Review
The Use of Real Energy Consumption Data in Characterising Residential Energy Demand with an Inventory of UK Datasets

Lesley Thomson * and David Jenkins

Institute of Sustainable Built Environment, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh EH14 4AS, UK
* Correspondence: lt2013@hw.ac.uk

Abstract: The availability of empirical energy data from Advanced Metering Infrastructure (AMI)—which includes household smart meters—has enabled residential energy demand to be characterised in different forms. This paper first presents a literature review of applications of measured electricity, gas, and heat consumption data at a range of temporal resolutions, which have been used to characterise and develop an understanding of residential energy demand. User groups, sectors, and policy areas that can benefit from the research are identified. Multiple residential energy demand datasets have been collected in the UK that enable this characterisation. This paper has identified twenty-three UK datasets that are accessible for use by researchers, either through open access or defined processes, and presents them in an inventory containing details about the energy data type, temporal and spatial resolution, and presence of contextual physical and socio-demographic information. Thirteen applications of data relating to characterising residential energy demand have been outlined in the literature review, and the suitability of each of the twenty-three datasets was mapped to the thirteen applications. It is found that many datasets contain complementary contextual data that broaden their usefulness and that multiple datasets are suitable for several applications beyond their original project objectives, adding value to the original data collection.

Keywords: residential energy demand; real energy consumption data; energy demand dataset application; energy policy

1. Introduction

In accordance with agreements such as the 2015 Paris Agreement and 2021 Glasgow Climate Pact, countries, including the United Kingdom (UK), have targets to reduce greenhouse gas (GHG) emissions [1]. A portion of these reductions are expected to be attributed to the replacement of fossil fuels with renewable energy sources, but a simultaneous component is a requirement to reduce overall energy demand.

Despite upgrades to the housing stock and improvements in the efficiency of heating systems and appliances over the last two decades, the residential sector accounted for 33% and 21% of the UK’s 2019 annual final energy consumption [2] and GHG emissions [3], respectively. The built environment has been identified as a cost-effective sector to make energy demand and CO₂ savings [4–6], and upgrades to the housing stock are particularly important in the backdrop of market volatility to limit fuel poverty and improve occupant comfort.

The Energy Performance Certificate (EPC), initiated through the introduction of the European Union’s Energy Performance of Buildings Directive (EPBD) in 2002 [7], is one of the main vehicles in Europe for rating energy efficiency and gathering data on the residential building stock. It captures some aspects of modelled residential energy demand (in the UK, for example, lighting is the only type of electrical appliance considered) and outputs performance indicators. The indicators are typically annual energy demand per unit floor area (kWh/m²) and annual CO₂ emissions (kgCO₂/m²). An EPC is commonly produced.
on the basis of a steady-state calculation, with simplifications of thermodynamics [8] and standard assumptions about use and occupancy [9] to give a rating of the ‘asset’ rather than of its ‘operation’ [10]. While this approach allows for a level of energy assessment standardisation and inter-building comparison, discrepancies have been identified between calculated and actual energy demand [11–14], referred to as the performance gap. Older, less efficient buildings, in particular, have been shown to have their energy demand overestimated [15,16]. The implication of this is that if real residential energy demand is lower than that modelled, the potential energy savings across the sector through retrofit and renovation are overestimated, and energy efficiency measures are less cost-effective than predicted. This has been partly addressed by some European Union (EU) Member States, which have implemented alternative frameworks within which measurements of real energy consumption can form the basis of the EPC performance rating [17,18].

Although not generally used to generate an EPC energy rating—with the exception of lighting and auxiliary devices such as those in ventilation systems (note that what is included in the EPC calculation is country-dependent)—the energy used for household electrical appliances is an important component of residential energy demand and is an active research area. With an increase in non-dispatchable renewable energy sources, gaining an understanding of the potential for demand flexibility—which has several sub-topics such as load shifting and load shedding—with regard to appliance use is of interest to researchers.

Research into residential energy demand—whether it be for heating or electrical appliances—has been facilitated by the rollout of Advanced Metering Infrastructure (AMI) in households, which includes ‘smart meters’. By 2024, there are expected to be ~225 million (77% of consumers) electricity and ~51 million (44% of consumers) gas smart meters in the EU [19]. Alongside providing a record of real energy consumption, an advantage of smart meter data is the high temporal resolution of the recordings (30 min in the UK). From a researcher’s perspective, the volume and resolution of data from AMI (of which household smart meters are one example) open up opportunities for the large-scale study of transient energy demand characteristics and building energy performance [20–22]. There are barriers to accessing household smart meter data, data privacy regulations being one. However, empirical energy demand datasets have been collected for academic research in the UK and are now in open access. Nevertheless, there is a balance between aligning energy data with contextual metadata (and therefore providing crucial context for purposes of analyses) and data privacy (where such contextual data can cause barriers to anonymisation).

To make confident predictions of energy savings that can be achieved through upgrading the energy efficiency of the building stock and understanding the impact of introducing an increasing climate-based renewable energy supply, reliable measurements of current energy demand, load patterns, and their driving factors are needed. Therefore, there has been an effort to research the ways in which real energy consumption data can be used to characterise residential energy demand.

The aims of this paper are to (i) review the applications of real energy consumption data in characterising residential energy demand in practice and in research, (ii) identify user groups, sectors, or policy areas that benefit from the outputs of such characterisation, (iii) present a detailed inventory of UK residential energy demand datasets available to researchers, and (iv) categorise the identified datasets by their suitability to different data applications. Assessment of the useability and quality of each dataset is beyond the scope of this review. It is intended that this paper brings together an introductory overview of groupings of a wide variety of applications of real energy consumption data that are commonly presented in the literature for interested parties. There are future applications of data that could be formed by obtaining feedback from end users of energy demand data—this is intended to be a focus of future research. The datasets considered here are from the UK, but the applications of the datasets are relevant internationally.
2. Applications of Real Energy Consumption Data

This section presents a selection of applications of real energy consumption data in practice from a range of different countries. This includes examples of heating (e.g., gas) and non-heating (e.g., appliance electricity demand) energy demand data, but also recorded contextual data that are used to provide further meaning to empirical energy demand data. The review is categorised in order of increasing temporal resolution of the data being collected, where low resolution is defined as annual or monthly, medium resolution is defined as data on a daily scale, and high resolution is hourly or better. Table 1 summarises the applications of real energy consumption data and the criteria that are considered to make a dataset suitable for the application. Each application is described in turn throughout the section.

Table 1. Applications of real energy consumption data.

<table>
<thead>
<tr>
<th>Data Resolution</th>
<th>Application</th>
<th>Dataset Suitability Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Develop, monitor, and evaluate energy policies</td>
<td>Annual energy consumption data that could be used for policy development and/or to evaluate the effect of implemented policies (Section 2.1)</td>
</tr>
<tr>
<td></td>
<td>Benchmarking annual energy consumption</td>
<td>Annual energy consumption data that could form a database used to benchmark the energy performance of a residential building against similar types (Section 2.2)</td>
</tr>
<tr>
<td></td>
<td>Variability in overall consumption (including enabling analysis of drivers of consumption)</td>
<td>Annual energy consumption data coupled with contextual data that can be used for the applications described in Section 2.3</td>
</tr>
<tr>
<td>Low-to-high</td>
<td>Validating assumptions in EPC calculations</td>
<td>Datasets that record heating patterns over at least 1 full year or include internal temperature measurements that can be used to evaluate if assumptions made in EPC calculations (about variables such as length of heating season, number of heating hours and internal temperature) are valid (Section 2.4)</td>
</tr>
<tr>
<td>Medium</td>
<td>Prediction of building thermal properties</td>
<td>Data that have been collected in field trials with the aim of inferring building thermal properties or are suitable for use as input data to the Deconstruct method (Section 2.5)</td>
</tr>
<tr>
<td>Medium-to-high</td>
<td>Variability in consumption patterns (including enabling analysis of drivers of consumption)</td>
<td>Energy data recorded at a resolution higher than annual coupled with contextual data that can be used for the applications described in Section 2.6</td>
</tr>
<tr>
<td></td>
<td>Research into heating load profiles (This does not have a separate section here as it is considered a subsection of ‘Variability in consumption patterns (including enabling analysis of drivers of consumption)’ but is included as a separate item for the purpose of identifying datasets that collected heating data (presented later in the paper))</td>
<td>Electricity consumption data for electrically heated dwellings (including heat pumps) or gas consumption data for dwellings with gas-fired boilers Only datasets where energy for heating is available separate from other loads are considered suitable</td>
</tr>
<tr>
<td></td>
<td>Impact of electrification of heat</td>
<td>Datasets that can be used to model the impact on the existing electricity network from the electrification of domestic heating systems (Section 2.7)</td>
</tr>
<tr>
<td></td>
<td>Demand flexibility and dynamic electricity tariffs</td>
<td>Datasets that can be used for research into demand-side response or reduction, load flexibility and/or the influence of time-of-use tariffs on energy use (Section 2.8)</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Data Resolution Application</th>
<th>Dataset Suitability Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Magnitude and timing of peak demand</strong></td>
<td>Datasets that enable the timing of peak energy demand across households to be established, including those that contain contextual data that can be used to understand the drivers of peak demand and its relationship with outdoor temperature (if applicable)</td>
</tr>
<tr>
<td><strong>Linking time use with energy demand profiles</strong></td>
<td>Datasets are considered suitable if the temporal resolution is 1 h or better, and measurements are available for at least 1 year (Section 2.9)</td>
</tr>
<tr>
<td><strong>Energy disaggregation</strong></td>
<td>Datasets that contain information on time use of household occupants such that these activities can be linked with energy demand profiles (Section 2.10)</td>
</tr>
<tr>
<td><strong>Evaluating machine learning methods</strong></td>
<td>Datasets with energy load profiles that can be used to test machine learning algorithms to investigate the suitability of different techniques (Section 2.12)</td>
</tr>
</tbody>
</table>

**Low resolution**

2.1. Develop, Monitor and Evaluate Energy Policies

With the aim of improving the quality of residential buildings to both reduce GHG emissions and energy consumption and alleviate fuel poverty, governments have implemented policies and schemes designed to encourage the uptake of energy efficiency technologies and retrofit measures. Examples in the UK (past and present) include The Warm Front Scheme, Green Deal, Energy Company Obligation, Domestic Renewable Heat Incentive, Smart Export Guarantee, and Boiler Upgrade Scheme. To develop, monitor, and evaluate these types of policies, evidence of their effectiveness is beneficial, and metered energy data are one such source [23]. Annual electricity and gas consumption is one form of data that has been used to quantify the energy savings that can be achieved by the implementation of energy efficiency measures. One example is the annually published National Energy Efficiency Database [24], which collects annual gas and electricity consumption data for residential buildings in Scotland, England, and Wales. An estimate of the energy savings achieved through certain energy efficiency measures installed through government-funded schemes is quantified. This allows the effectiveness of the measures in reducing energy demand to be understood, including the variability by property type and how energy savings change over time.

2.2. Benchmarking Annual Energy Consumption

Measured annual energy consumption is used in practice to output energy performance indicators and assign ratings to residential buildings. In this section, the implementation of this in three European countries will be described, and its limitations. Measured annual energy consumption is also used to build databases against which dwellings can be compared to assess their energy efficiency. This form of benchmarking is the basis of schemes in the USA and Australia, both of which will be described.

2.2.1. Energy Performance Certificates

As permitted by the EPBD, measured annual energy consumption is used to assign energy efficiency ratings to residential buildings in Sweden, Germany, and Poland.

In Sweden, real energy consumption can be used to generate an EPC for a newly constructed building or one that has undergone renovation. The non-heating electricity consumed and the energy used for heating, cooling, and domestic hot water (DHW) in one year (collected by an independent assessor within two years of the building being completed) are aggregated, corrected to primary energy, and divided by the heated surface...
area of the building. A correction is made to the heating energy consumption to account for the regional climate [17]. To produce a rating, this energy consumption is compared to the obligatory requirements of new buildings built in the present day. The rating system is on a scale from A to G, where A is $\leq 50\%$ and G is $>235\%$ of the new building’s requirement [25].

Germany’s ‘usage certificate’ is based on the actual energy use of a domestic property over the previous three years. Only the energy used for heating is considered, and the usage certificate is only permitted for residences constructed after 1978 [26]. The energy demand is weather-corrected and divided by the heated floor area to calculate the kWh required to heat 1 m$^2$. In the rating system, $\leq 50$ kWh/m$^2$ is considered ‘good’ and $\geq 400$ kWh/m$^2$ is considered ‘poor’ [27].

In Poland, 36 continuous months of utility bills measuring gas, electricity, or heat can be used to generate an EPC for a dwelling [18].

Weather correction is common for isolating the impact of weather on space heating, cooling, and, to some extent, DHW (usually using the heating or cooling degree day method), as described for the Swedish and German examples. To enable inter-building comparison in the same way that calculation-based EPC methodologies do, the measured energy use should be corrected to standard user behaviour, but the data required to carry this out are not always easily accessible [28]. Therefore, while this measured approach can make the process of constructing an EPC simpler, because data collection can be less invasive and there is no need to make assumptions about properties of the building fabric, it has limitations in terms of standardisation. From 1 July 2021, the French national framework moved from allowing energy ratings to be constructed based on measured energy consumption to solely a calculation methodology, which takes into account the property characteristics, such as the level of insulation [29]. A limitation cited as one reason for the change was the inability to produce an EPC rating for holiday homes as the required data were not available [29] (the metered energy data from three previous years were required [10]).

2.2.2. Other Schemes

In the Swedish example described earlier, an energy rating is calculated by comparing or benchmarking the measured energy consumption against that expected from a theoretical building of a similar type. This comparison against a reference building is one method that can be used to benchmark energy performance. Other benchmarking approaches are historical energy performance, energy performance based on the results of dynamic simulations [17], comparing against national averages [30], and, the primary focus of this review, comparing against the real energy consumption of similar buildings.

The USA Energy Star ‘Home Energy Yardstick’ scheme allows homeowners to compare their 12-month actual energy use with similar homes in an online tool. The result is displayed on a ‘Yardstick’ with a scale from 0 to 10. The higher the score, the better the home performed (used less energy) relative to a similar home over the 12-month period. Linear regression analysis is used to identify the building and occupant variables (data on which are collected as part of a periodically updated Residential Consumption Survey database) that have the greatest impact on energy consumption [31]. Energy consumption is adjusted for the highest impact variables, and the adjusted samples are ranked in the database [31]. This adjustment allows for inter-building comparison and is used by the Home Energy Yardstick online tool when a householder enters the required information. The householder’s adjusted annual energy use is compared with the database-ranked samples to determine a score on the 0 to 10 scale [31]. The NABERS (National Australian Built Environment Rating System) scheme in Australia also uses measured annual energy consumption and linear regression to benchmark the energy performance of similar buildings, including aspects of residential apartment buildings [32,33].

In research, Lomas et al. [13] propose a Domestic Operational Rating scheme for UK dwellings where measured annual energy consumption normalised by floor area is benchmarked against national averages and used to rate their operation. Separately, a
measured energy performance indicator has been proposed by an EU Horizon 2020 project as an innovative feature of a next-generation EPC. The methodology determines the real energy consumption of a building based on measured annual energy use, which includes a weather and standard use correction to enable comparison between buildings [17]. Real energy consumption data also have applications in other benchmarking approaches, including those based on the results of dynamic simulation. Simulation results can be improved by calibrating models with measured annual energy consumption data [14]. By using actual energy data, inputs to simulations are a better match to reality, and the performance gap at the individual building and stock level can be reduced [34].

2.2.3. Benefitting Sectors

The rating and benchmarking approaches described here have the following advantages:

- There is evidence that calculation-based EPCs are viewed as untrustworthy, and recommended measures are disregarded [5,10]. Introducing real energy consumption could mitigate this;
- A comparison of energy use against that of similar dwellings has been shown in research to be perceived as being beneficial to household occupants [35];
- Capturing measured data gives a better characterisation of the true energy consumption of the residential building stock and as-built energy performance, allowing a more reliable estimate of annual energy savings and the economics of retrofit and renovation.

2.3. Variability in Overall Consumption (Including Enabling Analysis of Drivers of Overall Consumption)

Real energy consumption data at the low-resolution scale coupled with household survey data have been used in research to understand the drivers of overall residential energy demand. This is predominantly through statistical analysis (regression methods), which is used to determine the factors that have the greatest influence on overall energy consumption. Five sets of variables are predominantly considered in the literature:

- Dwelling physical characteristics (e.g., dwelling type, dwelling age, and floor area);
- Socio-demographic factors (e.g., household size, income level, and employment status);
- Occupant behaviour;
- Electric appliance stock;
- External conditions (location and weather).

There exists a large body of research investigating the factors that influence overall residential energy demand. The focus of the existing literature differs, and different factors are investigated across studies. Many studies focus on electricity consumption [36–40], with or without heating or cooling loads, whilst others also investigate the drivers of gas consumption [41–43]. Some studies have focussed solely on the factors influencing the energy demand for space and DHW heating [44–47]. In several worldwide studies, dwelling physical characteristics are found to have a significant influence on energy consumption [37,41,48]. However, research focusing on the drivers of heating energy demand emphasises that multiple factors are influential, and the contribution of occupant behaviour is particularly important [44,45]. The relevance of such studies is in identifying the factors that are important contributors to energy use so that they can be targeted for energy efficiency improvement through policy and regulation [49].

Low-to-high resolution

2.4. Validating Assumptions in EPC Calculation

The EPC calculation engine used in the UK makes standard assumptions about energy use to output a relative comparison of energy performance between households. Different authors have used annual energy consumption data and internal temperatures to gather evidence on the heating behaviour of households and validate the assumptions made in the EPC calculation. One example is Hughes et al. [50], who compared the results of EPC-based
modelling with measured annual gas consumption data at the average and individual household levels. They showed that a better agreement between modelled and actual annual gas consumption could be obtained by changing the assumptions about internal demand temperature (from 21 °C to 20 °C), number of heating hours (from nine hours per day on weekdays and sixteen hours per day on weekends to ten hours per day on all days), and heating season duration (from eight months to six months).

Datasets that include empirical measurements of internal temperature have been used to gather evidence on the heating behaviour of households in studies by Huebner et al. [51], Huebner et al. [52], and Oreszczyn et al. [53]. Their results from monitoring heating patterns are in agreement with Hughes et al. [50] in that some EPC assumptions, including a bimodal heating pattern, different heating durations on weekdays and weekends, and a living room heated to 21 °C, commonly do not hold true.

### Medium resolution

#### 2.5. Prediction of Building Thermal Properties

Residential energy demand is driven partly by the properties of the building envelope [21,54,55]. Crawley et al. [10] are referred to for an overview of static and dynamic methods to empirically determine a building’s heat transfer coefficient (HTC) and heating power loss coefficient (HPLC). Both parameters are a measure of thermal performance and, therefore, influence energy demand. The HTC and HPLC (W/m²K) are defined by Equations (1) and (2), respectively [22]:

\[
HTC = \frac{q}{\Delta T} \tag{1}
\]

\[
HPLC = \frac{HTC}{\eta_{HS}} \tag{2}
\]

where \(q\) is the heat flow rate (W/m²), \(\Delta T\) is the internal-to-external temperature difference, and \(\eta_{HS}\) is the efficiency of the heating system.

##### 2.5.1. Heat Transfer Coefficient

The HTC is a key input to the calculation engine to produce an EPC in the UK [56]. For a given HTC, annual energy demands are calculated based on standard use, occupancy and climate assumptions. Required inputs can be based on default values for the given construction type if they are not known. Errors in these assumptions contribute to the performance gap. Therefore, a more robust measurement of the HTC could bring the EPC-calculated annual energy demand performance indicator closer to reality.

A measure of the HTC can be obtained through a co-heating test, but the dwelling is required to be vacant during the test, and the monitoring period is in the order of 1 to 3 weeks [57]. A less intrusive method of quantifying this parameter would be beneficial. This was recognised by the UK Government, which initiated the SMETER (Smart Meter Enabled Thermal Efficiency Ratings) Innovation Programme [58], where organisations were invited to demonstrate technologies that can record data that can be used to infer the HTC of dwellings. This included real energy consumption data recorded through AMI. An analysis of the results found that two of the twelve approaches were able to provide a more accurate prediction of the HTC than that produced by an expert EPC assessment [56].

As implied by Equation (1), the HTC requires a measure of the heat delivered, which is not directly measured in a typical gas-fired boiler central heating system (an assumption would need to be made for the conversion efficiency from gas demand to heat [10,59]). The requirement for a direct measure of heat does, however, lend itself to district heating systems. These are prevalent in Denmark, where 63% of space and DHW heating in private homes is provided through district heating [21]. Publications by Gianniou et al. [21] and Leiria et al. [60] propose methodologies using smart heat data from homes connected to Danish district heating networks to estimate the HTC and building thermal characteristics,
respectively. The methodologies presented by both studies are useful to utility companies to identify inefficiencies in dwellings and in the district heating system itself, and to optimise user behaviour and the network. It has been highlighted that smart heat data research is still at a relatively early stage [60].

2.5.2. Heating Power Loss Coefficient (HPLC)

While the HTC takes into account the thermal performance of the dwelling fabric only, the HPLC additionally considers the efficiency of the building’s heating system to fully characterise heating performance (Equation (2)) [22]. The HPLC metric was first introduced in the ‘Deconstruct’ method by Chambers and Oreszczyn [10,22]. Deconstruct employs steady-state grey-box modelling and requires the daily total metered gas and electricity use (and therefore does not require a back calculation from energy to delivered heat) to estimate the HPLC to within ±15% [22]. Its application enables the use of smart meter data to characterise, at scale and unobtrusively, the in situ thermal performance of residential buildings—a key driver of energy demand. This is proposed to be an improvement on the current methods of assessing the as-built performance of dwellings, including the calculation-based EPC.

2.5.3. Benefitting Sectors

The following are areas that could benefit from estimates of HTC and HPLC based on measurements of real energy consumption:

- A quantification of as-built thermal performance is provided, which can be different from the designed thermal performance, that could form the basis of an ‘empirical EPC’ [61];
- A measure of the parameter(s) could be kept live (e.g., reassessed annually), giving occupiers information about their homes to inform decisions on making energy efficiency improvements [61];
- Evidence can be collected on the effectiveness (or not) of retrofit measures, supporting energy policy [22].

Medium-to-high resolution

2.6. Variability in Consumption Patterns (Including Enabling Analysis of Drivers of Consumption Patterns)

The medium-to-high-resolution data obtained from AMI facilitates research into transient energy demand characteristics. This section is a review of the applications of energy demand data that record the temporal variability of energy load profiles.

2.6.1. Electrical Load Profiles (Non-Heating)

Following on from the analysis of overall residential energy demand in Section 2.3, the study of the temporal variability of electricity load profiles is another active area of research. In this regard, real electricity consumption data are being used to study consumer electricity behaviour and the factors influencing it to identify opportunities for interventions in electricity use practices [62].

There are an extensive number of studies worldwide that have used data mining techniques to group consumers into user groups based on the similarity of electricity load profiles [63–69] and define profiles that are representative of the overall population’s electricity use. Groups could be characterised by the timing and magnitude of peak and minimum demand, relative difference in peak-to-trough magnitude, demand ramp-up time, and daily and seasonal variation. Differences between groups are influenced by physical dwelling and occupant characteristics. Statistical analysis (including regression) is used to understand the most important factors in determining the electricity load profile of a given consumer [62,63,66,67,70]. With the most likely influencing factors identified, dwelling and/or household characteristics can be inferred from consumption data for other homes, or vice versa (the consumption pattern can be inferred from dwelling and/or household
characteristics). Tureczek and Nielsen [71] conducted a comprehensive literature review on the classification of electricity consumption patterns obtained from smart meter data.

2.6.2. Heating Load Profiles

There are fewer published studies that have applied these types of analyses specifically to residential heating load profiles. This is because (i) if a dwelling is electrically heated, the heating load needs to be disaggregated from the electricity profile unless separately sub-metered, and (ii) natural gas boilers are the most common form of heating in some countries (85% of, or 23 million, residential buildings in the UK [72]), but the rollout of smart gas meters is less advanced and the data less accessible than that of electricity [59]. Table 2 summarises the existing research for residential buildings. Locations in Scandinavia have been the subject of several studies to date. Sample sizes tend to be limited relative to the overall size of the building stock.

Table 2. Studies into the temporal variability of heating load profiles based on measurements of real energy consumption (residential buildings only).

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Country</th>
<th>Number of Sites</th>
<th>Data Temporal Resolution</th>
<th>Heating Technology</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>[73]</td>
<td>Denmark</td>
<td>139</td>
<td>1 h</td>
<td>Ground source heat pump (GSHP) and air source heat pump (ASHP)</td>
<td>To identify heating load profiles and how they correlate with household socio-technical characteristics. To develop a methodology to randomly generate simulated thermal power demand profiles from the actual typical load profiles. To record electricity and DHW energy use at 6 s frequency to obtain a higher-resolution dataset than was previously available. To analyse the impact that data resolution has on the consumption values measured. (i) To propose a clustering approach for analysis of district heating consumption data; (ii) to study the correlation between consumption intensity and building and occupant characteristics; and (iii) to cluster normalised daily consumption profiles to identify representative patterns and study their variability. To propose a methodology using clustering to define representative consumption profiles of consumers of natural gas.</td>
</tr>
<tr>
<td>[74]</td>
<td>Denmark</td>
<td>25</td>
<td>15 min</td>
<td>Not specified</td>
<td>To model present and future national domestic heat demand in GB.</td>
</tr>
<tr>
<td>[75]</td>
<td>Sweden</td>
<td>72</td>
<td>6 s</td>
<td>DHW only (no space heating), not specified how heat is delivered</td>
<td>To propose a methodology using clustering to define representative consumption profiles of consumers of natural gas.</td>
</tr>
<tr>
<td>[76]</td>
<td>Denmark</td>
<td>8293</td>
<td>1 h</td>
<td>District heating</td>
<td>To model present and future national domestic heat demand in GB.</td>
</tr>
<tr>
<td>[77]</td>
<td>Ireland</td>
<td>&gt;1000</td>
<td>30 min</td>
<td>Gas-fired boiler</td>
<td>To model present and future national domestic heat demand in GB.</td>
</tr>
</tbody>
</table>
Table 2. Cont.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Country</th>
<th>Number of Sites</th>
<th>Data Temporal Resolution</th>
<th>Heating Technology</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>[79]</td>
<td>USA and Canada</td>
<td>408 and 480</td>
<td>1 day</td>
<td>Gas furnace, electric furnace, and heat pump (USA); heat pump and resistance heating (Canada)</td>
<td>To develop a method that uses smart meter data to extract building thermal characteristics for retrofit analysis.</td>
</tr>
<tr>
<td>[80]</td>
<td>Portugal</td>
<td>19</td>
<td>15 min (integrated to 1 h for analysis)</td>
<td>Electric heater for space heating and cooling (gas for DHW)</td>
<td>To conclude if variations in 1 h electricity consumption data can be used as a proxy for the occupants' space cooling and heating behaviour, and the influence of different minimum and maximum external temperatures.</td>
</tr>
</tbody>
</table>

2.7. Demand Flexibility and Dynamic Electricity Tariffs

The predicted wider electrification of heat, described in the previous section, and other sectors, such as transport, will increase electricity demand [67]. To meet decarbonisation targets, the GHG emissions of the electricity supply itself will continue to decrease. This will mean that increasing proportions of electricity will be generated by renewable sources [81]. Climate-based sources of renewable energy can be intermittent; therefore, coupled with expected greater electricity demand, energy demand flexibility strategies could be necessary to ensure supply and demand balancing [82].

Demand flexibility could take different forms, for example, demand side management (DSM) or offering consumers time-varying (dynamic) electricity tariffs. Energy consumption data at the medium-to-high-resolution scale have been shown to have applications in this space. More accurate load profiles that are based on true consumer behaviour can be generated, which is of benefit to utility companies [64,66,68]. Research shows this can enable better classification of new customers [64], design of tailored energy efficiency campaigns [68,70], and supply and demand management [83] by identifying sets of customers that could be eligible for DSM schemes [63,67,68]. With the knowledge of how energy is being used by groups of consumers, time-of-use tariff structures can be designed and tailored to particular energy-use patterns [62,63,84].

Research into demand flexibility of heating is mainly centred around the use of simulation modelling. With the availability of high-resolution residential energy demand datasets, there is now an effort towards attempting to quantify and generate metrics for heating energy flexibility using measurements of gas and electricity use and internal temperature [85].

2.8. Magnitude and Timing of Peak Demand

Of particular interest in the area of demand flexibility is the issue of peak demand. Load shifting and load shedding as demand flexibility strategies have received considerable attention in publications to date [86]. These have the potential to reduce the magnitude of peak power, which can, in turn, save energy and reduce system operating costs and GHG emissions [86]. To enable these types of demand flexibility in the residential sector, an understanding of peak demand—and the factors that drive it—is required, but Torriti [82] highlights that little is known about this.

While energy demand data analysed at the medium-resolution scale can quantify the magnitude of peak demand, high-resolution energy demand data (hourly or better) can give insights into the timing of peak demand and how this varies across groups of consumers. Coupled with qualitative data about the household, greater insight into peak demand can
be obtained. Examples of studies include the following: using 30 min smart meter electricity data collected in Ireland to identify the dwelling and occupant characteristics with the greatest influence on timing and magnitude of peak demand and the electrical appliances with the greatest load shifting potential [36]; using 1 h district heating smart meter data from Denmark to categorise profiles based on the characteristics of their peaks and applying numerical modelling to investigate peak load shifting and quantify the potential rate of peak load reduction [87]. The relationship between peak load and external temperature is also of interest—heat pump field datasets have been used to model the impact that different external temperatures and different penetrations of heat pumps have on peak load [88].

2.9. Linking Time Use with Energy Demand Profiles

In addition to the influence of dwelling and/or occupant characteristics on energy demand profiles (described in Section 2.6), it is desirable to have an understanding of the occupant activities that drive particular demand profiles—particularly the peaks—if demand side management is to be achieved [68,89]. To characterise the relationship between occupant activity and residential energy demand, research has been carried out into linking activity to energy demand with the objective of identifying the key activities that drive energy consumption and help identify where demand reduction or response can be implemented [68,82,90–92].

The UK Household Electricity and Activity Survey, 2016–2019 [93], had the objective of collecting data to provide an understanding of the relationship between occupant activities and electricity demand patterns [68]. Several publications have stemmed from the dataset, exploring the activities that have the greatest influence on electricity demand, prediction of household electricity consumption, peak demand reduction and load shifting, the influence of different electricity tariff structures, and the effect of energy efficiency interventions on household occupants [68,94,95].

The UK Household Electricity and Activity Survey dataset fits within the broader category of time-use surveys (TUSs). These have been used by researchers to link occupant activity to energy demand and generate deterministic and probabilistic occupancy-dependent energy demand profiles that can be used in dynamic building simulation tools [96,97]. The IEA-EBC Annex 66 [98] states the importance of integrating occupant behaviour into building simulation because the accuracy of building energy demand predictions will be improved if the energy behaviours of occupants are more realistic [97]. TUSs are, however, carried out infrequently [82], and the information is typically collected over short periods (one or two days) [91]. Real energy consumption data from smart meters, for example, offer an improvement over TUSs in that they could be used to derive data-driven transient occupancy schedules for input into a dynamic simulation [20,99]. This is less intrusive than conducting a TUS because the data can be gathered over longer periods and for a larger number of buildings. Anonymisation of this type of data at the individual building level would be important, however, due to the level of detail that could be ascertained relating to a household’s day-to-day schedule.

2.10. Energy Disaggregation

The availability of data at the individual appliance level has numerous benefits across the value chain. Armel et al. [100] describe the benefits of data at this granularity at the consumer (e.g., appliance level feedback has been shown to result in energy savings), research and development (e.g., appliances could be redesigned to improve energy efficiency), and policy (e.g., targeting energy efficiency programmes) levels. However, sub-metering of appliances can be impractical [101]. Most commonly, energy use is monitored at the whole house level. Thus, to obtain end-use data, disaggregation techniques can be applied to extract it from the aggregated signal [90]. UK datasets that have appliance-level data contain a relatively small number of buildings in the samples [101–105]. So, although sub-metering can be intrusive, datasets containing this information are necessary to validate disaggregation techniques that are used to extract appliance-level consumption from over-
all load profiles [101,106]. And in turn, where appliance-level data are not available, these real-data-informed disaggregation techniques can be applied to aggregated whole-house energy use to target efficiency improvements or demand management [90].

2.11. Impact of the Electrification of Heat

Following on from the previous section regarding variability in energy consumption patterns, a specific application of this type of research is in investigating and predicting the impact on heat and electricity demand and on electricity generation, transmission, and distribution requirements through the electrification of heat at local or national levels. This is in the context of an expected increase in the use of heat pumps as a technology to provide space and water heating in the effort to decarbonise heating systems.

Rather than using building simulations with synthetic load profiles to model the effect of heat pump uptake, measured energy data can be used, which captures the real intra-building diversity of consumers [59]. Knowledge about the current timing and magnitude of consumption is important for planning grid requirements as the gradual electrification of heat will increase demand on existing electricity networks [78,107].

Data from other forms of heating (for example, conventional gas heating, direct electric, or potentially other forms such as district heating) could theoretically be used to inform predictions of future electricity demand from the widespread rollout of heat pumps [108]. However, authors argue that due to the difference in the operation of a heat pump compared to other technologies (for example, operating at a lower temperature and longer operating period), real-world datasets monitoring the energy consumption of heat pumps themselves are then the most reliable to predict future demand scenarios [78,107].

2.12. Evaluating Machine Learning Methods

To facilitate the study of large high-resolution energy demand datasets, advanced data analysis techniques using machine learning can be employed [63,109]. A popular analysis technique for energy demand research is clustering, which provides a data-driven grouping of previously unknown patterns of consumer energy load profiles [110]. An overview of worldwide research on the clustering of electric load profiles is presented by Satre-Meloy et al. [68]. Datasets with large sample sizes have been used to test different machine learning algorithms to explore the differences in results and make recommendations for applicability to other studies [63,65,111,112]. These types of investigations help direct and advance future research. The high-resolution UK IDEAL (Intelligent Domestic Energy Advice Loop) household energy dataset [113] was collected, with one of the primary aims being to advance and evaluate methods of machine learning that can be applied to residential energy demand data [102].

3. UK Residential Energy Demand Datasets and Their Applications

This review has shown that there are numerous applications of real energy consumption data across a range of temporal resolutions that have been used to characterise and develop an understanding of residential energy demand. This section will present residential energy demand datasets that have been collected in the UK and identify their suitability for the applications described in Section 2 and presented in Table 1.

3.1. UK Residential Energy Demand Datasets

Residential energy demand datasets collected in the UK were identified through the review presented in Section 2 and a search of the UK Data Service (UKDS) [114]. The twenty-three datasets are listed in Table 3. It should be noted that the scope is limited to datasets that are openly accessible to researchers or accessible through defined processes; proprietary datasets are excluded. None have been excluded based on useability or data quality. The temporal resolution of the available data ranges from annual to 1 s. Further detail on the datasets, including if any additional contextual information was collected (i.e., on the categories of dwelling characteristics, occupant characteristics, electrical appliances,
and weather or internal temperatures), is provided in Appendix A. All but one dataset contains at least one category of contextual data, with five datasets collecting data related to all five categories.

Table 3. UK residential energy demand datasets accessible to researchers.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Temporal Resolution of Energy Data</th>
<th>Notes on Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Energy Efficiency Data (NEED) framework [115]</td>
<td>Annual gas and electricity consumption</td>
<td>Consumption data tables are available for download from UK Government website. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Energy Follow Up Survey (EFUS), 2011 [116]</td>
<td>Annual gas and electricity consumption; 10 s electricity for a subset of 79 homes</td>
<td>Data accessed through the UKDS SecureLab. Researchers are required to achieve accreditation through an approval process and the data can only be accessed inside a secure environment [118]. Data are available for download through the UKDS. No requirement for registration.</td>
</tr>
<tr>
<td>Smart Energy Research Lab (SERL) Observatory Data, 2019–2021 [117]</td>
<td>Daily and 30 min gas and electricity consumption</td>
<td>Data held by Loughborough University [120].</td>
</tr>
<tr>
<td>SMETER Technologies Project Phase 2 Data [119]</td>
<td>30 min gas and electricity consumption</td>
<td>Data are available for download from CLNR website through Creative Commons Attribution-ShareAlike Licence [121].</td>
</tr>
<tr>
<td>DEFACTO (Digital Energy Feedback and Control Technology Optimisation) Field Trial [120]</td>
<td>30 min (gas); 2 min (electricity)</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Customer Led Network Revolution (CLNR) [121]</td>
<td>30 min electricity consumption</td>
<td>Data are available for download from CLNR website through Creative Commons Attribution-ShareAlike Licence [121].</td>
</tr>
<tr>
<td>CLNR enhanced profiling of domestic customers with air source heat pumps (ASHP) [121]</td>
<td>30 min (household electricity consumption) and 1 min (heat pump and household electricity consumption)</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>REFIT Smart Home Dataset [122]</td>
<td>30 min gas consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>SmartMeter Energy Consumption Data in London Households [123]</td>
<td>30 min electricity consumption</td>
<td>Data are available for download from London Datastore [128].</td>
</tr>
<tr>
<td>Energy Demand Research Project (EDRP), 2007–2010 [124]</td>
<td>30 min gas and electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Domestic Heating Data from the Energy Systems Catapult Living Lab [125,126]</td>
<td>15 min gas and electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Solent Achieving Value from Efficiency (SAVE) Data, 2017–2018 [127]</td>
<td>15 min electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Low Carbon London (LCL) Project Heat Pump (HP) Load Profiles [128]</td>
<td>15 min electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Household Electricity Survey 2010–2011 [103]</td>
<td>10 min and 2 min electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>North East Scotland Energy Monitoring Project (NESEMP), 2010–2012 [129]</td>
<td>5 min electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
<tr>
<td>Renewable Heat Premium Payment (RHPP) [107,130]</td>
<td>2 min electricity consumption</td>
<td>Data are available for download through the Loughborough University repository [122]. Data are available for download from London Datastore [123]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Issues reported with data management and processing during project [59]. Data are available through the UKDS for registered users subject to the End User Licence Agreement. Access is restricted to users residing in European Economic Area. Data are available through the UKDS for registered users subject to the End User Licence Agreement.</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Temporal Resolution of Energy Data</th>
<th>Notes on Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornwall Local Energy Market (LEM) Residential Electricity Dataset with Solar Production and Battery Storage, 2018–2020 [132]</td>
<td>1 s and 1 min electricity consumption</td>
<td>Data are available for download through the UKDS for registered users</td>
</tr>
<tr>
<td>Low Effort Energy Demand Reduction (LEEDR) [104]</td>
<td>1 min (electricity, gas, and DHW consumption); 1 s (gas and DHW consumption)</td>
<td>Data are available for download through the Loughborough University repository [104]</td>
</tr>
<tr>
<td>One-Minute Resolution Domestic Electricity Use Data, 2008–2009 [133]</td>
<td>1 min electricity consumption</td>
<td>Data are available through the UKDS for registered users subject to the End User Licence Agreement</td>
</tr>
<tr>
<td>REFIT Electrical Load Measurements [105]</td>
<td>8 s electricity consumption</td>
<td>Data are available for download through University of Strathclyde [105]</td>
</tr>
<tr>
<td>UK Domestic Appliance-Level Electricity (UK-DALE) [101]</td>
<td>6 s electricity consumption. Three homes also have 1 sec whole-home active power and apparent power</td>
<td>Data are available for download through the UK Energy Research Council Energy Data Centre under Creative Commons Attribution 4.0 International license [134]</td>
</tr>
<tr>
<td>Measuring and Evaluating Time-use and Electricity-use Relationships (METER): UK Household Electricity and Activity Survey, 2016–2019 [95]</td>
<td>1 s electricity down sampled to 1 min and 10 min mean values</td>
<td>Data are available through the UKDS for registered users subject to the End User Licence Agreement Additional activity data can be accessed by UK-based users through a Secure Access application</td>
</tr>
<tr>
<td>Intelligent Domestic Energy Advice Loop (IDEAL) Household Energy Dataset [113]</td>
<td>1 s (electricity); 1 reading per 1 dm³ or 1 ft³ (gas)</td>
<td>Data are available for download through The University of Edinburgh [113]</td>
</tr>
</tbody>
</table>

The project timeline, temporal resolution, and type of energy data available for the datasets are shown in Figure 1. Where datasets comprise measurements at different temporal resolutions (for example, DEFACTO and Cornwall LEM), the end points of the black dashed lines in Figure 1 represent the highest and lowest temporal resolution. The project timeline, indicated by the x-axis bars, is the overall project duration or the timeframe for which data have been made available, but individual sites may have different monitoring periods within this timeline. The x-axis bars are coloured by order of magnitude of the number of sites monitored by the project, and the pattern of the bars indicates the type of energy data collected (gas, electricity, or both).

The majority of datasets contain data that have a temporal resolution of up to 30 min, and five datasets recorded data at 1 s resolution over relatively long time periods, albeit for lower spatial resolution (number of sites in the tens-or-fewer to hundreds scale). It can be noted that there are more datasets with a number of sites in the order of tens or fewer to hundreds than datasets with larger numbers of sites (Figure 2); just slightly over one-quarter of the twenty-three datasets have sample sizes in the order of thousands or more. To put this into context, in 2021, there were estimated to be 28.1 million households in the UK [135]. To investigate this further, data temporal resolution was cross plotted against the number of sites (Figure 3). Although there is scatter of the data points, in general, the higher-resolution datasets tend to have been collected for fewer sites. This may suggest that project investigators are cognisant of the trade-off between collecting data at high temporal and spatial resolutions. The choices made within this trade-off, and ensuring statistically representative conclusions can be made from a manageable data collection exercise, will tend to be linked to the specific outcomes of any associated project.
Figure 1. Temporal resolution versus project timeline for UK residential energy demand datasets listed in Table 3. For clarity, overlapping datasets have been shifted slightly on the y-axis.
3.2. Mapping Dataset to Application

The datasets listed in Table 3 were often collected to meet particular project objectives, but the data resolution and duration of the collection period means that datasets could be used in applications outside of their original intended focus. Section 2 and Table 1 presented groupings of thirteen applications of real energy consumption data. For example, applications related to demand reduction, response, management, flexibility, or investi-

Figure 2. Distribution of number of sites monitored in the datasets listed in Table 3. Percentage contributions within the sample size of twenty-three are shown.

Figure 3. Data temporal resolution versus number of sites. Datasets are plotted more than once if data were collected at different temporal resolutions.
gating the effect of time-of-use tariffs were grouped into ‘Demand flexibility and dynamic electricity tariffs’.

The suitability of each UK residential energy demand dataset (listed in Table 3) for each data application (listed in Table 1) was determined. This was based on the suitability criteria described in Table 1, which includes the temporal resolution of the data collected and the time period of data collection. Although covered in this literature review, the application of smart heat data has not been included here because the most common heating systems in the UK do not enable direct measurement of delivered heat.

The suitability of each identified energy demand dataset for a given application is shown in Figure 4. The application for which the greatest number of datasets are found to be suitable is in analysing the variability of consumption patterns and the factors that influence this variability. The LEEDR, DEFACTO Field Trial, and Household Electricity Survey datasets are found to be useful to the greatest number of applications identified here.
Figure 4. UK residential energy demand datasets that could be suitable for each data application.
4. Discussion and Conclusions

As identified in this review, at a top-down, non-granular level, real energy consumption recorded at annual resolution can be used to benchmark energy consumption and rate the energy efficiency of an in-use residential building. With datasets available in the UK that are coupled with contextual data, there could be an opportunity to create a scheme that would output an operational rating and enable households to compare their energy use with other buildings of a similar type. Performance indicators (such as annual kWh/m²) would then be based on actual energy demand rather than on a theoretical calculation where the underlying assumptions might misrepresent the energy use practices of a household. Datasets forming the basis of such a scheme would need to be checked for their representativeness of the residential building stock as a whole. However, it has been highlighted that in the EU, there are instances of a deliberate move away from this type of energy rating scheme, seemingly due to the problems that this presents for purposes of standardisation and cross-comparison.

The high temporal resolution of data obtained from AMI has been used to provide insights into both the as-built thermophysical properties of the building fabric and energy use patterns (including the physical and socio-demographic driving factors of those patterns) across residential buildings. It has been highlighted that research focusing on electricity use is prevalent, but smart heat data and heating loads, particularly in terms of the diversity of heating patterns, have been less widely studied. More than half of the UK energy demand datasets identified in this paper collected electricity data only, and although there will be some occurrences of electric heating within these samples, sub-metering to isolate the heating load is uncommon. However, thirteen datasets that could be used for some form of research into UK heating load profiles have been identified, primarily for gas-fired boiler heating systems or heat pumps, with fewer instances of electric space or DHW heating. Datasets that can enable research focused on the use of direct electric heating technologies, such as electric radiators and storage heaters, are mostly missing from the datasets identified here.

This paper has aimed to review the applications of real energy consumption data in characterising residential energy demand. To present a comprehensive review, this was not restricted to any particular energy end use; heating and non-heating energy use and electricity, gas, and heat data have all been included. The review encompasses what could be considered well-established applications of real energy data (for example, in generating an EPC in some EU Member States) to state-of-the-art applications that take advantage of modern computing power (for example, through machine learning and dynamic simulations). Complementary to this first aim, UK residential energy demand datasets available for use by researchers have been comprehensively reviewed and inventoried. Again, this was not restricted to a particular data type—the inventory includes datasets that collected data on electricity, gas, or both. Twenty-three UK datasets, with temporal resolution ranging from 1 s to annual, have been identified, and the inclusion of any additional data that can provide context to the energy data itself has been noted. Through this review, thirteen applications of real energy consumption data were identified. To combine the findings, energy data type, temporal and spatial resolution, and the existence of contextual information of each energy demand dataset were considered to determine its suitability for each of the thirteen applications. The LEEDR, DEFACTO Field Trial, and Household Electricity Survey datasets were deemed to be useful for the greatest number of applications. It is hoped that this data inventory and mapping to applications will be of benefit to researchers by increasing awareness of UK datasets that are available upon which to carry out studies related to characterising residential energy demand. Furthermore, where potential applications for datasets beyond the original project objectives have been highlighted, this could add value to the dataset and enable greater insight to be extracted.

This paper has attempted to highlight gaps in the currently available datasets and show where future data collection efforts may wish to focus. Higher temporal resolution
datasets tend to be collected for a smaller number of sites; therefore, future projects may endeavour to achieve both high temporal and spatial resolution to enable deeper demand analysis. Another point to note is that despite data being released into the public domain, it is not a given that the data quality will be adequate for immediate use. It is possible that substantial data processing could be required to obtain a dataset that is suitable for analysis, and so issues of useability and data quality could be obstacles to use. The accessibility of each dataset has been described, and there are instances where access must be sought through formal approval processes. A second challenge in relation to access is that there may be a time delay between the collection of a dataset and its release to the wider research community until the data collectors have maximised value from the dataset themselves. Although there are many areas of research that propose ways to take advantage of the rich content of empirical energy databases, few are in widespread practice within the UK. Sectors or policy areas benefitting from the characterisation of residential energy demand have been identified from across the value chain, ranging from individual consumers to utility companies and government bodies. However, more research is required to determine the outputs from energy demand data that would benefit end users in terms of the data resolution required by different end users and its visual representation. These are intended to be the focus of future research.

Author Contributions: Conceptualisation, D.J.; methodology, L.T.; formal analysis, L.T.; writing—original draft preparation, L.T.; writing—review and editing, D.J. and L.T.; supervision, D.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: This work is part of an Engineering and Physical Sciences Research Council-funded PhD project.

Conflicts of Interest: The authors declare no conflict of interest.
## Appendix A. Inventory of UK Residential Energy Demand Datasets

Table A1. Residential real energy consumption datasets collected in the UK, ordered from lower to higher temporal resolution. Individual sites can have different monitoring periods within the overall project timeline. ‘Appliances’ includes data collected through questionnaires on ownership and use of electrical appliances. Dw.—dwelling; Occ.—occupant; App—appliances; and Tint—internal temperature. Weather data include data that were collected by project investigators or third-party data. The “x” indicates availability of a particular type of data within the dataset.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Location</th>
<th>Number of Sites</th>
<th>Electricity</th>
<th>Gas</th>
<th>Heating Loads</th>
<th>Heating Technology</th>
<th>Project Timeline</th>
<th>Data Temporal Resolution</th>
<th>Contextual Data Gathered</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Energy Efficiency Data Framework (NEED) [115]</td>
<td>England and Wales, UK</td>
<td>&gt;4,000,000</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Not detailed but includes gas-fired boiler</td>
<td>2005–2012</td>
<td>Annual</td>
<td>x</td>
</tr>
<tr>
<td>Energy Follow Up Survey (EFUS), 2011 [116]</td>
<td>England, UK</td>
<td>1345 (electricity); 1197 (gas)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler; electric storage radiator; and gas or solid fuel fire</td>
<td>Dec. 2010–Jan. 2012</td>
<td>Annual gas and electricity; 10 s electricity for subset of 79 homes between Apr 2011 and Jan 2012</td>
<td>x x x x x</td>
</tr>
<tr>
<td>Smart Energy Research Lab (SERL) Observatory Data, 2019–2021 [117]</td>
<td>Scotland, England and Wales, UK</td>
<td>&gt;13,000</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler; electric storage radiator; electric radiator; heat pump; district heating; oil; solid fuel; and biomass</td>
<td>Aug. 2018–Dec. 2021</td>
<td>Daily and 30 min</td>
<td>x x x x</td>
</tr>
<tr>
<td>Digital Energy Feedback and Control Technology Optimisation Field Trial (DEFACTO) (main study) [120]</td>
<td>Midlans, England, UK</td>
<td>393 at beginning and 155 at end</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler</td>
<td>2015–2018</td>
<td>30 min (gas); 2 min (electricity)</td>
<td>x x x x x</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Location</th>
<th>Number of Sites</th>
<th>Electricity</th>
<th>Gas</th>
<th>Heating Loads</th>
<th>Heating Technology</th>
<th>Project Timeline</th>
<th>Data Temporal Resolution</th>
<th>Contextual Data Gathered</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB</td>
<td>−11,000</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>Not detailed but includes ASHP</td>
<td>2011–2014</td>
<td>30 min</td>
<td>x (External temperature recorded for properties with ASHP only)</td>
</tr>
<tr>
<td>GB</td>
<td>89</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>ASHP</td>
<td>May 2013–Apr. 2014</td>
<td>30 min (household) and 1 min (HP and household)</td>
<td>x</td>
</tr>
<tr>
<td>Loughborough, England, UK</td>
<td>20</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler</td>
<td>Feb. 2014–June 2014</td>
<td>30 min</td>
<td>x</td>
</tr>
<tr>
<td>Scotland, England, and Wales, UK</td>
<td>18,370</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler; electric heating</td>
<td>2007–2010</td>
<td>30 min</td>
<td>x</td>
</tr>
<tr>
<td>London, England, UK</td>
<td>4000</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>Not detailed, but heat sources are electric, gas, solid, biomass, and ‘other’</td>
<td>Jan. 2017–Dec. 2018</td>
<td>15 min</td>
<td>x</td>
</tr>
<tr>
<td>London, England, UK</td>
<td>19</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>ASHP, GSHP</td>
<td>Dec. 2011–Mar. 2014 (9 sites); Jan. 2014–Mar. 2014 (10 sites)</td>
<td>15 min (9 sites); 10 min (10 sites)</td>
<td>x</td>
</tr>
<tr>
<td>Dataset</td>
<td>Location</td>
<td>Number of Sites</td>
<td>Electricity</td>
<td>Gas</td>
<td>Heating Loads</td>
<td>Heating Technology</td>
<td>Project Timeline</td>
<td>Data Temporal Resolution</td>
<td>Contextual Data Gathered</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-----</td>
<td>---------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Household Electricity Survey 2010–2011 [103]</td>
<td>England, UK</td>
<td>250</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler; electric heating (less common)</td>
<td>May 2010–July 2011</td>
<td>10 min and 2 min</td>
<td>x x x x x x</td>
</tr>
<tr>
<td>North East Scotland Energy Monitoring Project (NESEMP), 2010–2012 [129]</td>
<td>Aberdeen and Aberdeenshire, Scotland, UK</td>
<td>215</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Oil-fired boiler; LPG fired-boiler; gas-fired boiler; electric heating; electric heating; biomass; and solid fuel fire</td>
<td>2010–2012</td>
<td>5 min</td>
<td>x x x x</td>
</tr>
<tr>
<td>Renewable Heat Premium Payment (RHPP) [107,130]</td>
<td>UK</td>
<td>417 ( Those that met thresholds for quality and completeness of data (“B2 dataset”) [130]–full sample is 703 sites)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>ASHP (318); GSHP (99)</td>
<td>Dec. 2011–Mar. 2015</td>
<td>2 min</td>
<td>x x x x</td>
</tr>
<tr>
<td>Cornwall Local Energy Market (LEM) Residential Electricity Dataset with Solar Production and Battery Storage, 2018–2020 [132]</td>
<td>Cornwall, England, UK</td>
<td>100</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Gas-fired boiler; electric storage heater; other electric; ASHP (5); and GSHP (1)</td>
<td>Apr. 2018–Dec. 2020</td>
<td>1 s and 1 min</td>
<td>x x x x</td>
</tr>
<tr>
<td>Low Effort Energy Demand Reduction (LEEDR) [104]</td>
<td>Loughborough, England, UK</td>
<td>20</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Space: gas-fired boiler; DHW; gas-fired boiler; hot water cylinder; and electric shower</td>
<td>2011–2014</td>
<td>1 min (electricity, gas and DHW) and 1 s (gas and DHW)</td>
<td>x x x x x x</td>
</tr>
</tbody>
</table>

Table A1. Cont.
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Location</th>
<th>Number of Sites</th>
<th>Electricity</th>
<th>Gas</th>
<th>Heating Loads</th>
<th>Heating Technology</th>
<th>Project Timeline</th>
<th>Data Temporal Resolution</th>
<th>Contextual Data Gathered</th>
<th>T&lt;sub&gt;int&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-Minute Resolution Domestic Electricity Use Data, 2008–2009</td>
<td>East Midlands, England, UK</td>
<td>22</td>
<td>x</td>
<td></td>
<td>x</td>
<td>Space: gas or oil-fired boiler. DHW: gas or oil-fired boiler; electric shower</td>
<td>Jan. 2008–Dec. 2009</td>
<td>1 min</td>
<td>x x x</td>
<td></td>
</tr>
<tr>
<td>UK-DALE (UK Domestic Appliance-Level Electricity) [101]</td>
<td>London, England, UK</td>
<td>5</td>
<td>x</td>
<td></td>
<td></td>
<td>Gas-fired boiler</td>
<td>Nov. 2012–Jan. 2015</td>
<td>6 s. Three homes also have 1 s whole-home active power and apparent power</td>
<td>x x x</td>
<td></td>
</tr>
<tr>
<td>IDEAL (Intelligent Domestic Energy Advice Loop) Household Energy Dataset [113]</td>
<td>Edinburgh, Lothians, and south Fife, Scotland, UK</td>
<td>255 (39 sites had sub-monitoring of a selection of electrical appliances and other more detailed monitoring)</td>
<td>x</td>
<td>x</td>
<td>x x x x</td>
<td>Gas-fired boiler</td>
<td>Aug. 2016–June 2018</td>
<td>1 s (electricity); 1 reading per 1 dm&lt;sup&gt;3&lt;/sup&gt; or 1 ft&lt;sup&gt;3&lt;/sup&gt; (gas)</td>
<td>x x x x</td>
<td>x</td>
</tr>
</tbody>
</table>


82. Torriti, J. Temporal aggregation: Time use methodologies applied to residential electricity demand. *Util. Policy* 2020, 64, 101039. [CrossRef]


86. Li, F.; Li, D. Clustering Load Profiles for Demand Response Applications. *Energies* 2022, 15, 5304. [CrossRef]


89. Torriti, J. Understanding the timing of energy demand through time use data: Time of the day dependence of social practices. *Energy Res. Soc. Sci.* 2017, 25, 37–47. [CrossRef]


96. Flett, G.; Kelly, N. Modelling of individual domestic occupancy and energy demand behaviours using existing datasets and probabilistic modelling methods. *Energy Build.* 2021, 252, 111373. [CrossRef]


114. UKDS. Available online: https://ukdataservice.ac.uk/ (accessed on 1 July 2022).


118. SERL. Accessing SERL Data. Available online: https://serl.ac.uk/researchers/ (accessed on 1 July 2022).


Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.