 5-7 confirms that there is relatively little scatter in this performance indicator. Similar costs within each class reflect the similar DE solutions generated.

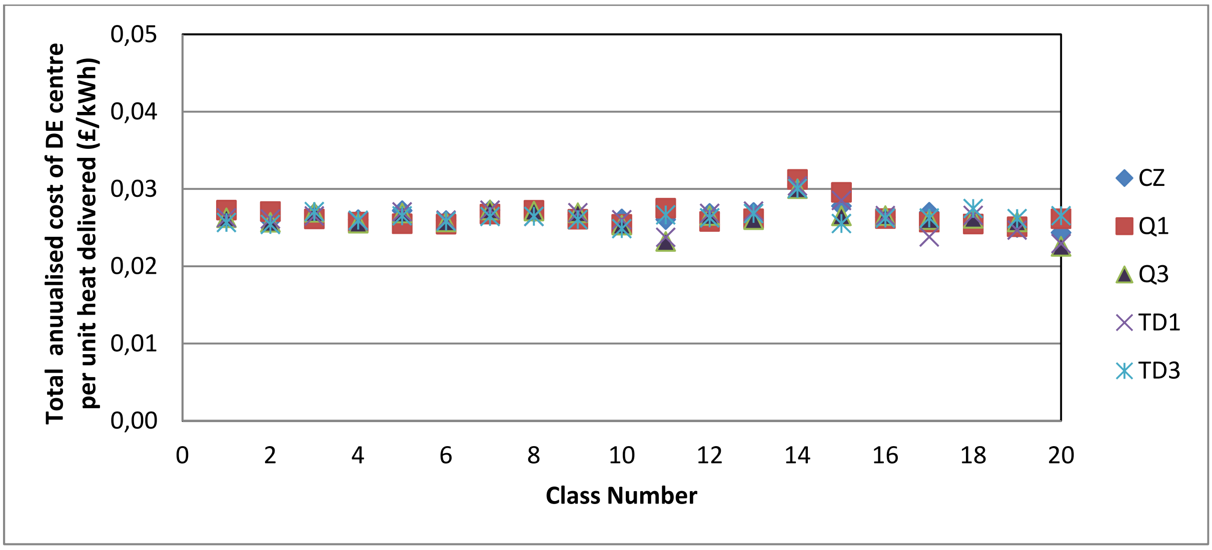


Figure H-1: Total annualised cost of DE centre per unit of heat delivered

(CZ: characteristic zone; Q1, Q3: zones representing first and third quartiles in terms of thermal demand; TD1, TD3: zones representing first and third quartiles in terms of thermal demand density)

Figure H-2: Annualised capital cost of DE centre per unit of heat delivered

Figure H-3 presents the annualised capital cost of the DHN where the model of Table 5-3 is applied; significant variation in cost may be observed. Table 5-7 indicates that costs vary by more than 20% in six classes (i.e. classes 1, 2, 5, 8, 10, 19). The source of variations is further investigated in Figure H-4 and in Appendix I —: Figure H-4 takes all zones in a given class to be of the same 'type' and Appendix I — demonstrates that variations in the numbers of domestic meters and tertiary meters also introduce variations in cost of 0.01 p/kWh to 0.02 p/kWh.

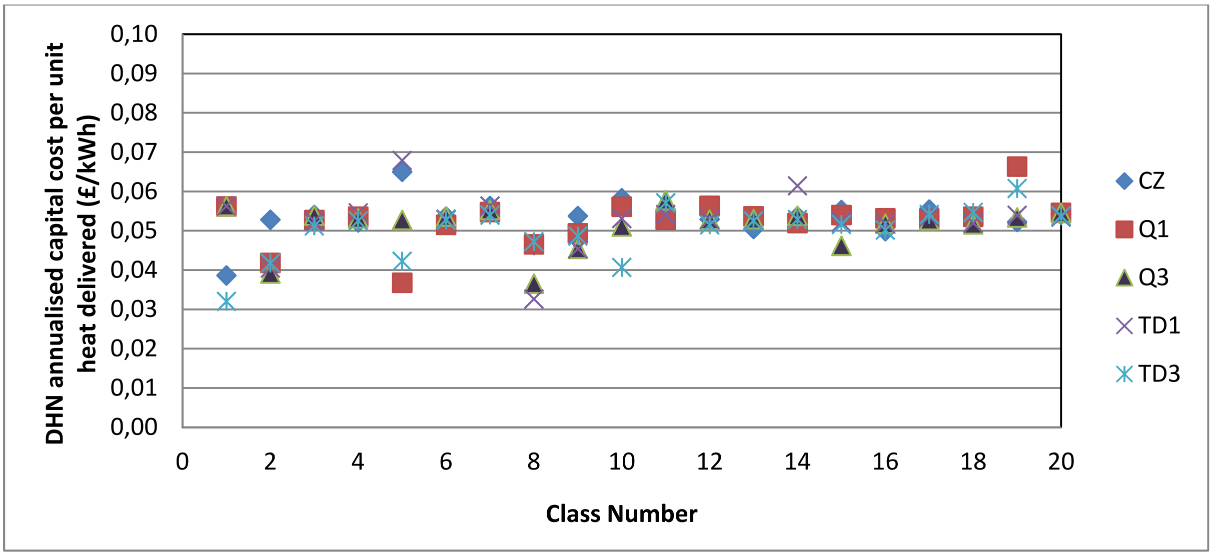


Figure H-3: DHN annualised capital cost per unit of heat delivered

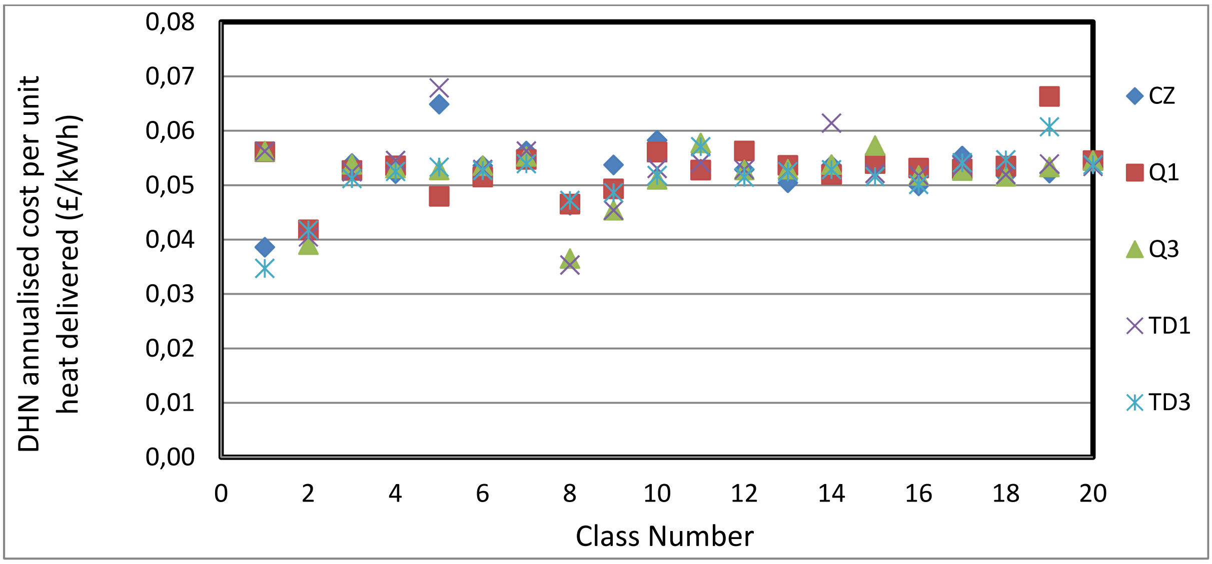


Figure H-4: DHN annualised capital costs recalculated: zones in each class are taken to be of the same 'type'

Figure H-5 presents the cost of the DE solution (DE centre and DHN) per unit of heat delivered. As noted in Table 5-7, the differences in cost of greater than 20% are observed for classes 1, 5 and 10.

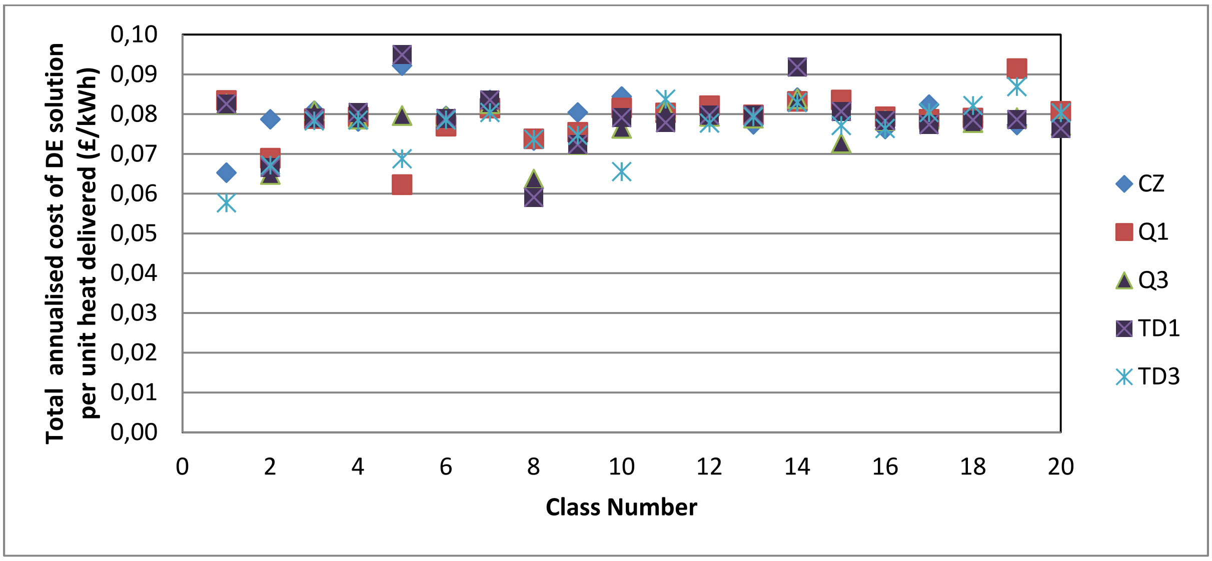


Figure H-5: Total annualised cost of DE solution (DE centre and DHN) per unit of heat delivered

Figures B.6 to B.7 present results relating to CO_2 emissions per unit of heat delivered. It may be observed that the performance within each class (as well as across all classes) is relatively similar, mainly because in all cases natural gas was the only fuel selected for heat generation. Figure H-8 presents 'global' CO_2 emissions results; significant scatter exists because the numbers being subtracted are similar. Figure H-9 (and Figure 5-16) confirm that, within each class, the contribution of the CHP units to heat production is similar, i.e. that the DE centre design is similar for the five zones within each class.

Figure H-6: Fuel CO_2 emissions per unit of heat delivered

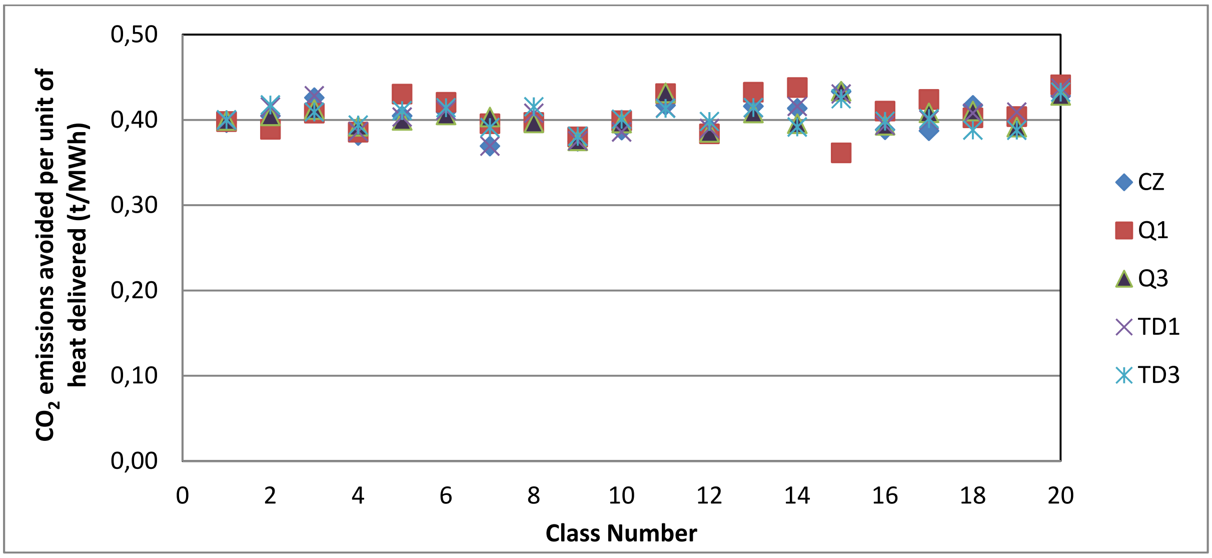


Figure H-7: CO_2 emissions avoided per unit of heat delivered

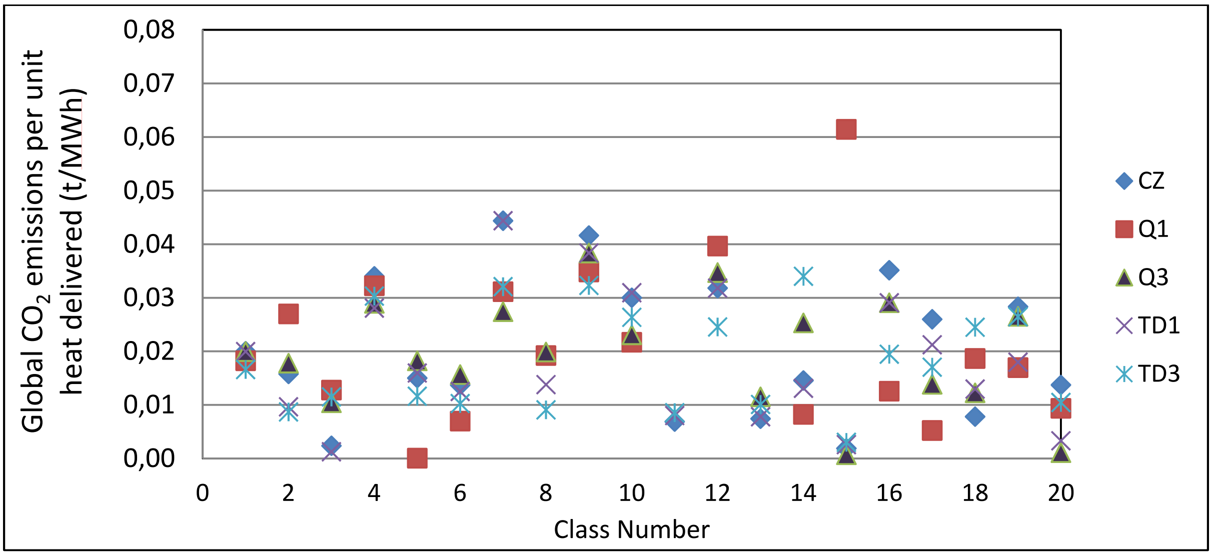


Figure H-8: Global CO_2 emissions per unit of heat delivered

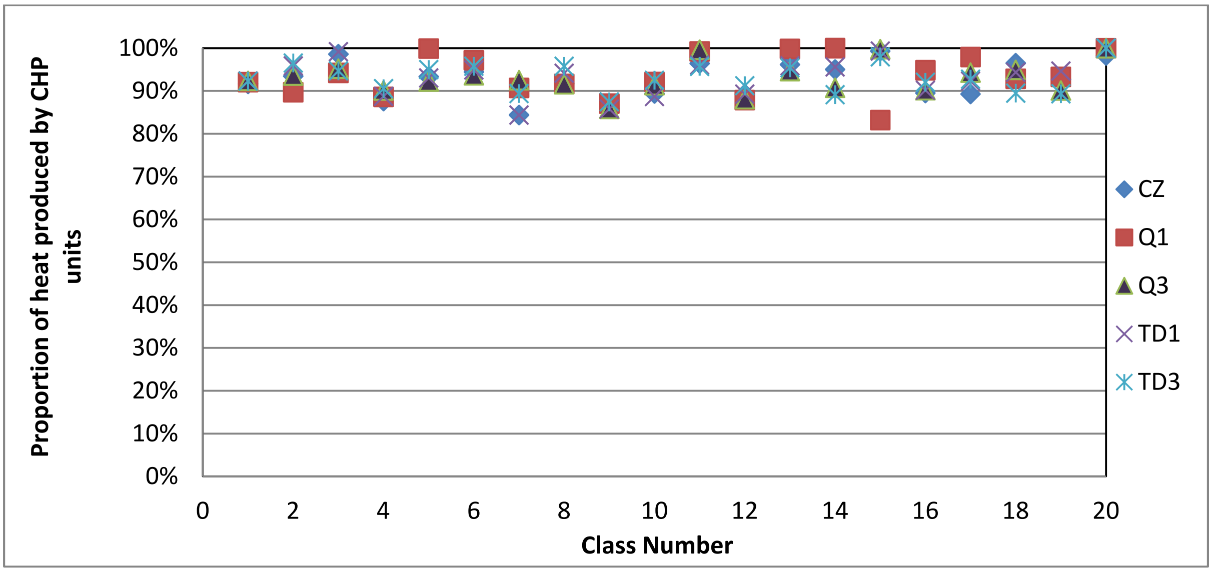


Figure H-9: proportion of heat delivered that is produced by CHP units

Appendix I — Contributions of heat interface units to DHN cost

The simple DHN cost model presented in Table 5-3 includes piping costs, which depend on the heat demand and heat density, as well as costs of heat interface units, which depend on the number of domestic and tertiary users of each type. Figures C.1 and C.2 present the contributions of the heat interface units to annualised DHN costs.

Figure I-1 demonstrates that the cost of domestic heat interface units is generally around £0.01/kWh, based on total thermal demand of the zone; the costs are relatively similar for the five zones investigated in each class. Figure I-2 shows that the cost of tertiary (commercial and industrial) heat interface units is relatively low, at less than £0.005/kWh based on the total thermal demand of the zone. Where the number of tertiary meters varies widely within in a class, the contribution of tertiary heat interface units to the overall cost of the district heating network can vary considerably, by £0.01/kWh to £0.02/kWh.

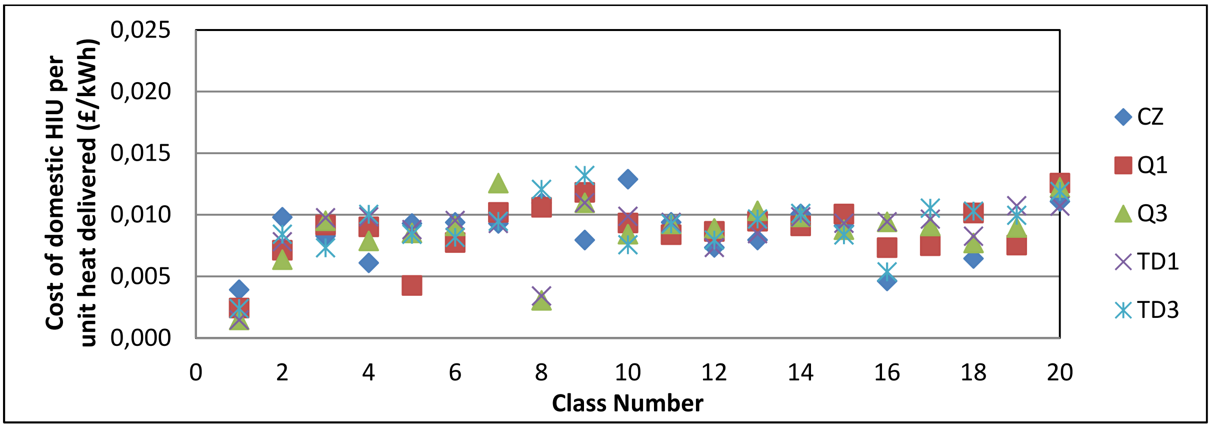


Figure I-1: Cost (per units of heat delivered) of domestic heat interface units

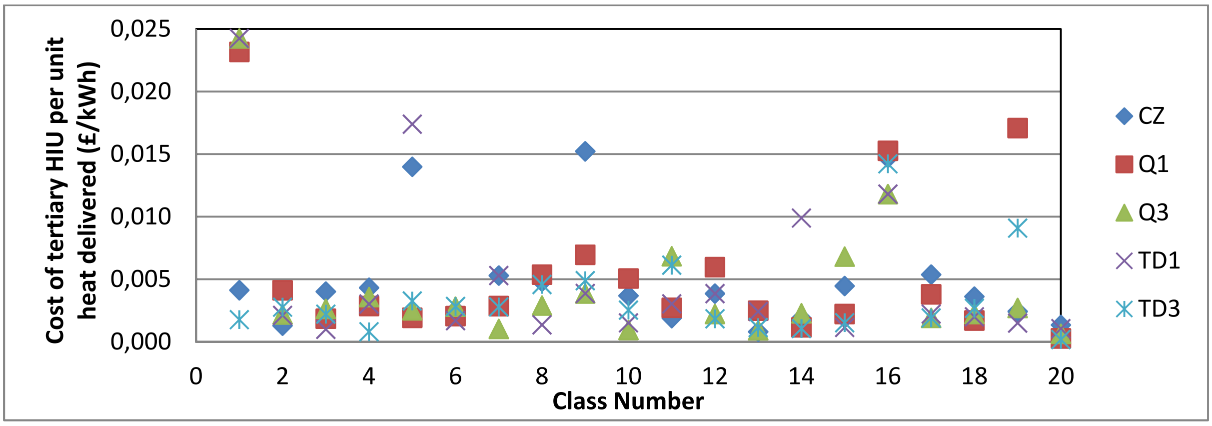


Figure I-2: Cost (per unit of heat delivered) of tertiary heat interface units

Appendix J — Outliers

J.1 Determination and identification of outliers and data anomalies

After calculating the annual energy demand of all MLSOAs in GB (see Chapter 1 — Energy Demand Calculation), 19 MLSOAs have been identified that show an abnormal behaviour towards their tertiary demands. Their heat demands were estimated larger than they should have been. Therefore, these MLSOAs can be found outside of the general set of MLSOAs or DE zones and are referred to as outliers.

Although, generally speaking, most outliers can be found in city centres, three different groups can be identified:

1) MLSOAs with a higher total heat demand than 500 GWh[36]. This value has been defined as the possible maximum demand of a DE zone that can be satisfied by a DE solution with an output of 50 MW_e. In total, 11 MLSOAs exceed this threshold level and therefore are excluded from the clustering and the classification processes as single MLSOA zones. Their demands vary between slightly above 500 GWh and about 3,000 GWh which is the factor of 6 of the maximum zonal demand (see Figure J-1).

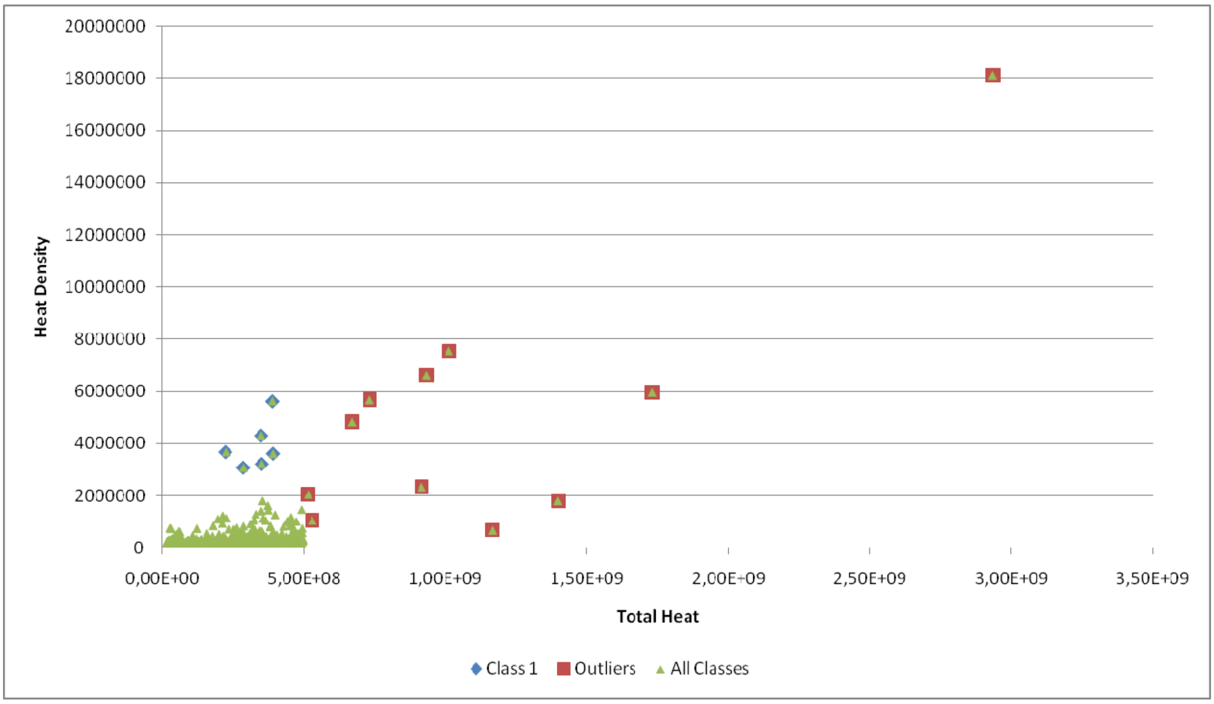


Figure J-1: Scatter plot of total heat (kWh) and heat density (kWh/ha) showing all created DE zones (green triangle), the 11 member of the outlier group 1 (more than 500 GWh total heat demand) and the 6 members of the outlier group 2 (Class 1) (blue diamonds)

[36] This value has been defined earlier in the project and will be discussed in section 2.2 Input Parameters.

2) During the classification process, class one has been created out of 6 outlying zones[37], consisting of 7 MLSOAs. These MLSOAs (or zones) differ from the rest of the total set of DE zones by their extraordinary high energy density (see Figure J-1). Although these zones fulfil all criteria that are required in this project for DE zones, there energy density excludes them from the main set and therefore it has been decided to take a second view at the calculated tertiary demand of these zones.
3) Two MLSOAs have been indentified whose numbers of the employees are significantly increased during the re-distribution process (see section 1.2.4). In one case, nearly 300,000 employees get re-distributed into this MLSOA. As these figures are challenging the improvements that have been achieved by the re-distribution (see section 1.2.4), these two MLSOAs will be analysed in details and treated as outliers before they are taken back into their original DE zones.

Table J-1 below shows the name of the outliers, the region they are in and the group of outliers they belong two.

MLSOA Name	GOR	Outlier Group
Birmingham 130	West Midlands	3
Bracknell Forest 005	South East	1
Bradford 038	Yorkshire and Humber	1
Broxbourne 010	East of England	1
Camden 028	London	2
City of London 001	London	1
Islington 021	London	2
Islington 022	London	2
Leeds 084	Yorkshire and Humber	1
Newcastle upon Tyne 024	North East	1
Sutton 012	London	3
Tower Hamlets 029	London	1
Wakefield 022	Yorkshire and Humber	1
Westminster 011	London	2
Westminster 013	London	1
Westminster 015	London	2
Westminster 018	London	1
Westminster 020	London	1
Westminster 021	London	2

Table J-1: Summary of the outlier MLSOAs, their regions, outlier groups and the DE zone they belong to

J.2 Explanation for outlier MLSOAs

The reason for these 19 MLSOAs to be outlying can be explained by data anomalies in the tertiary demand data. As the calculation is based on the number of employees, inaccuracies can be seen within the Marketsafe data that refer back to the company reports, the database is based on. Limited

[37] See classification section for details

companies with many small subsidiaries (e.g. food store chains or banks) tend to centre their employees at their headquarters (see Table J-2) which leads to an overestimation of the tertiary heat and electricity results in the city centres and to an underestimation in the outer areas. The underestimation is not as significant, because the affected MLSOAs were excluded from the zone creation by their low energy demand and energy density. Even a small increase of the employee numbers in these areas and thus a higher tertiary demand would not change the results.

Company name	Original Number of employee	Re-distributed number of employees
TESCO PLC	472094	472094,00
TESCO STORES LIMITED	257283	99872,44
SANDERS SUPERMARKETS LIMITED	0	99872,44
TESCO DISTRIBUTION LIMITED	13904	13904,00
HALESWORTH SPV LIMITED	0	13736,75
FLITWICK PHARMACIES LIMITED	0	4000,99
TESCO DISPENSING LIMITED	0	4000,99
WEYMOUTH AVENUE (DORCHESTER) LIMITED	0	4000,99
BROUGHTON RETAIL PARK NOMINEE 1 LIMITED	0	2019,55
BOSCOMBE RESIDENTIAL MANAGEMENT COMPANY LIMITED	0	2019,55

Table J-2: Top ten employers in one MLSOA in Broxbourne according to the Marketsafe data after re-distribution

Secondly, global companies, especially banks and insurances, seem to prefer continental headquarters in GB as is shown in Table J-3 for the top 10 employers one MLSOA in Tower Hamlets.

Company name	Original Number of employee	Re-distributed number of employees
BARCLAYS PLC	153800	153213,12
BARCLAYS BANK PLC	153800	112062,11
HSBC FINANCE (NETHERLANDS)	0	44685,83
HSBC OVERSEAS INVESTMENTS (UK) LIMITED	0	44685,83
HSBC LATIN AMERICA HOLDINGS (UK) LIMITED	0	44685,83
HSBC HOLDINGS PLC	309516	44685,83
HSBC LONDON HOLDINGS LIMITED	0	38958,26
HSBC OVERSEAS HOLDINGS (UK) LIMITED	0	22342,92
HSBC RETIREMENT BENEFITS TRUSTEE (UK) LIMITED	0	13015,29
CARNEGIE HOLDINGS LIMITED	0	6584,90

Table J-3: Top ten employers of one MLSOA in Tower Hamlets according to the Marketsafe data after re-distribution

The last indications for data anomalies in these MLSOAs are double or even triple counted employees. This occurs if a company has changed its name within the period of data accumulation or if a holding company is reporting its total number of employees for itself as well as its subsidiaries (see Table J-4).

Company name	Original Number of employee	Re-distributed number of employees
BROADSTREET GREAT WILSON EUROPE LIMITED	168666	168666,00
WAL-MART STORES (UK) LIMITED	168666	168666,00
ASDA STORES LIMITED	168666	139489,84
ASDA GROUP LIMITED	276673	139205,33
PORTH INVESTMENTS LIMITED	0	40545,24
POWER4ALL LIMITED	0	39193,73
ASDA C S LIMITED	0	29176,16
ASDA STORAGE LIMITED	0	28767,81
ASDA EMPLOYEE SHARE SCHEMES TRUSTEE LIMITED	0	13322,01
INTERNATIONAL PERSONAL FINANCE PLC	9454	9454,00
ASDA QUEST TRUSTEES LIMITED	0	6564,47

Table J-4: Top ten employers in one MLSOA of Leeds according to the Marketsafe data after re-distribution. The number of employees of Broadstreet Great Wilson Europe Limited, Wal-Mart Stores (UK) Limited and Asda Stores Limited are triple counted before the re-distribution and still extremely high after redistribution

J.3 Approach to solve the outlier issue

These inaccuracies in the tertiary demand calculation can be corrected. Two options have been developed:

a) Delete double counted employees or significantly high numbers of employees within one company in one MLSOA and re-calculate the tertiary demand.
b) Exchange the calculated demand with the gas demand for tertiary and industrial sites provided by DECC. The gas demand needs to be converted into heat demand, for which a conversion factor of 0.8 has been assumed. Although the demand measured by DECC is by far lower than the calculations based on the Marketsafe data for the outlier MLSOAs, there is still the potential of over-estimating the tertiary demand because of the inclusion of industrial gas demand. As these outlier MLSOAs can mostly be found in city centres, it has been assumed that the potentially included industrial gas demand is low and can be ignored.

It has been decided to exchange the tertiary demand calculated with the Marketsafe data with the industrial and tertiary gas demand for these outlier MLSOA. The methods will not be changed for all other MLSOAs. Thus, these 19 MLSOAs and their corresponding DE zones will be taken out of the further process at different point in time and reintegrated into the classes again at the end.

J.4 Impact of the outliers

The impact of the outliers is rather small. They are covering a share of 0.4 % of all MLSOAs that have been grouped together into DE zones (see Table J-5, column 2). This share is confirmed by the number of households and the calculated domestic heat demand for the outlier MLSOAs (see Table J-5, column 3 and 4). However, the calculated tertiary heat demand is incredibly high for these 19 MLSOAs, because they are covering nearly one quarter (23.7 %) of the total tertiary heat demand within all DE zones (see Table J-5, column 5). This number can be directly referred to the number of employees that are reported in these MLSOAs by the local companies, which share is with 25.6 % even higher than the tertiary demand share (see Table J-5, column 6). These corrupted employee data for the outlier MLSOAs can be seen when compared with the number of households (see Table J-5, column 4). In the area of these 19 MLSOAs 62,566 households are reported, but more than 9 million employees are working in the same area. This is a ratio of more than 148 employees per household. The figure for the tertiary sites in the area emphasises this statement (see Table J-5, column 7). For every household more than 3 tertiary sites are registered, which is a total share of nearly 9 % of all tertiary sites in all DE zones. It is lower than the share of employees, but still a large number for only 19 MLSOAs.

Hence, it has been tested if the tertiary demand of only these 19 MLSOAs can be replaced by other data, e.g. the gas demand for tertiary and industrial sites reported by DECC after conversion into heat demand. DECC is measuring gas demand of tertiary and industrial consumers, thus in general it is likely to be higher than only the tertiary demand. On the other hand, these outlier MLSOAs are mostly city centres and thus it is unlikely that large industrial sites can be found within these areas. The results are shown in Table J-5 as well (see column 8 and 9 of Table J-5). The calculated tertiary demand for all non-outlier zones including the converted gas demand from DECC for the outliers zones are showing a calculated heat demand of these outliers MLSOAs of 8.1 % (see Table J-5, column 8), which is very close to the share of the tertiary sites. If this calculated tertiary demand is added to the domestic demand than the outliers are covering a total demand share of 1.89 % (see Table J-5, column 9), instead of 5.61 % in the previous calculation (see Table J-5, column 10). Taking into account that the base line for all calculations here is about 50 % of the GB tertiary and domestic demand, than these figures will be reduced to about 1 % or 2.8 %. Thus, the district heating potential of each outlier MLSOAs can be more reliably estimated using the gas demand data for tertiary and industrial sites from DECC than the method using the Marketsafe data. For all other MLSOAs the energy demand calculation method will not be changed.

	# MLSOA	Heat Dom (kWh)	House-holds	Heat Tert (kWh)	Employees	Tert Sites	Tert DECC**	Total** DECC	Heat Tot (kWh)
Outliers*	19	8,43E+08	62,566	1,37E+10	9,276,663	191,587	3,87E+09	4,71E+09	1,45E+10
All Zones	4,660	2,01E+11	13,612,339	5,77E+10	36,176,724	2,130,859	4,79E+10	2,49E+11	2,59E+11
Ratio of Outliers	0.4 %	0.42 %	0.46 %	23.74 %	25.64 %	8.99 %	8.1 %	1.89 %	5.61 %
* including the 6 DE zones of Class 1 ** using DECC data for outliers and own calculation for other MLSOAs/zones									

Table J-5: Comparison of outlier MLSOAs to all MLSOAs in all zones

J.5 Conclusion

The impact of the 19 outlier MLSOAs on the national feasibility study for district heating in GB is very limited. These MLSOAs are covering a total demand share of less than 1.9 % of the total potential demand, which is less than 1 % of the total GB heat demand. The impact on the national benefits case is therefore very small.

On the other hand, it is not deniable that most of the outlier MLSOAs show a high potential for district heating, although their demand calculations are overestimated. Most of these 19 MLSOAs are located in city centres, 12 of them are even MLSOAs of London. There will be only a few MLSOAs where the demand is cut down so dramatically that it is uneconomic to install a district heating network. The outlier's exact potential will be analysed in a small follow-on study for all 19 MLSOAs separately.