Clustered spatially and temporally resolved global heat and cooling energy demand in the residential sector

Julia Sachs[1], *Diego Moya[1, 3, 4], Sara Giarola[2] and Adam Hawkes[1]

[1]Department of Chemical Engineering, Imperial College London, South Kensington, London SW7 2AZ, UK

[2]Department of Earth Science and Engineering, Imperial College London, South Kensington, London SW7 2AZ, UK

[3]Carrera de Ingeniería Mecánica, Facultad de Ingeniería Civil y Mecánica, Universidad Técnica de Ambato, Av. Los Chasquis y Rio Payamino, Ambato, 1801314, Ecuador

[4]Institute for Applied Sustainability Research, Av. Granados E13-55 e Isla Marchena, No.44, Quito, 170503, Ecuador

*Corresponding author: d.moya17@imperial.ac.uk, da.moya@uta.edu.ec

Abstract

Climatic conditions, population density, geography, and settlement structure all have a strong influence on the heating and cooling demand of a country, and thus on resulting energy use and greenhouse gas emissions. In particular, choice of heating or cooling system is influenced by available energy distribution infrastructure, where the cost of such infrastructure is strongly related to the spatial density of the demand. As such, better estimation of the spatial and temporal distribution of demand is desirable to enhance the accuracy of technology assessment. This paper presents a Geographical Information System methodology combining the hourly NASA MERRA-2 global temperature dataset with spatially resolved population data and national energy balances to determine global high-resolution heat and cooling energy density maps. A set of energy density bands is then produced for each country using *K-means* clustering. Finally, demand profiles representing diurnal and seasonal variations in each band are derived to capture the temporal variability. The resulting dataset for 165 countries, published alongside this article, is designed to be integrated in a new integrated assessment model called MUSE (ModUlar energy systems Simulation Environment), but can be used in any national heat or cooling technology analysis. These demand profiles are key inputs for energy planning as they describe demand density and its fluctuations via a consistent method for every country where data is available.

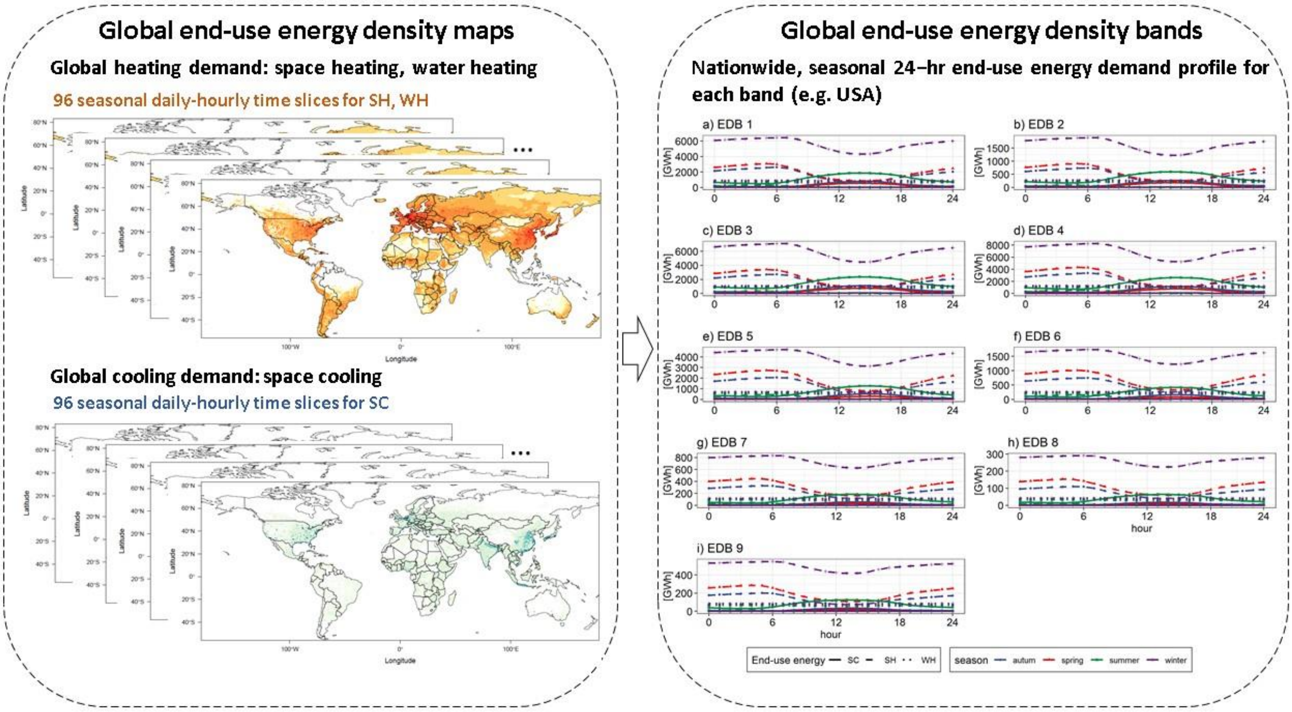
Highlights

- A high-resolution spatio-temporal approach for estimating global heat and cooling energy demand.
- K-means clustering for deriving energy density bands of each heat and cooling energy demand.
- Open-access data for spatial energy density bands for 165 countries covering 99.96% global energy users.
- 5% of heat demand is at very high energy densities worldwide, while >50% is at very low density.

Keywords

Heat demand; cooling demand; spatially resolved; temporally resolved; energy systems; integrated assessment

Graphical abstract



Nomenclature

DES	District energy systems
DHC	District heating and cooling
DH	District heating
DC	District cooling
MUSE	ModUlar energy systems Simulation Environment
IAM	Integrated assessment model
GIS	Geographic information system
HDH	Heating degree hours
CDH	Cooling degree hours
SH	Space heating
SC	Space cooling
WH	Water heating
$_c$	Cell
$_d$	Day
$_h$	Hour
$T_{c,d,h}$	Temperature in a cell for a specific day and hour
T_{ref}	Reference temperature
WH	Waiting factor for Water heating
kc	scaling factor
POP_{den_c}	Population density per cell
I_{SH}	Index for space heating
I_{SC}	Index for space cooling
I_{WH}	Index for water heating
E	Energy consumed
ED	Energy density
E_{SH}	Energy demand for space heating
E_{WH}	Energy demand for water heating
E_{SC}	Energy demand for space cooling
GWh	Gigawatt hours
km^2	Square kilometre
EDB	Energy density band
ONC	Optimal number of clusters
EM	Elbow Method
W_k	Average internal sum of squares
k	Number of clusters
n_r	Total members points in the cluster
R	Cluster
D_r	Sum of distances between points
d_i and d_j	Points belonging to a cluster

1. Introduction

The residential sector accounts for almost a third of final energy consumption globally and is an equally significant source of CO_2 emissions [1]. In particular, space and water heating account for approximately half of that consumption and are usually dominated by fossil fuels, thus representing a significant opportunity to reduce CO_2 emissions. Furthermore, cooling demand is growing rapidly both in high-income countries and in emerging economies such as India and China [2], where the electricity that powers these technologies is currently relatively emissions-intensive, [3]. To address the challenge of reducing building sector emissions, a number of advances in building design, technology, and policy have emerged, making it possible for energy use and emissions to decline significantly [4]. However, this will require dramatically increased use of low carbon technologies, alongside development or expansion of associated infrastructures and supply chains.

Energy systems models have emerged as a key tool to investigate different plausible climate change mitigation strategies. Several such models indicate that conventional unabated fossil-fuelled heating technologies must be replaced by lower carbon technologies if climate change mitigation targets are to be met [5]. Therefore technologies such as district heating and cooling (DHC), heat pumps, hydrogen-fuelled boilers or micro-combined heat and power will become more prevalent. When considering these technology options, it becomes apparent that the magnitude, and spatial and temporal variability of heat and cooling demand are of significant importance. This is because the costs of the infrastructure associated with each technology solution (e.g. district heating networks, gas/hydrogen networks, electricity networks) is an important determinant of its relative advantage [6]. As these costs depend upon the length, capacity and topology of those networks, detailed data on the spatial and temporal distribution of heat and cooling energy demand is needed. Space heating (SH), space cooling (SC) and water-heating (WH) demands are highly dependent on geographic location due to different climate conditions, population densities, economic situations and differences in heating practices. Understanding these variations is a central part of understanding which heating and cooling system of technologies and infrastructure will be most effective in each location.

The assessment of energy demand density is a significant factor in overall energy investments due to associated infrastructure costs (e.g. distribution pipes for district heating and cooling systems). Thus, energy demand density assessment plays an important role in seeking transition pathways to decarbonise the residential sector. The end-use energy demand density is the total demand for a heating or cooling service for a defined area; street, neighbourhood, city, country or region [7]. To estimate the end-use energy demand density of a defined area, studies have begun to use gridded population density data [8], combined with temperature profiles and other data to derive spatially-resolved energy demand [9]. Table 1 categorises these studies in terms of the approach used, the resolution of data inputs and results, and consideration of demand density. Most of these studies were focused on Europe, USA and China. The main approaches are the use of national energy balances such those presented by the International Energy Agency (e.g. India building energy demand estimations [10]), the national energy authorities (e.g. Europe's heat roadmap [11], USA and China's building energy use [8], Denmark's district heating assessment [12] and German's residential building heating and cooling demand [13]), the United Nations database as in [14], and other research initiatives as the use of climate model's output of mean monthly temperature [15], and the use of energy demand surveys [16]. In contrast to heating demand studies, cooling demand studies are rarely found in the scientific literature. Table 1 also presents relevant studies that estimate the demand for space cooling at country and regional level. Based on the reported literature, space cooling demand estimates are presented in more detail for the European Union at the country and regional level (e.g. nationwide estimations [11], accounting for measurement of 80 weather stations across UE [17], and at sub-national scale estimates [18]). Overall, a systematic assessment of the spatio-temporal variations in end-use heat and cooling demand for the world at the country scale was not present in the literature prior to this study.

Table 1: Summary of GIS-based end-use energy density assessment studies.

Country/region	**Approach**	**Resolution and/or Study level**	**Demand density considerations**	**End-use energy demand estimation**	**Ref.**
Global	Spatio-temporal	1 km^2	Heating and cooling demand density	SH, SC, WH	This study
Global	GIS Spatial	55 km x 55 km, for mean monthly temperature	–	SH, SC	[15]

Global	Regional demographic distributions	Regional level 11 regions worldwide	–	SH, SC	[14]
Europe	DC estimations	Regional level	Average cooling density demand	SC	[17]
EU28	GIS Spatial	City level	Population density	SC	[18]
EU28	National energy consumption	Country and Regional level	–	SH, SC, WH	[11]
Germany	GIS Spatial	Country level m^2	Spatially distributed residential density	SH, SC, WH	[19]
Germany	National statistics	Country and Regional level	Spatially distributed population density	SH, SC	[13]
UK	GIS Spatial	City level	Linear heat density demand	Heating	[6]
Denmark	GIS Spatial	Country level, results in million m^2	Heat demand density	SH, WH	[12]
Denmark	Spatio-temporal	District level	Population density	SH	[20]
Netherlands	Spatio-temporal	Municipality level	–	Heating	[21]
Switzerland	Spatio-temporal	Municipality to country level 1 km^2	Population density	SH	[9]
India	National statistics	Country level	–	SH, SC, WH	[10]
USA	Temporal for HDD/CDDs	Country level	Population density	SH, SC	[8]
USA	GIS Spatial	0.56 km^2 and 0.78 km^2	Heat demand density	SH, WH	[16]
USA	temporal	Building level	–	SH, WH	[22]
China	National statistics	Country level	Population density	SH, SC	[8]
China	GIS Spatial	Building to city level	Population density	SH	[23]

Advances in Geographic Information System (GIS) methodologies have created an opportunity to assess the end-use energy demand of a city or a country in a spatially and temporally-explicit manner. One example of this is the increasing use of GIS-based methods for analysing the potential of district energy systems (DES). Several GIS-based studies have been carried out on a local scale to identify suitable agglomerations of buildings. Studies include a range of approaches from optimization of the distribution network, surface modelling, and multispectral imagery, to satellite data and use of ancillary street vectors [24]. However, the large amount of required data makes this method challenging to apply at a global scale. For example, the small-scale study in [25] focuses on the optimization of the piping network in areas with high potential for district heating by mapping building characteristics, resource availability, and road networks. Likewise, suitable configurations of buildings for thermal microgrids based on analysis of geo-referenced building stock data is carried out in [26]. In that case, the linear heat density is estimated, taking into account the heat delivered to a building and the length of the pipe route. A further assessment of the building energy consumption for the Centro Residential Europa area of Turin–CRE is carried out in [27], where the space heating demand is evaluated using a thermal model, where a geo-referenced representation of the DH network enables the application of an optimization of the network layout. In [28], a geographic information-based mixed integer linear programming model for cost assessment of a DH network in northern Japan is presented. Although these studies have addressed the spatial density of the heating demand, the proposed methodologies (1) lack accounting for diurnal temporal variability, and (2) are only suitable at a small scale. Moreover, despite the increased research into the evaluation of DH potential, existing assessments almost universally focus on countries in Europe and the USA, and consideration of cooling demand is rarely undertaken.

Also of relevance to this study is the treatment of spatial detail in Integrated Assessment Models (IAMs). This is a class of global models are used to capture possible trends in the energy (and climate, land, etc) system, to guide decision makers on setting targets and to indicate the strategies that may be most effective in achieving them. The uptake and potential of energy-technologies is difficult to estimate due to a wide range of factors including energy price, capital cost, technical applicability, social acceptance, policy and regulatory uncertainties, to name but a few. Taking note of these uncertainties, IAMs strive to capture the interaction between the energy sectors, the demand for end-uses, the technologies available and the diffusion of those technologies in the market [29]. Some IAMs include spatial and temporal detail, but many lack a logical connection between the spatial aspects of energy consumption and the suitability of technologies.

Strachan, et al. [30], for example, integrate a more detailed spatial analysis of supply, demand and infrastructure requirements regarding hydrogen uptake in the UK. The methodology they present is integrated within the energy system model MARKAL. It focuses on the use of hydrogen as a transportation fuel and spatially locates the supply and demand centres. Another effort is apparent in highRes (high-Resolution Electricity Model), which analyses the development of the electricity sector, including a spatially and temporally-detailed renewable generation time series, to accurately determine electricity production and dispatch [31]. This model can be linked to a TIMES model to test the feasibility of long-term investment decisions in TIMES. Another study using spatially resolved information as an input to energy system models is presented in [32]. This uses spatial data for 2.5 million buildings in Denmark along with individual heating supply data to develop a country-wide heat demand atlas. Also, a heat atlas for Denmark [33] is used to provide inputs to an energy system model in the form of marginal cost curves for heat saving measures and district heating networks.

Looking at the residential sector and in particular at the role of demand density in IAMs, a focus on different core characteristics can be observed. MARKAL and TIMES are bottom-up, technology-rich intertemporal cost optimisation models covering any geographical area desired [34]; national, regional, state, province, or community level, [35]. Any MARKAL or TIMES model is able to consider any range of heating technologies, but implementations often do not include a spatially distributed assessment of their potential [36]. Another example is GCAM (Global Change Assessment Model) which includes a representation of demand density in various sectors of the energy system [37]. In this case, however, the market share of the technologies is determined based on a logit-share function on the investment cost, without separately considering additional cost for energy distribution systems.

NEMS (National Energy Modelling System) is an energy system model representing energy markets, demand and economic interactions for the USA. Here, investments in new technologies are split in retrofits, including fuel-switching and in energy systems of new buildings. The share of technologies in the system is determined using a log-linear function based on bias, capital costs and operating costs [38, 39]. Although NEMS is a technology-rich model, the spatio-temporal demand density phenomena and associated infrastructure costs are not considered [39]. In the PRIMES model, the building-sector follows a nested logit model approach to determine the technology share based on the cost of each technology, which is modified by the households' income and equipment maturity [40]. These approaches consider several energy technologies for buildings. However, the expansion of these systems in the energy market is only determined by the logit function, without direct consideration of spatio-temporal demand density characteristics and thus infrastructure costs.

Overall, it is clear that while researchers have begun to include the impact of spatial demand detail in energy systems models, many do not take into account both spatial demand density and the temporal variability of that demand. There is a further need for mathematical descriptions and computational methods to enable research teams to explore this issue, generate temporally-explicit energy demand and spatial energy density maps, and thus improve estimates of cost-effective uptake for heating and cooling technologies and their associated infrastructure. With this paper the authors aim to begin to fill this gap.

In summary, this paper presents such an analysis, producing a spatially and temporally-resolved worldwide atlas of space heating, water heating and space cooling demand, and deriving associated energy density bands for each of these service demands for each country. Results, which include the diurnal and seasonal profile of demand in each band and for each country, are published alongside this article. The approach is validated based on cross-checking against the International Energy Agency (IEA) Energy Balances for the residential sector. The novelty and contributions of this paper are as follows:

(1) A systematic, spatio-temporal data-driven approach for estimating end-use energy demand based on a combination of energy data from IEA Energy balances [41], hourly temperature data from NASA MERRA-2 [42], population counts [43] and population density [44] at 1 km^2 resolution. Interpolated temperature profiles are used to calculate heating degree hours (HDH) and cooling degree hours (CDH) for each location, season and hour, for each country in the world. This is combined with population density data and energy balance data to yield time-sliced energy demand by location, ultimately generating spatio-temporal energy demand raster data for 165 countries, covering 99.96% of global energy users in residential sector.

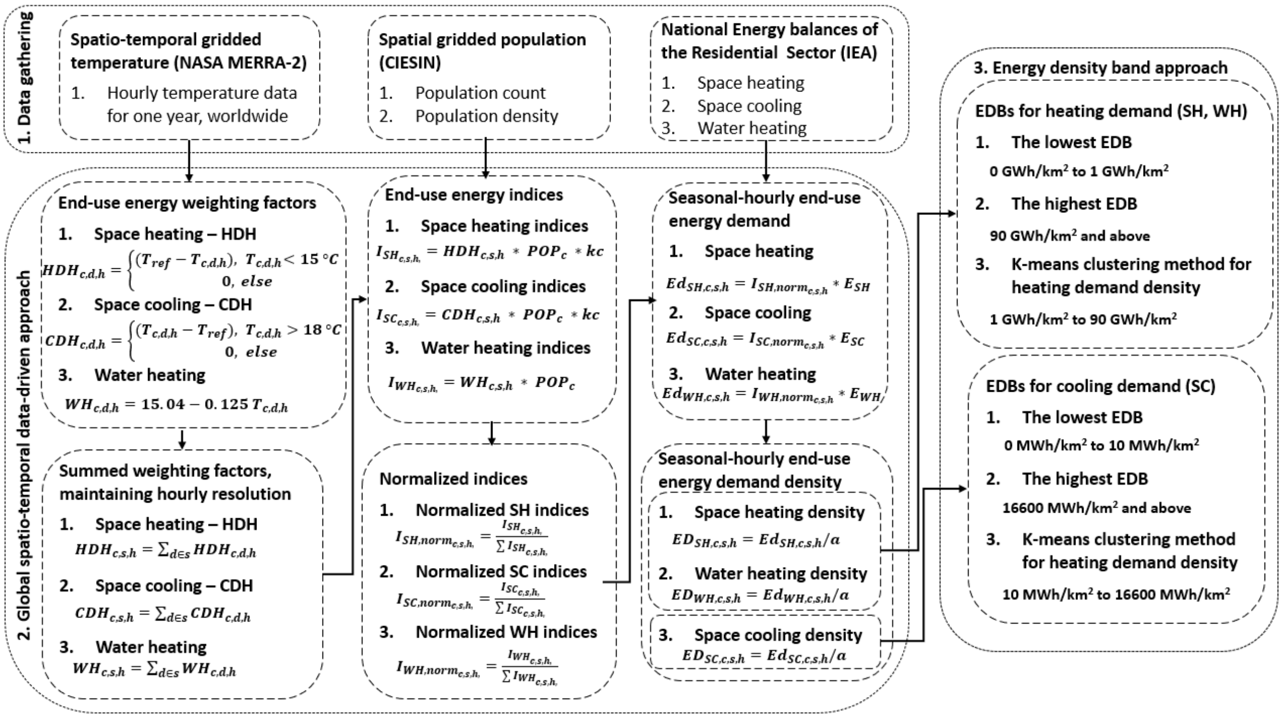
(2) A spatial clustering analysis for deriving the energy density bands (EDBs) of each end-use energy demand via the use of *K-means* techniques. Here, *K-means* clustering techniques are used to find distinctive spatial areas with similar energy demand characteristics. Thus lower and an upper limits are mathematically determined applying this spatial clustering approach. The heating demand is assessed separately from the cooling demand to any implicit assumption that these two types of demand are served via common infrastructure. Details of the method are presented below.

(3) Significant temporal detail is retained in the resulting energy density bands. After banding by energy density, the temporal profile of each band is derived based on the locations that were binned into each band. Results are presented using a 24-hour profile for each of four seasons, for each country.

The outputs of this analysis are intended to be integrated in the residential sector module in the MUSE[1] framework, but can be equally used for input in other energy systems models and are made freely available for the community as supplementary material here. The remainder of this paper is organised as follows. Section 2 describes the datasets and clustering methods used in this study. Section 3 gives a discussion of results followed by limitations, conclusion and further research provided in Section 4.

2. Methodology

This paper presents a GIS-based approach to spatially and temporally assess the global energy demand and energy density for all heating and cooling end-uses. Based on the total energy consumption taken from the IEA Energy balances, population density gridded maps, and national weather conditions, heat density maps of the world with a grid of $1 \times 1 \text{km}^2$ cells are generated. These heat density maps are used to identify the amount of energy consumed by each end-use in a set of energy density bands, which are in turn determined via a *K-means* clustering method. By using GIS, it is possible to conduct a study with high spatial resolution on a global scale, which is necessary in order to build a consistent model. This assessment is carried out in four main steps and schematically illustrated in Figure 1.



[1] MUSE is the ModUlar energy systems Simulation Environment, an integrated assessment model (IAM) developed at Imperial College London. MUSE is a bottom-up technology-rich model of the global energy system, focusing on the simulation of investment and operational decision making across the energy sector.

Figure 1: Schematic representation of the global spatial heating and cooling energy demand assessment per country for 2010. The gathered data contains three data sets: (1) spatio-temporal temperature variations of the world, from the NASA MERRA-2 program, at a spatial resolution of 0.625 x 0.5 degrees at the equator, which are used to determine the heating degree hours and cooling degree hours of a cell [42]; (2) the Gridded Population of the World data, which models the distribution of human population (counts [43] and densities [44]) on a continuous global raster surface using population census tables and corresponding geographic boundaries at 1 km^2 spatial resolution; and (3) the total energy consumption of the residential building sector from the International Energy Agency Energy Balances [41]. Two main data processing steps are described. First, for the global temperature data in combination with population data and energy balances data, global energy density maps along with seasonal-hourly demands are calculated for space heating, space cooling and water heating demands. Second, the global spatio-temporal datasets are clustered into energy density bands to find areas with similar demand density characteristics and demand profiles in each band.

2.1. Energy density maps

In the first step, the energy demand for the main heating and cooling end uses is determined. To obtain the demand for the different end-uses for each country the total energy consumption of the residential building sector from the IEA energy balances [41] is allocated to the main end-uses. Due to a lack of consistent household level data on energy consumption for all countries [45], an estimate of the energy consumption for cooling and heating purposes is determined by assuming a specific share of the total consumption of each end-use by region as indicated in [46]. This is shown for the example of the United Kingdom in Table 2.

Table 2: End-use energy share for the United Kingdom [46]. This is an example of the Energy Balances of the International Energy Agency for a selected country. In this study, the values corresponding to space heating, space cooling and water heating demands have been considered.

Region	Country	End-Use	Share [%]	Demand [$\frac{\text{Energy}}{\text{TWh}}$]
European Union	United Kingdom	Space heating	59	304.81
		Water heating	14	75.08
		Space cooling	3	16.52
		Lighting	6	30.03
		Cooking	3	15.02
		Appliances and other	15	78.08

To account for the spatial resolution of the heating and cooling demand, the country level end-use demand is refined based on two key drivers, the exterior temperature and the population in each 1 km^2 cell. The allocation of energy demand is carried out based on temperature dependent specific weighting factors for each end-use: water heating I_{WH}, space heating I_{SH} and space cooling I_{SC}. Key indicators of the energy consumption in each cell are the heating degree hours (HDH) and cooling degree hours (CDH), which are used to estimate the energy required to heat a building in a particular location. To determine the HDH and CDH of a cell, the freely available NASA MERRA-2 [42] dataset was used, which contains hourly temperature measurements with a spatial resolution of 0.625 x 0.5 degrees at the equator (approximately 64 km x 55 km). The HDH and CDH for cell c for a specific day d and hour h are calculated according to [47] by using Eq. (1) and Eq. (2):

$$HDH_{c,d,h} = \begin{cases} (T_{ref} - T_{c,d,h}), & T_{c,d,h} < 15\,°C \\ 0, & else \end{cases} \quad \text{Eq. (1)}$$

$$CDH_{c,d,h} = \begin{cases} (T_{c,d,h} - T_{ref}), & T_{c,d,h} > 18\,°C \\ 0, & else \end{cases} \quad \text{Eq. (2)}$$

where $T_{c,d,h}$ is the temperature and T_{ref} the reference temperature. The weighting factor $WH_{c,d,h}$ for water heating is assumed to be less dependent on the HDH and is defined according to [48] by using Eq. (3):

$$WH_{c,d,h} = 15.04 - 0.125\, T_{c,d,h} \quad \text{Eq. (3)}$$

The datasets are further summed by season s, where the hourly resolution is maintained:

$$HDH_{c,s,h} = \sum_{d\in s} HDH_{c,d,h} \quad \text{Eq. (4)}$$

$$CDH_{c,s,h} = \sum_{d\in s} CDH_{c,d,h} \quad \text{Eq. (5)}$$

$$WH_{c,s,h} = \sum_{d\in s} WH_{c,d,h} \quad \text{Eq. (6)}$$

The demand required for cooling or heating a building in a specific location is considered to be proportional to the number of HDH and CDH at the location. To map the population data[2] (counts [43] and density [44]) with the HDH, CDH, and WH rasters, the data in the grid cells need to be interpolated to match the accuracy of 1 km^2 of the population dataset. Further adjustments are taken into account for the reduced heat demand in cities by assuming a reduction for the per-capita heat demand in densely populated areas compared to rural ones according to the approach in [16], [49]. The indices for space heating and space cooling are multiplied by a scaling factor $kc = f(POP_{den_c})$ with $k \in [0.7,1]$ depending on the population density in a cell POP_{den_c}. This dependency on outside climatic conditions can be

[2] The Gridded Population of the World (GPW) collection, now in its fourth version (GPWv4), models the distribution of human population (counts and densities) on a continuous global raster surface using population census tables and corresponding geographic boundaries.

balanced by the housing conditions (namely building type, dwelling size, inhabitants per habitation). Since such data is not available on a global scale the SH, SC and WH indices are determined by combining the weighting factors and population data;

$$I_{SH_{c,s,h,}} = HDH_{c,s,h} POP_c f(POP_{den_c}) \quad \text{Eq. (7)}$$

$$I_{SC_{c,s,h,}} = CDH_{c,s,h} POP_c f(POP_{den_c}) \quad \text{Eq. (8)}$$

$$I_{WH_{c,s,h,}} = WH_{c,s,h} POP_c \quad \text{Eq. (9)}$$

where the population of the cell is POP_c. The indices are normalized to split the energy consumption of a country for the different end uses between all cells, while maintaining the overall sum:

$$I_{SH,norm_{c,s,h,}} = \frac{I_{SH_{c,s,h,}}}{\sum I_{SH_{c,s,h,}}} \quad \text{Eq. (10)}$$

$$I_{SC,norm_{c,s,h,}} = \frac{I_{SC_{c,s,h,}}}{\sum I_{SC_{c,s,h,}}} \quad \text{Eq. (11)}$$

$$I_{WH,norm_{c,s,h,}} = \frac{I_{WH_{c,s,h,}}}{\sum I_{WH_{c,s,h,}}} \quad \text{Eq. (12)}$$

The overall energy consumed in one cell for a country for each season and hour $E_{c,s,h}$ and energy density $ED_{c,s,h}$is defined by

$$E_{c,s,h} = I_{SH,norm_{c,s,h,}} E_{SH} + I_{SC,norm_{c,s,h,}} E_{SC} + I_{WH,norm_{c,s,h}} E_{WH} \quad \text{Eq. (13)}$$

$$ED_{c,s,h} = \frac{\left(I_{SH,norm_{c,s,h,}} E_{SH} + I_{SC,norm_{c,s,h,}} E_{SC} + I_{WH,norm_{c,s,h}} E_{WH}\right)}{a} \quad \text{Eq. (14)}$$

where the area is a, the energy demand for space heating is E_{SH}, water heating is E_{WH}, and space cooling is E_{SC}.

2.2. Energy density bands approach

Heat demand density and cooling demand density are used to describe and assess the related technologies, considering climatic and demographic conditions in each country [50]. This research classifies different energy demand density levels by implementing an energy density bands (EDB) approach. Firstly, the lower and upper energy density bounds for each EDB are defined using the GIS-based energy density global maps clustering results. Then, every cell within the bounds of each EDB is allocated to that EDB. All cells in an EDB are then summed up to determine the total energy demands within the band.

A range of energy density levels for the assessment of heating technologies are evaluated in [51] and [52]. The values considered in these studies are used to define the lowest and highest EDB at 0 GWh/km^2 to 1 GWh/km^2 and at 90 GWh/km^2 and above, respectively. Further banding between 1 GWh/km^2 and 90 GWh/km^2 is conducted by grouping energy density values into clusters using a *K-means* method. The clustering-based knowledge discovery method is applied to avoid arbitrarily dividing up the range of the energy demand density values.

Clustering analysis is an established technique for pattern classification into groups of similar characteristics denominated "clusters" which has been widely applied in the electricity sector for assessing spatio-temporal electricity consumption profiles [21]. Although a number of different clustering algorithms have been implemented in the energy field [53], Gianniou, et al. [54] found that the *K-means* algorithm has great potential in relation to energy demand. In this research, the *K-means* clustering is undertaken as follows. (1) From the global atlas of end-use energy demand density, all 1 km^2 cell-by-cell values between 1 GWh/km^2 and 90 GWh/km^2 are selected. Then, (2) this data is used to calculate the optimal number of clusters (ONC) according to the Elbow Method (EM). The EM uses Eq. (15) to calculate the sum of squared errors (SSE) of within-cluster distances between the cluster centres and their members. When the number of clusters k is plotted against SEE, the visually-determined location of the elbow of the curve indicates the appropriate number of clusters. Once the ONC is estimated, (3) the Hartigan-Wong algorithm (1979) for *K-means* is applied; this is defined by Eq. (15) [55] and Eq. (16) [56], and explained in [57]. The *K-means* clustering algorithm applies an iterative process aiming at minimizing the intra-cluster inertia criterion defined by Eq. (16).

$$minimise\ \left(W_k = \sum_{r=1}^{k} \frac{1}{n_r} D_r\right) \qquad \text{Eq. (15)}$$

$$D_r = \sum_{i=1}^{n_r-1} \sum_{j=1}^{n_r} \left\| d_i - d_j \right\|^2 \qquad \text{Eq. (16)}$$

where W_k is the average internal sum of squares, k is the number of clusters, n_r is the total members points in the cluster r, and D_r is the sum of distances between points d_i and d_j, that belong to a cluster.

3. Results and discussion

As a result of the energy density mapping methodology applied, a spatial mapping of global heat, cooling and water heating demand is presented for 165 countries, representing 99.96% of the global population. The seasonal and hourly data profile of end-use energy demands for each country are also presented. Further analysis of these large datasets is conducted with respect to two areas of interest (1) EDB for annual end-use energy demands (SH, WH and SC), and (2) seasonal and hourly end-use energy demand profiles for each EDB, for each country. These data can be used as inputs to modelling and are included as supplementary material alongside to this article.

3.1. Global heating and cooling demand

The global atlases of space heating demand and space cooling demand are illustrated in Fig. 2 and Fig. 3 respectively. It is found that 81% of the global SH demand is concentrated in Europe (44%), USA (18%), and China (19%). Overall, the highest energy demands for space heating are located in Northeast China, Northern and Central Europe, Northeast USA and the western part of Russia; countries with both extreme seasonal weather conditions and highly populated areas. Similarly, significant space cooling demand is also identified in India (8%), China (20%) and USA (30%).

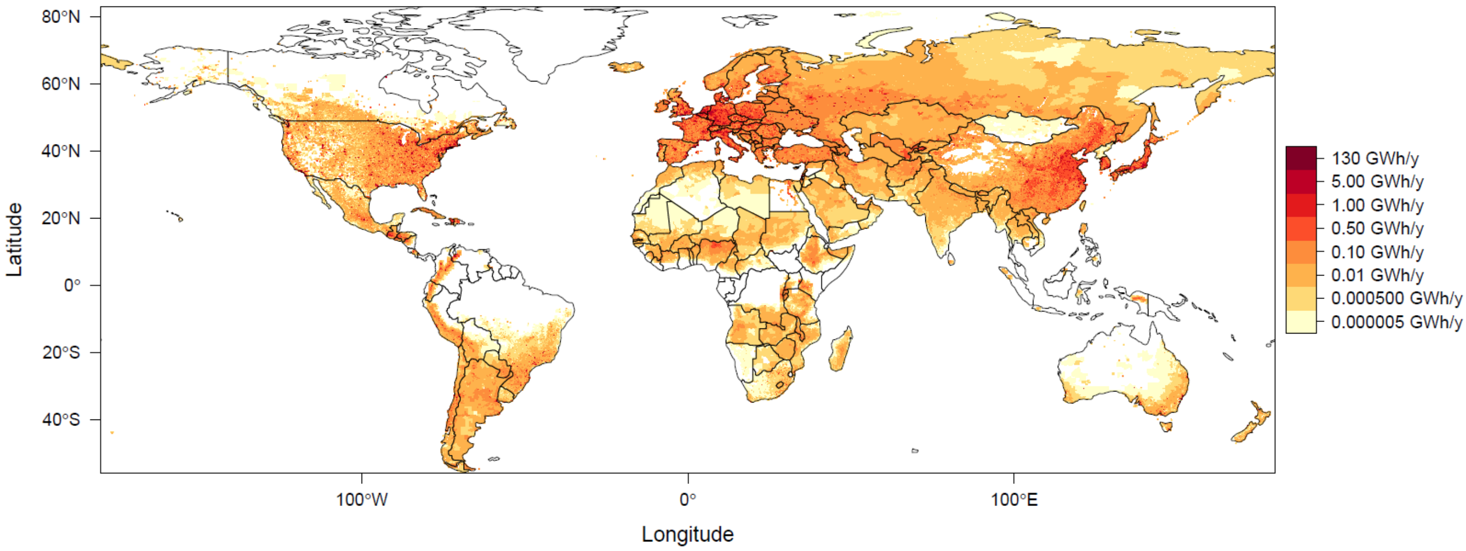


Figure 2: Global atlas of space heating demand in the residential sector. This atlas has been generated taking into account the global demand for spacing heating in each cell of 1 km^2 resolution. This atlas illustrates the annual spacing heating demand distribution globally. Similar atlases have been obtained for a seasonal hourly temporal resolution. A total of 96 time slices are considered for 165 countries covering 99.96% of global energy users.

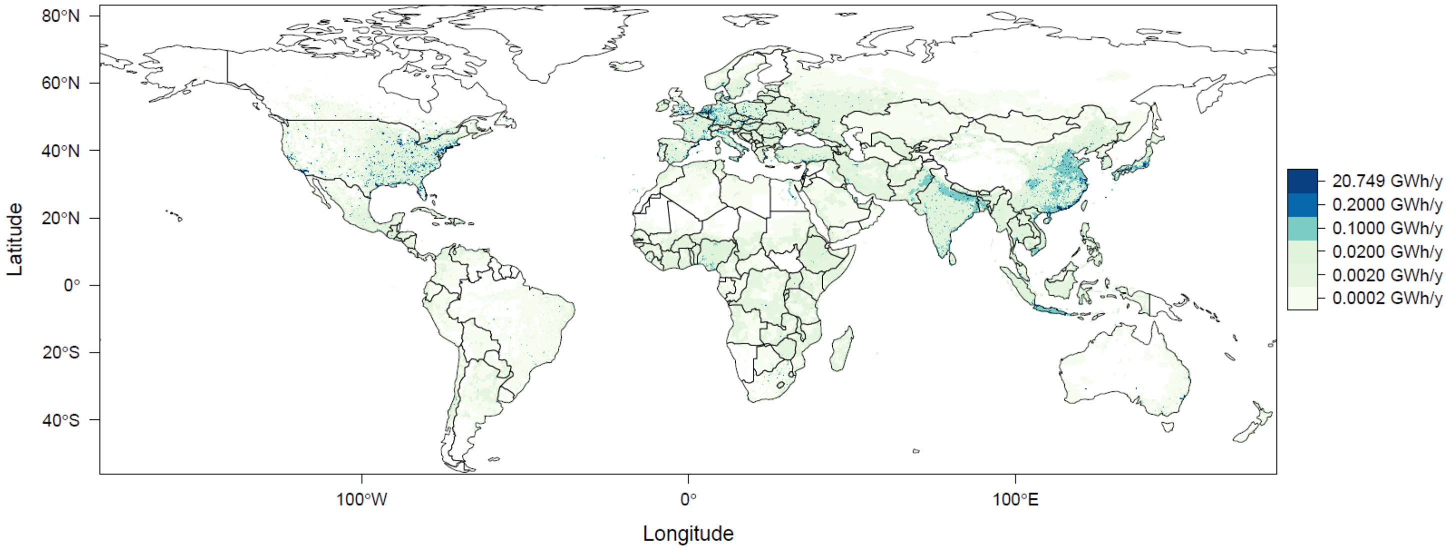


Figure 3: Global atlas of space cooling demand in the residential sector. This atlas has been generated taking into account the global demand for spacing cooling in each cell of 1 km^2 resolution. This atlas illustrates the annual spacing cooling demand distribution globally. Similar atlases have been obtained for a seasonal hourly temporal resolution. A total of 96 time slices are considered for 165 countries covering 99.96% of global energy users.

3.2. Regional heating and cooling demand

Fig. 4 shows the highest consuming twenty countries responsible for 77% of heat demand and 83% of cooling demand worldwide. It is clear that China and USA lead the demand of this end-use energy, as expected. High demand for space cooling in India is also apparent; mostly located in Northern India. Clearly local temperature variations along with population density are reflected in the demand for end-use energy results. Of particular note is the dominance of the highest five countries, responsible for approximately 50% of the global demand for these energy end-uses.

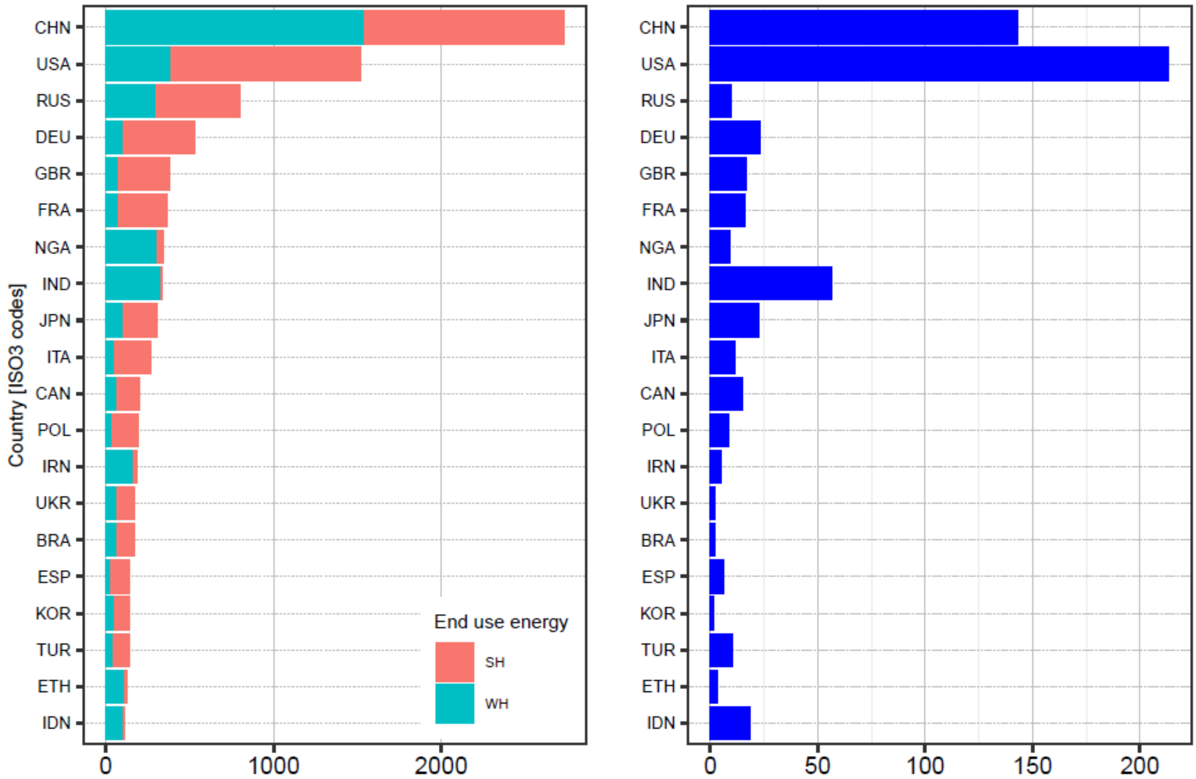


Figure 4: The twenty countries with the largest energy demand for heating (left), including SH and WH, and for space cooling (right) in the residential building sector. Units are in TWh/year. These countries are responsible for approximately 77% of heat demand, including space heating and water heating, and 83% of cooling demand worldwide.

3.3. Energy density banding results

Fig. 5 illustrates the evolution of the SSE values with increasing *k*. As *k* increases, SSE tends to zero, creating an elbow at the 6-cluster solution. This suggests that cluster solutions larger than 6 do not provide a substantial impact on the total SSE. Once the ONC is selected at 6, the *K-means* clustering algorithm is applied to obtain the clusters (centred around their respective centroids). Then, the EDB bounds are defined to be halfway between each consecutive centroid value, thus defining the limits of each band. Table 3 shows the lower bound and upper bound of each EDB. This is based on the literature for the extreme bands (to eliminate spurious k-means results in very high or very low energy density areas), and based on the clustering approach for bands within those extremes. Fig. 6 represents the resulting EDB within the annual heating energy demand for 165 countries representing 99.96% of global energy consumers. A similar method is also applied to cooling demand and presented in Appendix A.

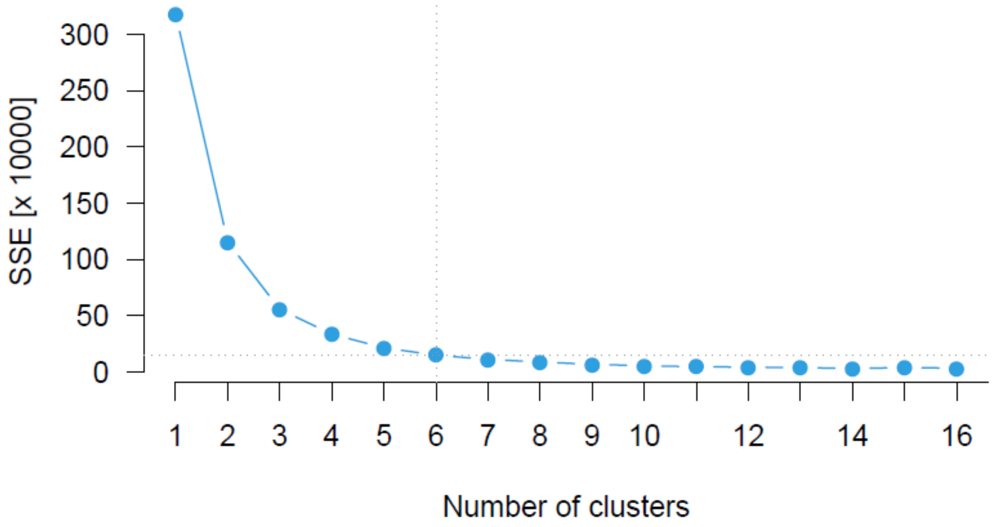


Figure 5: The Elbow Method to define the optimal number of clusters. This number of clusters is used as an input in the K-means algorithm. ONC = 6. The evolution of SSE values with increasing the number of clusters, *k* creates an elbow at the 6-cluster solution. This suggests that cluster solutions larger than 6 do not provide a substantial impact on the total SSE variation.

Table 3: Energy density bands widths based on clustering analysis results and previous research found in the literature. Bands are presented for heating end use demand. Separately, bands are also presented for cooling end use demand, see Appendix A. Heating end use refers to space heating and water heating while cooling end use refers to space cooling. SH plus WH energy density clustering was conducted at the global scale with a resolution of 1km^2.

EDB	Lower bound [MWh/km^2]	Upper bound [MWh/km^2]	Validation
1	0	1000	[6, 51]
2	1000	1790	Spatial clustering on SH+WH energy density
3	1790	5680	Spatial clustering on SH+WH energy density
4	5680	12080	Spatial clustering on SH+WH energy density
5	12080	21360	Spatial clustering on SH+WH energy density
6	21360	36930	Spatial clustering on SH+WH energy density
7	36930	63910	Spatial clustering on SH+WH energy density
8	63910	90000	Spatial clustering on SH+WH energy density
9	90000	Inf	[6, 51]

Fig. 6 provides the distribution of EDB for residential heat-related energy demand for the 165 countries considered in this study. From this figure, it is apparent that the larger part of aggregate demand occurs in the lower energy density bands. In essence, despite much higher energy density in cities etc, the sheer volume of less dense areas leads to greater aggregate demand there. The supplementary material contains the number of cells accounted in each EDB for each country.

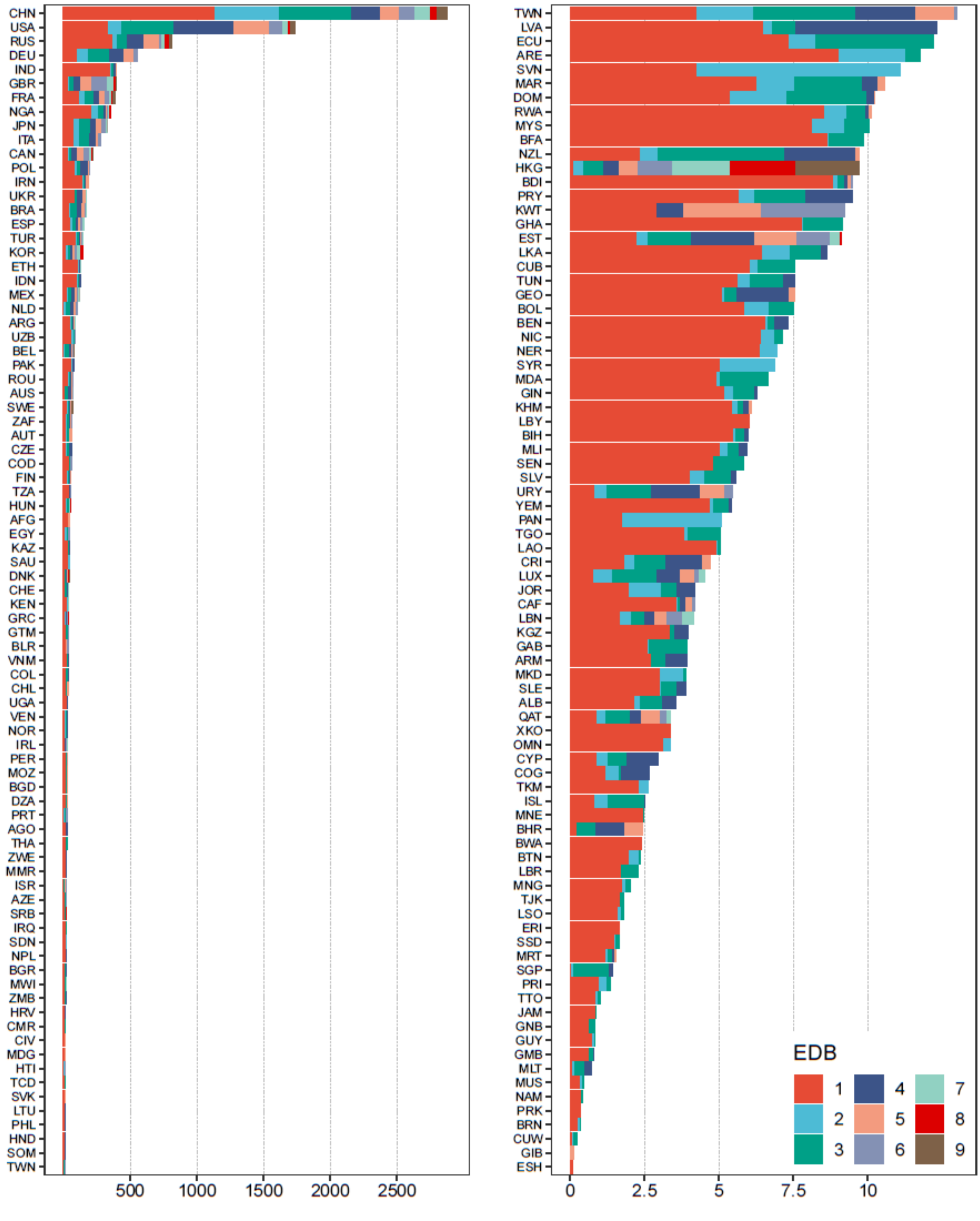


Figure 6: Energy density bands for global end-use energy demand per country [TWh/year]. The distribution of EDB for residential heat-related energy demand for the 165 countries is presented here. It is apparent that the larger part of aggregate demand occurs in the lower energy density bands. In essence, despite much higher energy density in densely populated areas, the sheer volume of less dense areas leads to greater aggregate demand there. In the supplementary material, the area accounted in each energy density band for each country is provided.

Fig. 7 shows the distribution of the EDB in the countries with the largest demands for heating (77%, left) and cooling (83%, right). In the case of China, the highest three EDBs represent 8.5% and 7.3% of heating and cooling demands respectively, and are associated with a total land area of 5708 km^2. On the other hand, the lower three EDBs account for 75% and 74.7% of heating and cooling demands respectively, associated with a total area of approximately 11 million km^2.

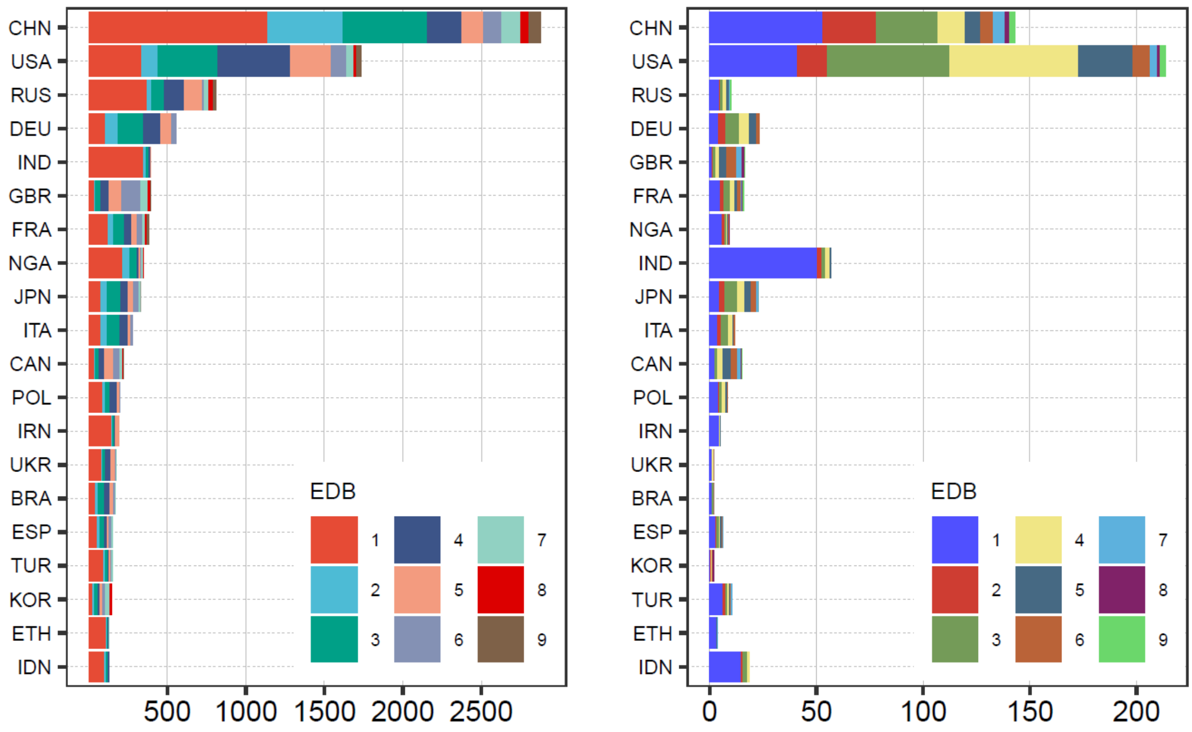


Figure 7: Energy density bands for the highest consuming twenty countries worldwide. The energy density bands are presented for heating (left) and cooling (right) for corresponding bands. Units are in TWh/year.

The case of USA is similar to China when comparing the highest three EDBs, but completely different for the lower three EDBs. The highest three EDBs represent 5.4% and 3.3% of heating and cooling demands respectively, associated with an area of 2183 km^2. However, bands 1, 2 and 3 represent 47% and 53% of heating and cooling demands respectively, associated with approximately 9.4 million km^2. This result reflects the higher urbanisation of the USA relative to China.

India is a unique case because only five EDBs are present due to the lack of occurrence of the extreme energy densities seen in China or the USA. In India, the highest two EDBs represent

4.1% and 4% of heating and cooling demands respectively, associated with 2394 km^2 land area. In contrast, 88.5% of heating demand and 88.6% of cooling demand are in band 1, covering approximately 4 million km^2 of land. The reader should note that the analysis presented represents a snapshot in time, and one would expect the spatial characteristics of demand to change as countries become wealthier [58] and more urbanised [59]. Also in particular for cooling demand, ownership of air conditioning equipment, which has a strong correlation with wealth, is an important driver of results [60].

Results are also presented with respect to the temporal variation of heating and cooling. The temporal variability is captured in seasonal profiles with an hourly resolution per EDB for each country. Energy demand profiles have been traditionally used by utility companies, covering large areas, and thus do not well represent the spatio-temporal phenomena of demand. Also, a utility company's data is not usually accessible to the public [21]. The use of high-resolution country-wide detailed spatio-temporal demand profiles developed in this paper can therefore contribute to the insights obtained regarding technology choice in the residential sector. Some possible applications of these data sets are (1) the realistic representation of the spatio-temporal energy demand variability in energy modelling, (2) the assessment of deployment of low carbon technologies in an area of interest, (3) the design of location-based electricity tariffs for distributed generation, and (4) the use of local renewable energy sources to meet the demand in the residential sector. Fig.8 illustrates the USA-wide example of 24-hr energy demand profile of SH, SC, WH in band 2. Fig. 9 additionally provides the spatio-temporal visualization of a 4-hr (13:00 – 17:00) SH demand for USA in each of the four seasons. The supplementary material supplied with this paper contains the data needed to distribute annual energy demand values in each EDB to arrive at the seasonal/diurnal profiles for each country.

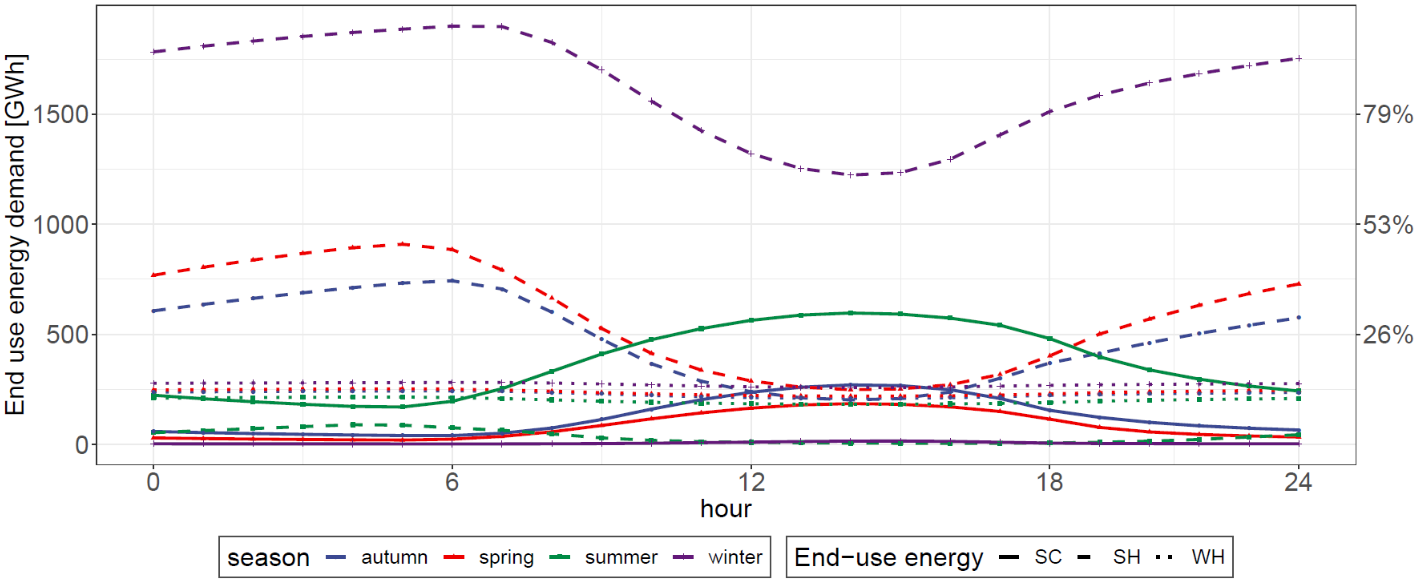


Figure 8: USA 24-hr profile of each service demand for each season for the second highest consuming energy density band (EDB = 2). The variation of SH demand in day peak hours can reach up to 40% less than night peak hours. The SC has a demand profile with a plateau between 11:00 and 17:00 in summer in EDB2.

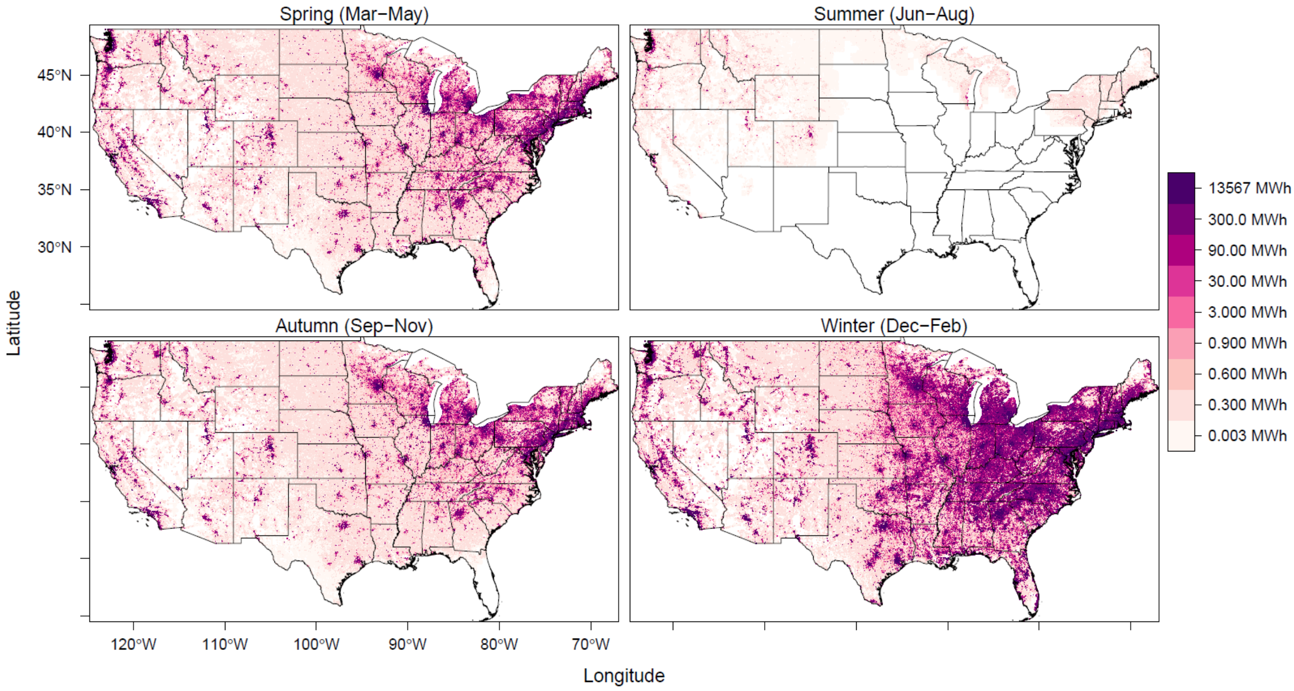


Figure 9: Four-hour total space heating demand for USA for selected time slices in each of the four seasons. Results are presented for USA SH demand [MWh], 1pm – 5pm. One cell is equal to 1 km^2.

3.4. Approach comparison

The approach presented here differs from earlier works that estimate the global spatio-temporal end-use energy demand (SH, SC, and WH) at a high resolution of 1 km^2. One approach found in the literature is that of Voulis, et al. [21]. Both Voulis et al. and this study share a number of key features, such as the estimation of energy demand profiles considering spatio-temporal characteristics, and the use of clustering techniques. However, there are important differences: Here, we estimate the spatio-temporal energy demand profiles for SH, SC and WH globally while Voulis, et al. [21] estimate only the spatio-temporal electricity demand for specific areas of a selected country. Although that approach presents novelty in terms of spatio-temporal considerations, it is also limited with respect to application at the global scale.

Another relevant study has been conducted by Rhodes, et al. [61]. In a similar way to our study, Rhodes, et al. use a *K-means* clustering algorithm to determine the pattern of demand profiles in the residential sector. Their study is different from the approach presented in this paper in a number of respects. For example, Rhodes, et al. conduct a spatio-temporal analysis for only 103 homes in a selected city of the USA. They analyse electricity demand data at the household level while we estimate the global demand of heating and cooling at the resolution of 1 km^2. What is interesting from Rhodes, et al. [61] is that they validate the applied clustering approach with the use of survey data of the same 103 dwellings. Although this approach gives valuable insight to the accuracy of estimations, is also presents data challenges for application at larger scales.

The scale of the datasets and the goal of the analysis presented in this study make comparison with other studies challenging. In the scientific literature, significant approach differences have been found, including (1) the estimation of demand density (e.g. USA and China [8], and Switzerland [9]); (2) the estimation of demand profiles for SH, SC, and WH [21]; (3) the spatio-temporal scale of the approach (e.g. national scale [15] and spatial considerations [18]), i.e. global 1 km^2 spatial resolution and seasonal daily-hourly variation.

3.5. Approach validation

The GIS-based methodology presented in this paper can be described as a bottom-up approach to estimate countrywide spatio-temporally resolved energy demand profiles for heating and cooling worldwide. At smaller spatial scales (i.e. mega-city, metropolitan areas, neighbourhoods, intra city areas), bottom-up approaches can result in detailed descriptions of spatio-temporal demand variations [62] and energy user compositions [63]. However, one of the limitations of these approaches is that it requires large location-specific datasets (i.e. building features, energy consumption description at the appliance level) that are not available on a consistent basis or tractable at larger scales [21]. To address this limitation in order to scale-up the methodologies from city level to national/region level, a combination of datasets from others fields has been presented here.

Therefore, datasets upon which to perform a validation are presently not widely available. Thus, the validation of the approach herein relies on (1) comparing the aggregated cell-by-cell energy demand with selected sub-national measurements of energy statistics data; and (2) comparing the aggregated cell-by-cell energy demand with the Annual Energy Balances from the IEA.

(1) The generated results presented in this study have been validated with the UK National Statistics of the Lower and Middle Super Output Areas (LLSOAs) with respect to gas consumption to meet heating demand in the residential sector. Two countries have been considered: England and Wales [64]. Although these energy statistics does not provide spatio-temporal detail, the annual country-level energy balance data is used for validation purposes, and thus enables the measurement of calculation accuracy undertaken in this study. First, energy consumption data was obtained from [64]. This data represents the total heating energy consumption in the residential sector in approximately 40,000 LLSOAs across England and Wales. A significant number of gas meters were considered in this study, as can be seen in Table 4. Second, data was aggregated by local authority for 335 authorities in total. Third, using the GADM database of Global Administrative Areas [65], all 1 km^2 cell-by-cell values of the annual heating demand estimated in this study were extracted from raster datasets, aggregated and matched with the same 335 local authorities. This step allows us to have comparable results with the data collected from LLSOAs. Finally, both datasets, this study and LLSOAs, are compared. As can be observed in Fig. 10, the approach presented here has a high accuracy of demand estimation in respect to LLSOA data. In Table 5, ab additional comparison is performed using absolute percentage error. We can see that the absolute percent error varies among boroughs, however, it is aligned with the error calculated for both countries overall (Table 4).

Table 4: Comparison of energy statistics of gas consumption in the UK's Lower Layer Super Output Areas (LLSOAs) with results presented in this study. England and Wales were analysed, accounting for 335 local authorities. Heating demand represents the energy required for space heating and water heating uses.

No. of LLSOAs	No. of gas meters	UK countries	Authori ties	LLSOA energy statistics [TWh/year]	This study estimation [TWh/year]	Percent error, abs [%]
39928	20,783,505	2	335	300	318	6

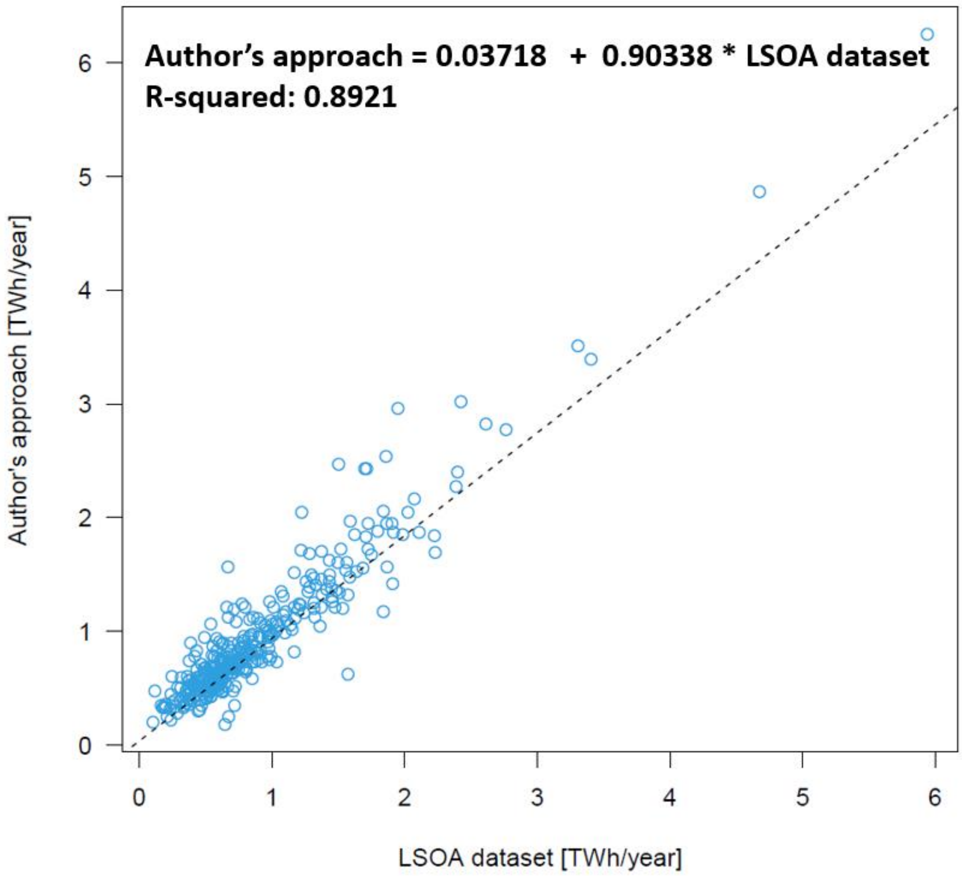


Figure 10: A comparison of selected LLSOAs data and results from this approach for England and Wales is illustrated. A linear regression shows that predictions in this paper are similar to LLSOAs measurements. At higher demand levels, the approach presented here shows a slight overestimation of demand, likely because the related areas are typically smaller and very dense, leading to higher potential errors where the spatial grid in this article does not match exactly with LLSOA boundaries.

Table 5: Comparison of energy statistics of gas consumption in boroughs of Greater London taking into account national statistics of the UK's Lower Layer Super Output Areas (LLSOAs) with results presented in this study. Ten local authorities are selected for comparison purposes. Here, heating demand represents the energy required for space heating and water heating uses.

Authority	**No. of LLSOAs**	**Area covered [km^2]**	**No. gas meters per km^2**	**LLSOA Energy statistics [TWh/year]**	**This study estimation [TWh/year]**	**Percent error, abs [%]**
Barnet	250	86.74	1447.91	2.39	2.28	4.61
Bexley	163	60.56	1463.24	1.38	1.32	4.37
Greenwich	158	47.35	1775.76	1.27	1.35	6.29
Hackney	160	19.06	4727.96	1.10	1.18	6.67
Hammersmith and Fulham	127	16.40	4546.59	1.00	1.06	5.75
Haringey	168	28.50	3349.16	1.43	1.50	4.25
Islington	137	14.86	5877.93	1.03	1.00	2.98
Lambeth	204	26.82	4416.41	1.56	1.54	1.03
Merton	144	37.61	1999.52	1.17	1.21	3.53
Waltham Forest	167	38.82	2350.46	1.33	1.41	6.29

(2) The obtained results have been compared with the Annual Energy Balances from the IEA. Although the IEA does not provide detailed spatio-temporal demand profiles, this is a self-validation approach that matches the annual country-level energy balance data of the residential sector, and thereby ensures the calculation steps undertaken are correct. All 1 km^2 cell-by-cell values were aggregated and matched the national energy balances presented by the IEA.

Future research will consider further validation measures, such as comparison with emerging spatially-resolved demand datasets that may be able to empirically validate the results further.

Conclusion

This article has formulated and applied a methodology for spatially and temporally-resolved estimation of global space heating, space cooling and hot water demand. It combines energy balances, spatially-resolved population density, and spatially and temporally-resolved temperature data to calculate these demands at 1km^2 resolution for 165 countries, representing 99.96% of global energy consumers. A k-means clustering algorithm has then been applied to this data to arrive at a set of representative energy density bands for each country, each of which is then further disaggregated into seasonal and hourly demand profiles. The result is the first self-consistent analysis of these three energy end-uses at a global scale, disaggregated into country and energy density categories, and provided in profile form to capture seasonal and diurnal demand variations.

End-use energy demand for heat and cooling varies from country to country depending on weather conditions, population density, and level of uptake of energy-consuming devices to provide these services. It is apparent from the results presented here that global space heating demand is concentrated in three regions of the world; Northeast China, Northern and Central Europe, Northeast USA. Conversely, global cooling demand is heavily concentrated in the USA, China and India, reflecting not only population and temperature characteristics, but also the correlation of gross domestic product with the ownership of air conditioning equipment.

In terms of energy density, the results show that a relatively small portion of demand (approx. ~5%) occurs at very high energy density locations (i.e. above 36.9 GWh/km^2). This is in contrast to what is usually upwards of 50% of demand that occurs in low energy density locations (i.e. below 1.79 GWh/km^2). Each country differs substantially in this regard, highlighting the importance of this consistent dataset that captures the core differences between each country, and is applicable for use in the global integrated assessment modelling that is crucial for climate change mitigation assessment.

The dataset produced is published as supplementary material alongside this article. It describes the amount of energy in each energy density band for each country, including disaggregation into seasonal and hourly demand profiles for each band. The number of cells in each energy density band is also provided. The dataset is provided for heat and cooling demand using (a) the approach of clustering based on space heating and water heating demand combined (with space cooling also reported for each band), and (b) the approach of clustering based on space cooling demand only. It is expected that this data will be a valuable resource for the energy systems modelling and integrated assessment modelling communities, in addition to providing a sound basis for techno-economic or decarbonisation assessment of thermal demand for any of the 165 countries considered.

Finally, the authors would like to draw attention to the limitations of the analysis undertaken, in order to inform focus areas for future research. The important limitations are that:

a) It is assumed that heat and cooling energy demand density is proportional to population density, with basic adjustment to reflect lower per capita demand in energy dense areas. Researchers must be aware that this is not always the case as in some regions of the world, energy demand density is not proportional to population density due to economic and other factors. Further research should consider income distribution spatially in order to have a better understanding on the economic capacity of areas. Here, there is an opportunity for future research as the spatial-temporal distribution of energy service demand, when combined with socio-economic factors, can lead to a better understanding of the potential for clean technologies in diverse contexts.
b) This analysis represents a snapshot in time, and the spatial and appliance ownership characteristics of demand would be expected to change as countries become wealthier and more urbanised. Further research can apply the methodology presented here to data available for different years. This would provide further insights of how the EDB along with its associated demand evolve in time.
c) Only final energy consumption has been spatially and temporally distributed in the study, implying that the efficiency of conversion of that final energy to the heat or cooling energy service demand has not been considered. As future research could focus on the technology choices to meet the demands in each EDB, it is recommended that heating and cooling technology efficiency should be considered in the analysis. Additionally, at the time of writing, data for spatially resolved income analysis is not yet available. Including the income spatial distribution would allow researchers to have a better overview of what technologies are economically feasible in different locations. This limitation has in particular been identified as a key area for future research.

Acknowledgements

Mr Diego Moya has been funded by the Ecuadorian Secretariat for Higher Education, Science, Technology and Innovation (SENESCYT), Award No. CZ03-35-2017, The Technical University of Ambato (UTA), Award No. 1895-CU-P-2017 (Resolución HCU), and supported by The Science and Solutions for a Changing Planet Doctoral Training Partnership, Grantham Institute, at Imperial College London. The Institute for Applied Sustainability Research supports international research on global sustainability applied to the Global South. Dr Hawkes and Dr Sachs were supported by the NERC Newton project NE/N018656/1, Sustainable Gas Pathways for Brazil. Andrew Northern and David Sandoval, Imperial College London, Orlando Sabogal, Universidad Tecnológica de Pereira, and Raúl Fernádez, The University of Edinburgh, are acknowledged for their valuable comments during the development of this manuscript. We acknowledge the important comments and suggestions made by four anonymous reviewers to improve the quality, clarity and strictness of this article.

Author Contributions

J.S., D.M wrote the article. A.H., S.G., J.S. conceptualised the energy density maps approach. D.M., J.S., A.H. conceptualised the energy density bands approach. S.G., J.S. gathered data for energy density maps. J.S. wrote the data processing scripts and processed the data for developing the energy density maps. D.M. wrote the data processing scripts and processed the energy density maps for developing the energy density bands and the spatio-temporal end-use energy demand profiles. D.M. took responsibility for the integrity of the data, the accuracy of the data analysis, and methodology validation. All authors contributed to the article development, critically read and approved the final version of the manuscript.

Appendix A: EBD for cooling demand

A similar approach described in Section 3.2 is applied for global SC demand. Fig. 10 illustrates the EM to determine the ONC for SC data (the ONC is determined to be 5). Furthermore, the lowest and highest EDB are set based on the work presented in [7, 11, 13, 17]. Table 4 defines the EDB for SC demand based on the clustering analysis. Here, bands are presented for the cooling end use energy demand only. SC energy density clustering was also conducted at the global scale with a 1 km^2 resolution. For researchers requiring SC demand estimations, Fig. 11 shows the distribution of the EDB within 165 countries.

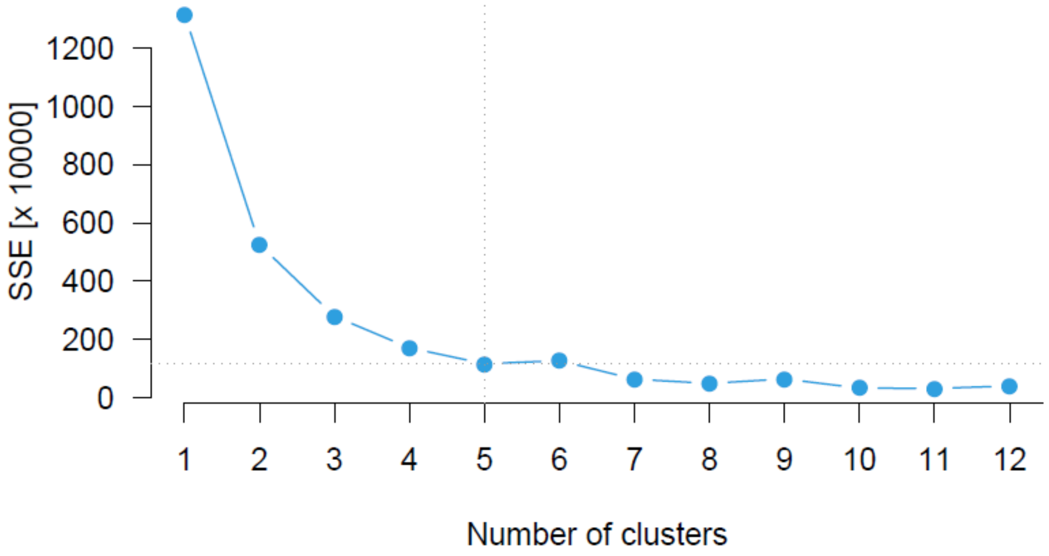


Figure 11: The Elbow Method to define the optimal number of clusters for SC demand. This number of clusters is used as an input in the K-means algorithm. ONC = 5. The evolution of SSE values with increasing the number of clusters, *k* creates an elbow at the 5-cluster solution. This suggests that cluster solutions larger than 5 do not provide a substantial impact on the total SSE variation.

Table 6: SC Energy density bands widths based on clustering analysis results and previous research found in the literature. Bands are presented for cooling end use demand. SC energy density clustering was conducted at the global scale with a resolution of $1km^2$.

Cooling energy density [MWh/km^2]			**Justification**
EDB	Lower bound	Upper bound	
1	0	10	[7, 11, 13, 17]
2	10	36	Spatial clustering on SC energy density
3	36	364	Spatial clustering on SC energy density
4	364	1087	Spatial clustering on SC energy density
5	1087	2420	Spatial clustering on SC energy density
6	2420	5228	Spatial clustering on SC energy density
7	5228	16600	Spatial clustering on SC energy density
8	16600	Inf	[7, 11, 13, 17]

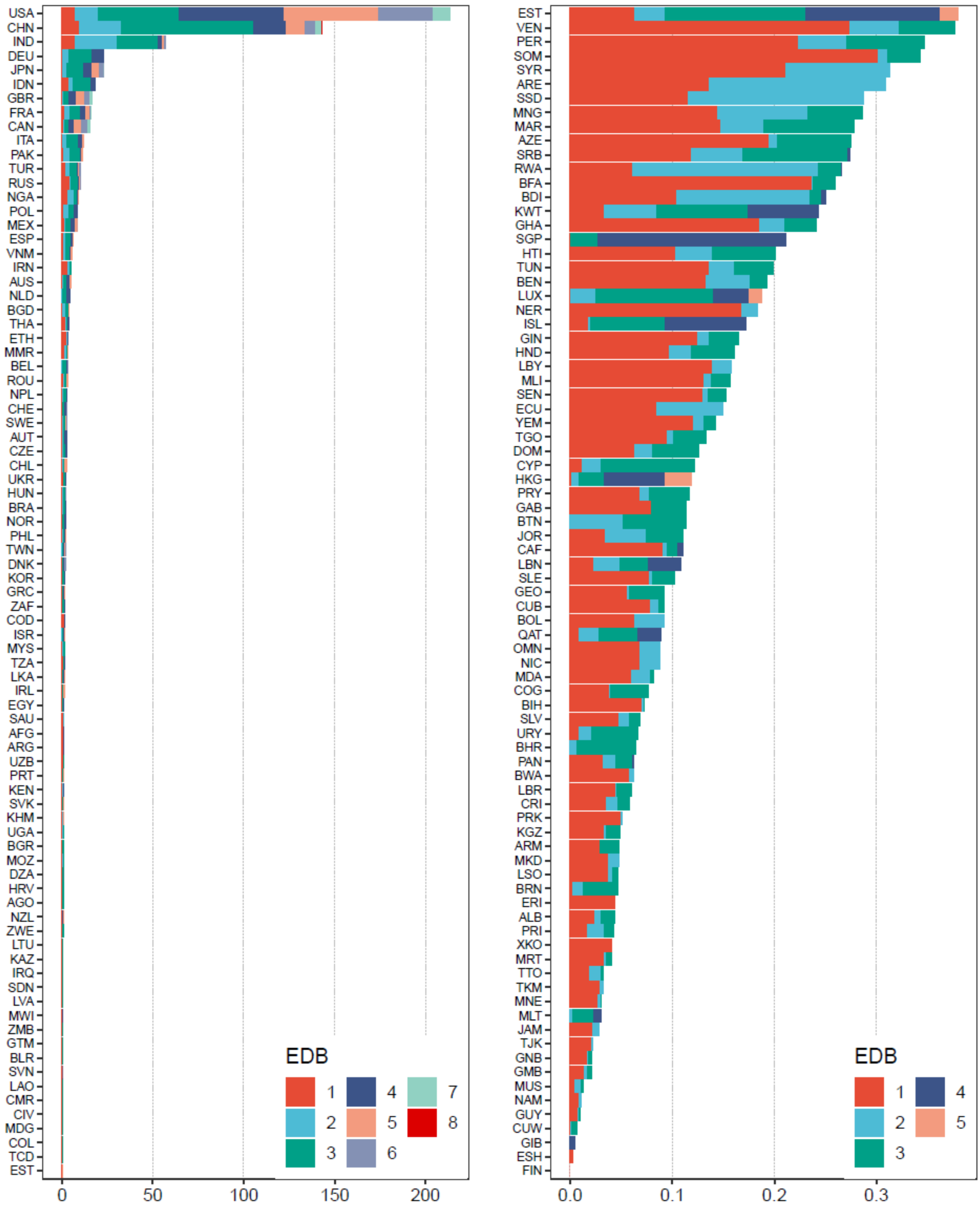


Figure 12: Energy density bands for global SC energy demand per country [TWh/year]. The distribution of EDB for residential cooling-related energy demand for the 165 countries is presented here. It is apparent that the larger part of aggregate demand occurs in the lower energy density bands. In essence, despite much higher energy density in densely populated areas, the sheer volume of less dense areas leads to greater aggregate demand there. In the supplementary material, the area included in each energy density band for each country is provided.

Appendix B: Supplementary data

The following are the supplementary data related to this article. Spreadsheet of the factors to allocate annual end-use energy demand into seasonal/diurnal end-use energy demand profiles for each EDB for each of the 165 countries. This is presented for each of the two clustering methods, heating and cooling separately. Units are in MWh/km^2 for end-use energy density demand and MWh for end-use energy demand.

References

[1] IEA, "Technology Roadmap Energy-efficient Buildings: Heating and Cooling Equipment," International Energy Agency2011.

[2] S. de la Rue du Can, A. Khandekar, N. Abhyankar, A. Phadke, N. Z. Khanna, and D. Fridley, "Modeling India’s energy future using a bottom-up approach," *Applied Energy,* vol. 238, pp. 1108-1125, 2019/03/15/ 2019.

[3] J.-L. Fan, J.-W. Hu, and X. Zhang, "Impacts of climate change on electricity demand in China: An empirical estimation based on panel data," *Energy,* vol. 170, pp. 880-888, 2019/03/01/ 2019.

[4] T. B. Johansson, A. P. Patwardhan, N. Nakićenović, and L. Gomez-Echeverri, *Global energy assessment: toward a sustainable future*: Cambridge University Press, 2012.

[5] Committee on Climate Change (CCC). (2018, 18 Nov). *Analysis of Alternative UK Heat Decarbonisation Pathways*. Available: https://www.theccc.org.uk/wp-content/uploads/2018/06/Imperial-College-2018-Analysis-of-Alternative-UK-Heat-Decarbonisation-Pathways.pdf

[6] F. Jalil-Vega and A. D. Hawkes, "The effect of spatial resolution on outcomes from energy systems modelling of heat decarbonisation," *Energy,* vol. 155, pp. 339-350, 2018/07/15/ 2018.

[7] S. Werner, "District Heating and Cooling," in *Encyclopedia of Energy*, C. J. Cleveland, Ed., ed New York: Elsevier, 2004, pp. 841-848.

[8] Y. Zhou, J. Eom, and L. Clarke, "The effect of global climate change, population distribution, and climate mitigation on building energy use in the U.S. and China," *Climatic Change,* vol. 119, pp. 979-992, August 01 2013.

[9] M. Berger and J. Worlitschek, "A novel approach for estimating residential space heating demand," *Energy,* vol. 159, pp. 294-301, 2018/09/15/ 2018.

[10] V. Chaturvedi, J. Eom, L. E. Clarke, and P. R. Shukla, "Long term building energy demand for India: Disaggregating end use energy services in an integrated assessment modeling framework," *Energy Policy,* vol. 64, pp. 226-242, 2014/01/01/ 2014.

[11] D. Connolly, "Heat Roadmap Europe: Quantitative comparison between the electricity, heating, and cooling sectors for different European countries," *Energy,* vol. 139, pp. 580-593, 2017/11/15/ 2017.

[12] S. Nielsen and B. Möller, "GIS based analysis of future district heating potential in Denmark," *Energy,* vol. 57, pp. 458-468, 2013/08/01/ 2013.

[13] M. Olonscheck, A. Holsten, and J. P. Kropp, "Heating and cooling energy demand and related emissions of the German residential building stock under climate change," *Energy Policy,* vol. 39, pp. 4795-4806, 2011/09/01/ 2011.

[14] D. Ürge-Vorsatz, L. F. Cabeza, S. Serrano, C. Barreneche, and K. Petrichenko, "Heating and cooling energy trends and drivers in buildings," *Renewable and Sustainable Energy Reviews,* vol. 41, pp. 85-98, 2015.

[15] M. Isaac and D. P. van Vuuren, "Modeling global residential sector energy demand for heating and air conditioning in the context of climate change," *Energy Policy,* vol. 37, pp. 507-521, 2009/02/01/ 2009.

[16] H. C. Gils, J. Cofala, F. Wagner, and W. Schöpp, "GIS-based assessment of the district heating potential in the USA," *Energy,* vol. 58, pp. 318-329, 2013.

[17] S. Werner, "European space cooling demands," *Energy,* vol. 110, pp. 148-156, 2016/09/01/ 2016.

[18] M. Jakubcionis and J. Carlsson, "Estimation of European Union residential sector space cooling potential," *Energy Policy,* vol. 101, pp. 225-235, 2017/02/01/ 2017.

[19] K. Schiel, O. Baume, G. Caruso, and U. Leopold, "GIS-based modelling of shallow geothermal energy potential for CO2 emission mitigation in urban areas," *Renewable Energy,* vol. 86, pp. 1023-1036, 2016/02/01/ 2016.

[20] F. Bühler, S. Petrović, F. M. Holm, K. Karlsson, and B. Elmegaard, "Spatiotemporal and economic analysis of industrial excess heat as a resource for district heating," *Energy,* vol. 151, pp. 715-728, 2018/05/15/ 2018.
[21] N. Voulis, M. Warnier, and F. M. T. Brazier, "Understanding spatio-temporal electricity demand at different urban scales: A data-driven approach," *Applied Energy,* vol. 230, pp. 1157-1171, 2018/11/15/ 2018.
[22] A. Rahman and A. D. Smith, "Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms," *Applied Energy,* vol. 228, pp. 108-121, 2018/10/15/ 2018.
[23] C. Su, H. Madani, and B. Palm, "Building heating solutions in China: A spatial techno-economic and environmental analysis," *Energy Conversion and Management,* vol. 179, pp. 201-218, 2019/01/01/ 2019.
[24] C. Geiß, H. Taubenböck, M. Wurm, T. Esch, and M. Nast, "Remote sensing-based characterization of settlement structures for assessing local potential of district heat," *Remote Sensing,* vol. 3, pp. 1447-1471, 2011.
[25] J. Unternährer, S. Moret, S. Joost, and F. Maréchal, "Spatial clustering for district heating integration in urban energy systems: Application to geothermal energy," *Applied energy,* vol. 190, pp. 749-763, 2017.
[26] A. Schlueter, P. Geyer, and S. Cisar, "Analysis of georeferenced building data for the identification and evaluation of thermal microgrids," *Proceedings of the IEEE,* vol. 104, pp. 713-725, 2016.
[27] C. Delmastro, G. Mutani, and L. Schranz, "The evaluation of buildings energy consumption and the optimization of district heating networks: a GIS-based model," *International Journal of Energy and Environmental Engineering,* vol. 7, pp. 343-351, 2016.
[28] I. Baldvinsson and T. Nakata, "Cost Assessment of a District Heating System in Northern Japan Using a Geographic Information–Based Mixed Integer Linear Programming Model," *Journal of Energy Engineering,* vol. 143, p. F4016006, 2017.
[29] J. Weyant, "Integrated assessment of climate change: state of the literature," *Journal of Benefit-Cost Analysis,* vol. 5, pp. 377-409, 2014.
[30] N. Strachan, N. Balta-Ozkan, D. Joffe, K. McGeevor, and N. Hughes, "Soft-linking energy systems and GIS models to investigate spatial hydrogen infrastructure development in a low-carbon UK energy system," *International Journal of Hydrogen Energy,* vol. 34, pp. 642-657, 2009.
[31] M. Zeyringer, H. Daly, B. Fais, E. Sharp, and N. Strachan, "Spatially and temporally explicit energy system modelling to support the transition to a low carbon energy infrastructure–case study for wind energy in the UK," 2015.
[32] S. N. Petrovic and K. B. Karlsson, "Danish heat atlas as a support tool for energy system models," *Energy Conversion and Management,* vol. 87, pp. 1063-1076, 2014.
[33] B. Möller, "A heat atlas for demand and supply management in Denmark," *Management of Environmental Quality: An International Journal,* vol. 19, pp. 467-479, 2008.
[34] R. Loulou, G. Goldstein, and K. Noble, "Documentation for the MARKAL Family of Models," *Energy Technology Systems Analysis Programme,* pp. 65-73, 2004.
[35] R. Loulou, U. Remne, A. Kanudia, A. Lehtila, and G. Goldstein, "Documentation for the TIMES Model, PART I: Energy Technology Systems Analysis Programme," ed: International Energy Agency Paris, 2005.
[36] R. Loulou and M. Labriet, "ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure," *Computational Management Science,* vol. 5, pp. 7-40, 2008.
[37] JGCRI, "The Global Change Assesment Model (GCAM)," ed: \url{http://jgcri.github.io/gcam-doc/index.html}, 2016.

[38] J. T. Wilkerson, D. Cullenward, D. Davidian, and J. P. Weyant, "End use technology choice in the National Energy Modeling System (NEMS): An analysis of the residential and commercial building sectors," *Energy Economics,* vol. 40, pp. 773-784, 2013.
[39] EIA, "Residential Demand Module of the National Energy Modeling System: Model Documentation," U.S. Energy Information Administration2014.
[40] E3MLab, "Primes Model - Detailed model description," E3MLab/ICCS at National Technical University of Athens2014.
[41] IEA. (2016). *Summary Energy Balances: World Energy Balances, 1960-2014*.
[42] R. Gelaro, W. McCarty, A. Molod, M. Suarez, L. Takacs, and R. Todling, "The NASA Modern Era Reanalysis for Research and Applications, Version-2 (MERRA-2)," in *AGU Fall Meeting Abstracts*, 2014.
[43] CIESIN, "Gridded Population of the World, Version 4 (GPWv4): Population Count," ed: NASA Socioeconomic Data and Applications Center (SEDAC) Palisades, NY; and the Center for International Earth Science Information Network —Columbia University, 2016.
[44] CIESIN, "Gridded Population of the World, Version 4 (GPWv4): Population Density," ed: NASA Socioeconomic Data and Applications Center (SEDAC) Palisades, NY; and the Center for International Earth Science Information Network —Columbia University, 2016.
[45] D. Uerge-Vorsatz, T. B. Johansson, A. P. Patwardhan, N. Nakićenović, and L. Gomez-Echeverri, *Chapter 10-energy end-use: building. Global energy assessment: toward a sustainable future*: Cambridge University Press, 2012.
[46] OECD/IEA, "Technology Perspectives, Harnessing Electricity's Potential," International Energy Agency2013.
[47] Eurostat. (10 September). *Metadata on Energy Statistics – Heating Degree Days*. Available: https://ec.europa.eu/eurostat/cache/metadata/en/nrg_chdd_esms.htm
[48] D. S. Parker and F. S. E. Center, "Factors Influencing Water Heating Energy Use and Peak Demand in a Large Scale Residential Monitoring Study," in *Teaming for Efficiency: Proceedings, 2002 ACEEE Summer Study on Energy Efficiency in Buildings: Aug. 18-23, 2002, Asilomar Conference Center, Pacific Grove, California*, 2002, p. 157.
[49] H. C. Gils, "A GIS-based assessment of the district heating potential in Europe," *Deutsches Zentrum für Luft-und Raumfahrt (DLR): Graz, Austria,* 2012.
[50] U. Persson and S. Werner, "Heat distribution and the future competitiveness of district heating," *Applied Energy,* vol. 88, pp. 568-576, 2011/03/01/ 2011.
[51] F. A. Jalil Vega, "Development of a systems approach for studying decarbonisation pathways of heat demand in the UK," 2018.
[52] F. Jalil-Vega and A. Hawkes, "Spatially resolved model for studying decarbonisation pathways for heat supply and infrastructure trade-offs," *Applied Energy,* vol. 210, pp. 1051-1072, 2018.
[53] F. McLoughlin, A. Duffy, and M. Conlon, "A clustering approach to domestic electricity load profile characterisation using smart metering data," *Applied Energy,* vol. 141, pp. 190-199, 2015/03/01/ 2015.
[54] P. Gianniou, X. Liu, A. Heller, P. S. Nielsen, and C. Rode, "Clustering-based analysis for residential district heating data," *Energy Conversion and Management,* vol. 165, pp. 840-850, 2018/06/01/ 2018.
[55] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* vol. 63, pp. 411-423, 2001.
[56] Y. Zhang, J. Mańdziuk, C. H. Quek, and B. W. Goh, "Curvature-based method for determining the number of clusters," *Information Sciences,* vol. 415-416, pp. 414-428, 2017/11/01/ 2017.
[57] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics),* vol. 28, pp. 100-108, 1979.

[58] B. C. O'Neill, E. Kriegler, K. L. Ebi, E. Kemp-Benedict, K. Riahi, D. S. Rothman, "The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century," *Global Environmental Change,* vol. 42, pp. 169-180, 2017/01/01/ 2017.
[59] S. Kc and W. Lutz, "The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100," *Global Environmental Change,* vol. 42, pp. 181-192, 2017/01/01/ 2017.
[60] R. V. Jones, A. Fuertes, and K. J. Lomas, "The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings," *Renewable and Sustainable Energy Reviews,* vol. 43, pp. 901-917, 2015/03/01/ 2015.
[61] J. D. Rhodes, W. J. Cole, C. R. Upshaw, T. F. Edgar, and M. E. Webber, "Clustering analysis of residential electricity demand profiles," *Applied Energy,* vol. 135, pp. 461-471, 2014/12/15/ 2014.
[62] A. Mastrucci, O. Baume, F. Stazi, and U. Leopold, "Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam," *Energy and Buildings,* vol. 75, pp. 358-367, 2014/06/01/ 2014.
[63] J. A. Fonseca and A. Schlueter, "Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts," *Applied Energy,* vol. 142, pp. 247-265, 2015/03/15/ 2015.
[64] Department for Business EIS. (March 29, 2019). *Lower and Middle Super Output Areas gas consumption 2010*. Available: https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-gas-consumption
[65] GADM. (2018, 13 Sep). *GADM database of Global Administrative Areas, version 3.6*. Available: https://gadm.org
[66] T. Räsänen, D. Voukantsis, H. Niska, K. Karatzas, and M. Kolehmainen, "Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data," *Applied Energy,* vol. 87, pp. 3538-3545, 2010/11/01/ 2010.