Modelling community electricity demand for UK and India

D.P. Jenkins*, S. Patidar, P. McCallum, K.B. Debnath

Institute of Sustainable Building Design, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh, EH14 4AS, UK

ARTICLE INFO

Keywords:
- Building modelling
- Energy systems
- Community energy

ABSTRACT

The energy assessment of single buildings and of larger areas of built environment, although exhibiting similarities in terms of technique, have in the past often used different approaches to energy modelling. The growing availability of empirical data and the capability of building modelling software has, more recently, allowed these differences to be reduced. This paper demonstrates, across two very different case-studies in UK and India, that techniques for community energy modelling can be used in a way that maintains detail in energy demand characteristics, thus helping to bridge the gap between detailed building assessment and higher-level energy system modelling. However, understanding the portability of such techniques requires an understanding of energy characteristics that can be specific to a geographic area. This study documents these important differences and proposes a more transferrable approach to detailed community energy modelling.

1. Introduction

Our energy systems are relatively complex structures encompassing the supply, transmission, distribution and demand of energy. When based on known, well-understood parameters, such systems perform well and are robust. However, developing urban landscapes, changes in building technology, climate change and changes in energy practices can, in combination, create an uncertain picture from which future strategies of energy provision have to be formulated. Rather than relying on singular, deterministic predictions of our energy futures, it is arguably more important to develop a range of tools and techniques that are flexible and adaptable enough to cope with a range of futures, thus providing key end-users with information from which sensible decisions can be made.

Concerns about future uncertainties are present in all countries. In the UK, the use of energy system models such as MARKAL/TIMES (ETSAP, 2008; Taylor, Upham, McDowall, & Christopherson, 2014) are prevalent to provide a policy-conversant, high-level picture of how energy supply/demand can be optimised. Additionally, the use of future scenario projections, such as those delivered by the National Grid (National Grid, 2017b), provide some estimation of cause-and-effect within our structure of energy provision. Barriers and challenges will still need to be overcome for a resilient energy supply to be achieved in a low-carbon future, but such problems exist within a relatively data-rich, and model-rich, country.

This is in contrast to some other parts of the world which, as well as having different data and modelling landscapes, are subject to a much steeper gradient of change. India, for example, has seen a recent growth in urbanisation and forms of distributed generation within the local energy infrastructure. Over 40 % of the population in India will be urban by 2030 (Roy et al., 2018), creating socio-cultural changes with implications for energy practices. Stresses with respect to clean energy provision and network management are emerging as key challenges but, conversely, this sheer pace and scale of change presents opportunities for developing demand reduction strategies married to the evolving energy infrastructure. Through this period of evolution, tailored approaches are required to determine practicable and quantifiable guidance for selecting energy demand measures in residential buildings that are cognisant of evolving infrastructure requirements and also the needs and constraints of the building occupant. Without a modelling framework that reflects these uncertainties, the risk is that demand characteristics develop that are discordant to changes in supply occurring over the same timescale.

Whilst supply-demand matching issues caused, or influenced, by a low carbon transition may be country-specific, the methods, tools and expertise used to address these problems may in part be translatable across different regions. This study aims to explore where techniques may be similarly beneficial to two very different countries, in UK and India. However, by looking at important differences in these countries, the limitations of translating UK-developed tools will also be explored, noting how such problems may be addressed. In this way, the extent to which energy demand models can be generalised and ported across to different geographical areas will be presented. The benefits of being able to do so is clear, with the potential for knowledge exchange and
extrapolation of application.

Informing this study will be the research conducted by the authors across two active projects. The £20 M UK National Centre for Energy Systems Integration (CESI) (CESI, 2018) is a large consortium investigating modelling solutions to capture how the UK energy system may evolve in the coming decades. Using a suite of new and existing modelling techniques, informed by empirical data, CESI aims to provide guidance to inform future government policy for optimising energy networks in the UK. Amongst several case-studies, the project has extensive data from the Findhorn EcoVillage (Findhorn, 2018), where energy use and behaviour of occupants are being monitored.

The project Community-scale Energy Demand Reduction in India (CEDRI) (CEDRI, 2018) aims to propose energy demand reduction solutions for residential areas of India that are sensitive to the building stock, householders, and constraints on the local energy networks. Within this diverse country, chosen case-studies are being used to test some of the modelling options for estimating the impact of various demand reduction strategies. This includes the village of Auroville (2018).

This paper will use the case-studies of Findhorn and Auroville to help explain the different challenges for energy demand models in these two different countries. Whilst these case-studies should not be seen as being statistically representative of their respective countries (and, indeed, no single case-study should ever be presented as thus), they do offer country- and climate-specific problems to test some of the developed tools of the aforementioned projects. The paper will provide context to these case-studies by reviewing energy assessments on a wider scale, such as those used in energy system/network analysis. The full spectrum of energy modelling, across scales, will therefore be discussed for these two different locations.

The CESI and CEDRI projects are looking at several different energy vectors, but this particular paper will focus on electrical demand modelling of residential buildings, albeit within wider objectives of future work in the aforementioned projects.

2. Review of energy modelling of communities across different locales

2.1. Understanding local energy landscapes

As discussed below, data availability and research into the energy use of specific building stock will vary with country. Projections of changes within those countries can be similarly diverse. To reflect the work of CEDRI and CESI projects, this paper will focus on the UK and India as examples of such differences. There will be particular attention paid to residential buildings, due to the impact of such buildings on the pressures faced by local energy networks.

2.1.1. Energy use in UK buildings

In 2015, residential buildings (approximately 27.2 million (ONS, 2018)) accounted for 29% of total energy consumption in the UK (BEIS, 2017b). Approximately 14% of total electricity usage in the UK is due to residential buildings (BEIS, 2017a), not including electricity consumed due to infrastructure losses (which exist, in part, to deliver that electricity to homes and other buildings). Whilst year-on-year variations are common, and often driven by weather, there is a longer-term trend in the aforementioned publications where annual gas usage (a 21% decrease in 2016 compared to a peak in 2004) and electricity usage (13% in the same time period) are decreasing significantly. This has been affected by well-recorded energy efficiency improvements (Decc, 2012) although there are also socio-economic factors that are important to note (Jones & Lomas, 2015).

However, to characterise energy demand of residential buildings specifically within the context of energy systems requires information on both the building stock and energy network data (as well as several other data source pertaining to the occupants). It is also evident that, for the type of energy system modelling analysis proposed in this paper, a researcher is required to access data at different spatial scales, ranging from national scale to building-specific scale, by way of interim levels of regional and community energy demand. In the UK, data representing the building stock is relatively well documented, within known limitations of characterising diverse building and household types. For designing tailored energy efficiency guidance, and related subsidy/support schemes, such information is of great importance.

A range of data sources exist for documenting the effect UK buildings have on total energy use across the country. In addition to the modelling approaches of Section 2.2, empirical data (and semi-empirical data that has undergone some degree of inference or extrapolation) exists that can identify both fuel source and end-use. For example, the Digest of UK Energy Statistics (DUKES) (BEIS, 2017a) uses data for fuel trading (and energy/commodity balancing) to gain an understanding of the carbon intensity of different sectors in the UK (buildings, industry, transport etc). Generally speaking, as further disaggregation is sought, the reliance on modelling increases where empirically measured disaggregation is not available – though this can be carried out in such a way that the modelling results of, say, energy use in residential buildings is consistent with top-down datasets such as DUKES.

Electrical demand data is available at different scales, though accessibility becomes more difficult as the scale becomes more building-focussed. National demand profiles, informed by empirical data, are widely used by the like of the National Grid and regional data is also available in different forms (including longer-term projections (National Grid, 2017a)). Substation data (at Low Voltage network level) can give profiles of electricity use for communities of residential buildings (e.g. ∼200 dwellings) but will also exhibit other electrical loads not emanating from residential buildings. Availability of such data is not always in the public domain, though the National Energy Efficiency Database (NEED) (DECC, 2011) provides a more accessible example of gas and electricity consumption data in the residential and non-residential sectors. Empirically, the energy data landscape is improving with the use of smart meters, though this provides a “Big Data” challenge, where we need to characterise temporally precise electrical demand for a statistically significant number of buildings for a given region. This is quite a different challenge to that of traditional stock modelling (Section 2.2), which is generally not used for specifying variations beyond the monthly timescale, and therefore does not require the same quality or detail of sample. The difficulty in getting data that is over a long enough duration (e.g. a calendar year), suitable temporal resolution for demand analysis (e.g. 1−5 minutes), and over a large enough sample for more general conclusions to be made is a common dilemma. Datasets with high temporal precision (and detailed appliance data) are often from a lower number of homes (Murray, Stankovic, & Stankovic, 2017). Such data can still be immensely valuable, but these limitations must be noted to understand the application of the findings of that data.

2.1.2. Energy use in Indian buildings

In some respects, for India as well as the UK, there is a general picture of higher resolution data becoming more available to describe energy use in the built environment, as noted in some of the below studies. Whilst there may be optimism that such a trend continues, it is also true that some of the risks that we may try to quantify for electricity networks (increasing peak demands, changing load factors, etc) require particularly high resolution data on a scale that might not, yet, be available.

Additionally for India, the scale of the country, and the difficulty in getting the same quality of datasets, does present different challenges for an energy modeller. In particular, the difficulty in connecting characteristics of individual dwelling demand profiles with trends observed at a regional (or even national) scale is more obvious. Even data relating to number of homes must be placed in specific context, with the...
definition of a home more complex than in other countries. Figures often refer to census data that is several years old, which is particularly a problem for a rapidly changing country. 2001 census data (Ministry of Home Affairs, 2018) reports of 187 million homes across India, 5 % of which are described as “dilapidated”. 58 million of these homes are categorised as “semi-permanent” and 35 million as “temporary”. This, clearly, should impact our approach to categorising future energy demand in India, where such categories may remain unrecorded within stock data or be brought “online”.

Meta-level studies in India have attempted to represent the degree of change projected for that country, further complicating the above picture. In 2012, India’s residential electricity consumption was 186 TW h/yr, up from 80 TW h in 2000 (Shukla, Rawal, & Shnapp, 2015). This vast increase, in a country where that growth is expected to continue (with 40 billion m² of new buildings projected by 2050 (Yu et al., 2017)), has resulted in households contributing 23 % of the electrical consumption of India (Ministry of Statistics and Programme Implementation, 2016).

Disaggregating that electrical consumption into household type, and improving spatial and temporal resolution, is limited by a lack of data (GBPN, 2014). Furthermore, due to diversity of households, climate and construction, relatively large samples of such data would be required to adequately capture current demand characteristics, where energy is being used/wasted, and demand reduction strategies that may be appropriate for those homes. Some research has attempted to bridge this gap between bottom-up and top-down energy modelling of buildings (Yu et al., 2017). This is of value in understanding regional variations in building stock and associated energy use, with clear policy impact, though tends not to focus on the more detailed demand characteristics that are required for an analysis of the risks facing energy networks in the future.

Energy efficiency in buildings is an area of concern in India, as evidenced by the introduction of the Energy Conservation Building Code (ECBC) (IMFR, 2015). It is notable, when comparing with European Union countries in particular, the goal of such policy is to reduce gradients of increased energy consumption, not to reverse it. This is due to the aforementioned increase in building completion, but also a projected population increase (from 1.3bn in 2015 to 1.5bn in 2030 (United Nations, 2019)) and the impact of climate change on a cooling-dominated building stock. There are, however, some signs that green building legislation, and the market associated with that (for both assets and technology), is maturing; in the beginning of the 21st century, green buildings were said to cost 18 % more than traditional buildings. In 2013, the difference was quoted as only 5 % (Smith, 2015).

Historical and projected future change therefore creates an uncertain picture from which a robust and low-carbon energy system, taking many years to plan and develop, must be designed. This creates a need for flexible methodologies to be developed that can assist the co-evolution between those technologies and buildings creating energy demand, and those systems aiming to serve that demand. The challenge, and potential feasibility, to do this at national level in India has already been noted, but focussing instead on discrete communities (as proposed within the CEDRI project) may allow an understanding of electricity demand patterns to be formed that has a wider application.

2.2. Tools and techniques for community energy modelling

Techniques for estimating building energy (thermal and electrical) performance are varied, and have been reviewed elsewhere for both transient and steady-state estimations (Jenkins, 2018; Sousa, Jones, Mirzaei, & Robinson, 2017), but approaches can be grouped as: empirical (using data, such as that in Section 2.1); semi-empirical (through use of statistical modelling of samples of empirical data); or purely theoretical (such as the use of thermo-physical models of buildings). Regional or country-wide building energy consumption is often modelled through stock modelling approaches (Hughes, Armitage, Palmer, & Stone, 2012) to provide some connection to policies aiming to promote, for example, energy efficiency measures within a standardised energy accreditation procedure (such as the role of Energy Performance Certificates (EPC) emanating from the European Union Energy Performance in Buildings Directive (EPBD) (European Commission, 2002)). This form of modelling also provides a basis for describing typology of buildings (in the form of archetypes) across a large area. However, defining specific household behaviour, or any aspect of energy use that has a strong temporal variation within diurnal scales, is more difficult. Such approaches are therefore often more successful with thermal demand (which have a strong correlation with physical variables of the building) rather than non-heating electrical demand.

Traditional dynamic building models are well-documented (Lomas et al., 1997) within the area of building design and related energy performance assessments. Although such calculations can require considerable quantity and detail of input, improvements in modelling efficiency (e.g. processing power of computers, interfaces of software etc) allow for this form of modelling to be used for groups of buildings. Much research has been conducted in this area, raising the potential of dynamic, temporally precise building modelling to be linked with the larger-scale energy pictures provided by energy system models (McCallum et al., 2019). This higher temporal resolution can be important as thermal demand becomes electrified, and the effect of aggregated heating controls across areas of built environment impact the electrical demand characteristics of those areas.

Energy system modelling (such as those stemming from the MARKAL/TIMES family of models) provides the potential to optimise across a range of energy demand and supply solutions, whilst attempting to achieve a future carbon target. The optimisation tends to occur over relatively long time periods, though work has been carried out to use higher-resolution supply and demand data with such models (Zeyringer, Daly, Fais, Sharp, & Strachan, 2014). One notable example of energy system modelling in practice is the Scottish Government Energy Strategy (Scottish Government, 2017) that used a version of the TIMES model to compare, amongst other aspects of energy policy, the effectiveness of heat decarbonisation against building energy efficiency. Again, whilst much of this work is often focussed on thermal energy demand (influenced by the importance of this within the UK built environment), projections suggesting an increase in electrification of heat and transport in the UK (United Kingdom Committee on Climate Change (UKCCC), 2019) will make the understanding of electricity demand evolution across regions/countries more important in the near future.

Better use of data can produce statistical models that are more likely to reflect actual electricity use patterns. Diary-based data studies of domestic consumption (Suomalainen et al., 2019) and time-of-use data (McKenna, Hofmann, Merkel, Fichtner, & Strachan, 2016; Torriti, 2017) can allow behavioural and occupancy patterns to be linked quite directly to demand profile characteristics, but rely on having significant qualitative and quantitative inputs from the household to make such correlations. This can also, due to the case-study specific nature of such data collection, limit the ability to extrapolate any findings to a wider sample.

2.3. Defining the need for further modelling

Reviewing the types of model described above, it is clear that aspects of each of these can be valuable when trying to estimate community electricity demands, and how to project these for given future scenarios. It would therefore be desirable to achieve a modelling framework that has the following properties:

- An empirical basis to account for sub-diurnal and seasonal variations that are difficult to model entirely theoretically.
- Within spatial/numerical limitations, provide energy demand characteristics that are scalable to a level that is useful for energy
suppliers and local networks (e.g. Low-Voltage network level, for which these aggregated building demands are so important).

- A modelling framework that has some indication of causal factors behind demand characteristics, such that future demand profiles can be generated at similar resolution based on defined scenarios of climate, occupancy, building typology, and technology.

- A modelling framework that exploits the availability of larger, higher resolution data sources that are becoming (within data protection protocols) more available for characterising energy demand.

A model that captures the above will allow those designing wider energy systems, and formulating policies for buildings within those systems, to reflect on how a superposition of changing parameters might impact on the performance of that system. The case-studies and model application of Sections 3 and 4 provide an initial indication of how that might work, and why it might be useful.

3. Case-studies of community energy use

Examining empirical data through real case-studies can be instructive in terms of understanding the requirement of what, for example, statistical models can do, and what constitutes a suitable application. As already discussed, the quantified findings of such applications (demand profiles, characteristics of energy use, etc) should not be generalised across large geographies; but the usefulness of that application, and validation process, can be demonstrated. With this objective in mind, two case-studies from very different locales are described below. Information of the studied dwellings is also summarised in Table 1.

3.1. UK case-study – Findhorn EcoVillage

Findhorn EcoVillage is a community located in Moray, in north east Scotland. It was established in 1972 and has seen various phases of expansion, including a number of coordinated housing developments in the last decade. There are around 160 dwellings in total under various tenure arrangements, with a significant proportion being owner-occupied. With over 500 residents, the community is one of the largest of its kind in the UK. In addition to the dwellings on site, there are a large number of guest houses, shops, an arts centre, town hall/theatre, community-run restaurant, a commercial printing press, and a number of offices and workshops.

Findhorn EcoVillage and its residents identify with a longstanding culture of understanding and protecting the natural environment, with these values underlying both lifestyle and construction practices (including sustainable food production, transport and energy/water/waste management). Aside from a number of older “Park home” dwellings, of relatively lightweight construction, the majority of residential buildings in Findhorn EcoVillage are built to very high thermal integrity, in excess of national standards. Along with three biomass-fed district heating systems, a mixture of air source heat pump (ASHP), LPG boiler and electric resistive heating can be found in the individual dwellings. Distributed renewables (solar photovoltaics and solar hot water) and thermal storage are included across a large number of sites. Further to this, there is a community-owned wind park (with three 225 kW turbines) and a private wire electricity grid. A number of the dwellings were monitored between 2014 and 2016 (including 43 listed in Table 1). Active/reactive power, voltage and frequency were also monitored at the village substation.

3.2. Indian case-study – Auroville

Auroville is an experimental township in Tamil-Nadu (established 1968) with a population of approximately 50,000 from different countries. Two sets of apartment buildings – “Citadines” and “Inpiration” – are being monitored as case studies within the CEDRI project, containing 34 and 14 flats respectively (of which 21 and 9 units were chosen to be metered).

23 occupants live in the selected 21 dwellings in Citadine. Most of these are single occupancy, though the majority have housemaids who work 4–8 h weekly. All the occupants work in the Auroville community. The monitored households use one or more type of lighting out of four categories (Incandescent, CFL, LED and T5 with Electronic Ballast) – suggesting a higher degree of diversity than a typical UK household. Electric fans and/or air conditioning (A/C) are commonly used in the dwellings. Other common household appliances are present (such as refrigerators) but, for cooking, Citadines has a community kitchen where most of the residents have their lunch. Again, from a demand characterisation perspective, the lack of a clear demand signal in the home representing times of cooking is somewhat unusual compared to UK equivalent data. Most of the dwellings use gas cylinders for cooking in general, though a small number have electric stoves. The dwellings have single phase meters for monitoring electricity consumption.

In the case of Inspiration, ten occupants live in the selected nine dwellings. The household size and working patterns are similar to occupant of Citadines, though differences exist in the use of lighting (incandescent, CFL and LED) and HVAC (electric fans in all selected households and A/C in one). Most of the dwellings use gas cylinders for cooking, as with Citadines, with one household using an induction stove. In the case of refrigerators, eight dwellings have one, and one household has two. The other appliances in use in Inspiration are similar as households in Citadines, except eight households use a geyser for hot water in the bathroom. The dwellings have three-phase meters for monitoring electricity consumption.

4. Characterising community electricity demand

Although working in different countries, and having somewhat different scopes, the work of the CEDRI and CESI projects have a number of common goals. Amongst these is the desire to overcome the boundaries between energy system and bottom-up energy demand models. As discussed elsewhere (Jenkins, Patidar, & Simpson, 2012; Jenkins, 2018), part of this challenge is being able to upscale detailed assessments of energy demand profiles of buildings to a level that is communicable, and useful, to those assessing the performance of an energy network (e.g. Low Voltage network or regional gas network).

<table>
<thead>
<tr>
<th>Case-study</th>
<th>Type</th>
<th>No. of dwellings</th>
<th>No. of occupants</th>
<th>Heating technology</th>
<th>Cooling technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auroville, Tamil-Nadu, India</td>
<td>Flat</td>
<td>21</td>
<td>23</td>
<td>None</td>
<td>Fans, A/C</td>
</tr>
<tr>
<td></td>
<td>Flat</td>
<td>9</td>
<td>10</td>
<td>Geyser</td>
<td></td>
</tr>
<tr>
<td>Findhorn, Moray, Scotland</td>
<td>Detached</td>
<td>21</td>
<td>63</td>
<td>LPG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Terraced</td>
<td>6</td>
<td>18</td>
<td>Biomass (district)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Terraced</td>
<td>5</td>
<td>15</td>
<td>Electric</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Terraced</td>
<td>9</td>
<td>27</td>
<td>ASHP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Detached</td>
<td>2</td>
<td>6</td>
<td>ASHP</td>
<td></td>
</tr>
</tbody>
</table>
though potentially at higher spatial scales than this). The authors have previously developed tools with this objective in mind for both thermal modelling (Jenkins, Patidar, & Simpson, 2015) and electrical demand modelling (Patidar, Jenkins, & Simpson, 2014). An updated approach of the latter will now be demonstrated for UK and Indian locations, thus indicating (within the wider issues discussed in Sections 3 and 4) the suitability of such tools for allowing building-level demand data to be extrapolated beyond just the individual building.

4.1. Decomposing electricity demand profiles for purposes of synthesis

A key part of pattern recognition of electricity demand is the distinction between signals that occur at definable time functions and those that are broadly stochastic and, as a percentage of a total day, low frequency. Modelling the latter is particularly important to account for diversity in aggregated demand profiles; in essence, households doing similar things at slightly different times.

The synthesis approach taken by the authors is to adopt a Seasonal Trend decomposition procedure based on Loess (STL) (Cleveland, Cleveland, McRae, & Terpenning, 1990). This allows for composite signals within a high resolution demand profile to be recognised. As well as raising the prospect that these different signatures can be modelled separately (with different sensitivities to external parameters, such as weather), it also allows for a stochastic component to be modelled discretely, with a view to capture diversity through the aggregation process. The stochastic component itself is modelled through a Hidden-Markov Model Generalised Pareto (HMM-GP), also detailed elsewhere (Patidar, Allen, Haynes, & Haynes, 2018), that allows for “states” of demand at a certain time to be linked to previous values, with the probability of being in those states trained on real data from a given sample. The use of a Generalised Pareto approach allows for statistically extreme values (i.e. those representing high power “spikes” in a typical high resolution demand profile) to be better characterised. The process can therefore be summarised as:

**Step 1:** Individual dwelling Demand data series is transformed into an additive series (from multiplicative series) using a logarithmic transformation. An STL decomposition algorithm is applied to decompose the electricity demand series into three components for each individual dwelling: i) Trend, ii) Seasonal and iii) Residual (stochastic) variations.

**Step 2:** A HMM model comprising of five elements is fitted to the residual component. The five elements of HMM model include defining i) a set of observed states using a percentile analysis of the observed values, ii) a set of unobserved (hidden) states, iii) a state transitional probability matrix of observed states, iv) an emission probability matrix of hidden states, and v) an initial probability matrix of observed states.

**Step 3:** The Residual component is simulated using a HMM model fitted to the observed series of Step 2.

**Step 4:** Synthetic electricity demand profiles are constructed by combining simulated residual components with the trend and seasonal components of the observed series.

**Step 5:** A GP distribution is fitted to the extreme values (i.e. over 95th percentile) of the observed electricity demand profiles. Extreme values in the simulated synthetic demand series are resampled from the fitted GP distribution to facilitate better estimation of peak electricity demand values.

**Step 6:** A fine percentile-based bias correction is applied to process and correct any potential bias introduced by logarithmic transformation of the original series and application of inverse function to achieve the final profiles.

The advantage of this approach is that it produces individual dwelling profiles that, when generated repeatedly, produce different stochastic components that are therefore suitable for aggregating (due to inherent diversity). One of the weaknesses is the reliance on, potentially, a relatively small sample and this can restrict the appropriate scale of extrapolation. This is explored further below. The tool will now be applied to the two case-studies under investigation, noting the comparison with empirical data (at individual and community level) and the differing performance of the approach in both locations.

4.2. Statistical modelling of electrical demand data (UK)

4.2.1. Comparison of individual dwelling data

Fig. 1 is an example of a measured 24-h electrical demand profile in a home in Findhorn, placed alongside a synthesised version of that profile from the discussed demand profile algorithm. A visual inspection of Fig. 1(a) can allow for inference of typical...
activities that are present in many UK homes, including short “spikes” of demand from kettles, toasters, and electric showers, and oven usage (comprising a cycling heating elements of, typically, 2−3 kW). Some of these features have strong relationships to time of day, whilst others occur more stochastically. Other underlying features at lower power consumption can be seen that occur throughout the day (e.g. refrigeration cycles) or for long periods of the day (e.g. lighting/consumer electronics during occupied hours).

The procedure documented in Section 4.1 allows for a synthesis of these patterns of demand as seen in the sample. Fig. 1(b) provides an example of a synthesised demand profile produced from this technique, with a breakdown of this decomposition (required for synthesis) shown in Fig. 2. The advantage of Fig. 2, for this particular study, is the ability to quantify differences in characteristics across different sample selections. As noted in Section 4.3.1, the variation of demand at defined periodicities is fundamentally different across the UK and Indian samples, and this can be related to tangible differences in how energy is being used.

Although visualising typical 24-hr profiles is of value to demonstrate the type of signals being generated, this does not provide an adequate validation for the success of this synthesis. Also, the synthesis is not designed to exactly replicate all data points – this would be contrary to the need for modelling diversity (e.g. if a real household has switched a kettle on at 9.03am, a synthesised profile would not necessarily model the same feature at exactly the same time, though it may reproduce this feature at a similar, but non-identical, time). Therefore, it is to be expected (and, actually, desired) that the synthesised and observed profiles in Fig. 1 do not exactly match every minute. They should, however, be returning similar characteristics of demand at similar times. With this in mind, Fig. 3 creates a more robust validation by carrying out a percentile analysis of modelled demand values, compared to those measured.

For a single observed dwelling, 20 synthetic dwellings are generated – where the stochastic element of the synthesis will produce a slightly different percentile distribution for each synthesised dwelling. Across percentiles the match between the synthetic demand generator and that observed is reasonable. The Generalised Pareto method (described in Section 4.1) proves to be an effective function to ensure very high percentile values (i.e. 95 % and above) are well matched. There appears to be slight discrepancies between 80−95 % percentiles; this may be due to residential demand itself being particularly diverse for such ranges, where demand values of ∼1−2 kW potentially being the result of a wide range of residential appliances, and therefore more uncertain in terms of when they might be visible on a real demand profile.

4.2.2. Comparison of aggregated community electricity usage

The analysis of 4.2.1 is instructive in terms of the mechanism of the synthesis and the performance at an individual dwelling level. However, a key function of the process is to produce aggregated demand profiles from a bottom-up process that is, in turn, linked to some understanding of individual demand profiles. Therefore, validating the final aggregated profiles against an empirical measurement of community electricity demand is of value.

For this study, data will be used from substations that are known to be serving the studied regions. The aim is to replicate, in the synthesised aggregation, the features of a substation profile that result from residential electricity demand. This can be challenging when a substation is known to serve significant non-residential demands (non-residential buildings, industry, electricity used by other services within...
4.3. Statistical modelling of electrical demand data (India)

A similar exercise is now carried out for the Auroville dataset. The data has been recorded slightly differently to Findhorn (using pulse meters that record information every 0.06Wh for single-phase properties and every 1.25Wh for three-phase), but the result is still a high resolution profile which is appropriate for the STL analysis referred to in Section 4.1.

4.3.1. Comparison of individual dwelling data

Fig. 5 shows an example of a real 24-h profile taken from an Auroville dataset, and compared alongside a synthesised equivalent.

Again, the full characteristics of demand of a household cannot be inferred in totality from a single day but, aesthetically, the Auroville profile has somewhat different features to Findhorn. The cycling refrigeration profile is more significant (perhaps driven by higher ambient temperature), there are fewer features that are associated with lighting/consumer electronics (suggesting reduced prevalence of such technologies), and the low frequency, high power events occur at different times. Other known issues about these households (e.g. relatively little electrical cooking, some communal washing services) are also evident in the profile.

In theory, the difference in physical activities and technologies in Fig. 5(a) (compared to Fig. 1(a)) should not impact the performance of the STL analysis, which just requires a high resolution profile that has discernible signals occurring at different time periods. As with Fig. 1, Fig. 5 provides only an indication of the differences in synthesised and observed demand profiles for an individual dwelling. As noted below for Fig. 7, the observed Indian demand profiles have a different distribution of demand values over a given range. In this case, for example, the fact that the synthesis of Fig. 5(b) is less capable of representing a refrigeration profile becomes more visible (and potentially more important) than for a UK home. Fig. 6 shows the results of the STL decomposition used to generate this synthesis, over a period of two months.

Comparing Fig. 6 with Fig. 2, there is a stronger signal discerned in the Seasonal profile and a lower degree of stochasticity in the Random profile. The former may be the result of a dominance of a smaller number of appliances in the home (a possibility which will be explored through household interviews at a later stage of the CEDRI project), which will produce a pattern repeating at a more regular periodicity. The notable trend in the Random profile may be due to the relatively short period over which the data has been collected – with 12 months of data, the “Trend” profile may identify a stronger link with long-duration variability (e.g. temperature changes throughout the year), and thus the technique will “remove” such patterns from the Random profile. However, the applied algorithm shows versatility in being able to identify quite different variations over a range of periodicities in these two different locations.

It is instructive to compare Fig. 7 with Fig. 4. Firstly, comparing the

the community etc.), or is impacted by distributed electricity generation within that community (though that can be modelled discretely if information is available).

Fig. 4 compares measured substation data from Findhorn with an aggregated profile that is modelled from individually synthesised dwellings over the same time period. With the substation known to serve 181 dwellings, the synthesis is applied for the same number of dwellings. The building stock in Findhorn fits four construction classes: Findhorn Construction (FC), FC+, Parkhome (PC) and Timber Construction (TC). In addition, the sample has different floor areas, heating technologies and solar photovoltaic usage. This breakdown was used to assign weightings through the aggregation process, to ensure the aggregated profile had a suitable representation of these different demand profiles.

Further validation of this data (at substation level) is being carried out as part of a separate task within the CESI project – with longer duration data becoming available. However, the synthetic aggregation appears to be reflecting a general diurnal trend that is seen in the real data, but doing so with more noticeable “noise” signals superimposed on top of that diurnal cycle. This may be due to the relatively small sample size of empirical data, and future work will test the performance of the synthetic aggregation with larger, and more diverse, sample sizes. Also, although this is primarily a residential community, the transformer will be serving non-residential loads which are not accounted for in the demand synthesis. Therefore, as also discussed below, substation data provides a somewhat imperfect validation for this data synthesis model.
observed demand in both graphs, it is clear that the two locations exhibit a different demand distribution (though this can also be confirmed by looking at Figs. 1 and 5). Findhorn has significant features in the 1–3 kW range (as demonstrated by the 80–95% range of values in Fig. 4) whereas Auroville does not. It might be expected that a HMM-GP model would perform differently for these two different locations, where HMM “states” are calibrated based on discrete percentile ranges relative to the maximum value. One reason for the relative similarity in model performance (in Figs. 4 and 7) is due to the bias correction step, as documented in Section 4.1. This ensures that, otherwise, outlying values within a percentile band are corrected by the empirical dataset to produce more characteristic values of that location. As well as being important for transferring the model across to different regions, this function is of potential value for looking at future demand profiles of a region – where the distribution of demand is likely to become quite different for profiles exhibiting electrified heat and transport.

4.3.2. Comparison of aggregated community electricity usage

As with Section 4.2.2, the aggregation of the synthesised demand profiles will now be compared to a local substation. The same limitations are present here in terms of non-residential loads, but an illustration of the performance of the synthesis can still be gained, as shown in Fig. 8.

It is evident that certain features in the substation data are not being characterised adequately by the synthesis, which has potential implications on sample selection and application of the synthesis to particular demand profiles. Firstly, there are clear outages in the substation data that have (intentionally) not been characterised by the synthesis and one must be aware of these when carrying out this type of analysis in parts of the developing world. Of more importance to the function

![Fig. 5. (a) Observed and (b) synthesised Auroville individual dwelling electricity demand profile.](image1)

![Fig. 6. Decomposed Auroville individual dwelling electricity demand.](image2)
(and sampling) of the model, of the 21 sample dwellings being used for Auroville only one exhibited a clear air-conditioning profile. When attempting to upscale this sample to a larger number of dwellings (specifically, the 108 being served by the above substation), the aggregated cooling profile is likely to be poorly specified. Even allowing for weighting of the dataset (i.e. increasing the weighting of the mechanically-cooled dwelling within the sample), the aggregation will not have captured real diversity in cooling usage across a wide enough sample. Also, for Auroville, there are a number of energy uses that might be seen as non-typical to other communities (both inside and outside India). This includes communal cooking and washing within that community, effectively removing those types of energy use from the individual demand profile. Clearly, if carrying out a demand synthesis based only on the characteristics found in residential demand profiles, there is likely to be a mismatch when comparing to substation data if that substation serves significant demands that are not due to residential buildings. This may suggest the need to change the type of validation analysis, such as comparing synthesised aggregation with an empirical aggregation (i.e. summated individual dwelling profiles) rather than the raw substation data. These factors, and other lessons, are noted in Sections 5 and 6.

5. Summarising differences and synergies

Comparing two different locations allows for an investigation into the appropriateness of some of the generalisations that might be applied for an individual location (e.g. a heating dominated country like the UK that has relatively little electrical heating). Based on the case-study analyses and extended review, the below areas of interest are identified as being more generalisable to community energy analysis of electricity.

5.1. Data availability

The statistical model proposed here requires data at appropriate temporal and spatial resolution, with key characteristics of residential demand difficult to distinguish above 5 min resolution (for individual dwellings). However, this is not just a limitation of the proposed model but a fundamental consequence of the types of electrical devices in the home. To characterise such features purely from theoretical modelling is very difficult, but the greater access to data in recent years does open up the possibility of more robust empirical models, populated and informed by large datasets.

The role of, more theoretical, physical models can still be of importance if/when heat is electrified (see Section 5.3) but the modelling of partly stochastic, high resolution signatures from real demand data benefits from an empirical starting point.

Both countries noted here have such data available but within certain constraints; in particular, as more variables are investigated with a greater level of data resolution, there is a greater need for significantly sized samples of households from which such correlations can be investigated and, potentially, generalised. An upscaling to larger regions (and nationally) is not investigated here but is a central challenge to any desire to take bottom-up modelling to a greater level of extrapolation.

![Graph](image-url)

**Fig. 7.** Percentile analysis of model performance for individual dwellings in Auroville.

**Fig. 8.** Substation data compared to aggregated synthesis of electricity demand in Auroville for example day.
5.2. Services required and related technologies

The two countries investigated here have significantly different energy service requirements, and it should not be a given that the same methodology is suitable for both. For many UK electricity demand profiles, heating will not be present (due to the prevailing choice of gas boiler heating from the mains gas grid), apart from relatively small usage from a boiler pump. Where electrical forms of heating are in use (e.g. rural areas or when investigating future electrification of heat), quite different profiles will be observed (see also Section 5.3).

Furthermore, whereas a UK home might have centrally-controlled heating (i.e. heating control responding to a whole-house average, or single thermostat), an Indian home will not necessarily be centrally cooled, so assumptions about control of temperature (and the demand profile emanating from this) are not necessarily transferrable between the two countries. Even if comparing Indian demand profiles with UK profiles that have electrical heating, the methods for achieving thermal comfort are quite different. Households may: i) not have any mechanical cooling and rely on passive measures, ii) not have mechanical cooling but rely on mechanical ventilation, or iii) have both mechanical cooling and ventilation.

The link between options for summertime comfort and household income, and the assumed change in disposable incomes for many Indian households in the coming decades, makes this all the more important, and difficult, to characterise. However, by distinguishing between current demand profiles that exhibit different forms of comfort provision (as might be identified through an STL-type decomposition), it is possible to construct a future aggregation of demand profiles that reflect different future scenarios. This is currently being developed by the CEDRI project.

5.3. Models of building stock

As previously mentioned, when modelling electricity demand, the relative importance of stock modelling (or any form of physical modelling) will depend on the presence of electrical heating or cooling in that stock. Again, with discussions about data availability in Section 5.1, constructing a representative sample size is challenging – in this case, in terms of building “archetypes”. For both countries, being able to model thermal demand (to be considered here as relating to both cooling and heating) across a large number of buildings but whilst maintaining a suitable temporal resolution to those demand profiles will often require an upper spatial limit to that modelling. Specifically, whilst such an approach has been applied here for communities of ~100 s of dwellings, imagining the same level of bottom-up detail for something approaching national level is more difficult.

If, in the UK, electricity demand profiles become more correlated with physical building parameters (due to electrification of heat), a community-scale stock model of heat will be of value, and this is being developed by the authors elsewhere (McCallum et al., 2019). For India, if wanting an analogous exercise around cooling, a further challenge exists due to the poorer characterisation of the building stock, and prevalence of informal settlements. There is also the question of whether building quality or household socio-demographic is the driver for cooling in Indian communities – and, like many of these questions for India, whether such a relationship would be in any way generalisable across a region or country due to the previously discussed cultural and climatic diversity. A conclusion of this might be that a stronger link is required between empirical/statistical models of energy demand, physical models of building archetypes, and the energy behaviour/practices that are intrinsic to the empirical electricity data that is being collected and, for CEDRI, embedded into a modelling environment.

5.4. Homogeneity of use

This study has attempted to constrain the multi-building analysis to relatively small communities (e.g. as might be served by a substation), rather than extrapolating too widely. However, any assumptions of homogeneity of electricity use for even small samples of households should be applied with caution – and, for this reason, the validation exercises proposed here are important. Electrification of heat, should this reach a high level of penetration, would potentially reduce homogeneity in electricity demand even further, with building parameters playing a greater role in the shape and characteristics of those demand profiles. As discussed above, the role of thermal models in such a future would be of increased importance.

For India, a national-level demand model that has strong bottom-up components (of the type described here) is unlikely due to the scale and diversity of the country. However, with data becoming more available in that country that can be used to quantify that diversity (and potentially explain its causes), an empirical, replicable methodology of demand characterisation is still of value – but within clearly defined spatial limits. As noted in Section 4.3.2, the task of choosing a sample that is then weighted to reflect wider heterogeneity in, for example, cooling demand is a non-trivial task.

5.5. Climate

The relationship between climate and whether residential demand is heating or cooling dominated as already been touched upon. However, the importance of future climate should also be noted. The quality of future climate projections, and accessible formats of these projections for demand modelling, differs with country. The work of the UK Climate Projections group (Met Office, 2019) has provided probabilistic future climate projections in the UK for some years, with various projects (ARCC, 2019) interpreting this for use with building modelling. These options are, in part, a legacy of the climate modelling (such as the Hadley model) led by that country over many years.

At present, India does not have the same level of current weather data or future climate modelling available. With the importance of cooling in Indian residential electricity demand, and the growth expected in the penetration of residential cooling into the market, a higher resolution of climate projection would be of benefit to the modelling of future electricity demand in India.

6. Discussion and conclusions

As part of a more general need to model the energy use of communities at higher resolution, an approach has been proposed that uses empirical electricity data to characterise, and synthesise, electricity demand in a scalable way (noting that the limits of this scalability require further investigation). The application of this method to two, very different, communities has allowed for a wider discussion around modelling requirements and applications.

Amongst other issues, the study has demonstrated challenges relating to lack of diversity present in a small empirical sample when attempting to upscale those characteristics. In particular, the representativeness of that sample in describing both the control and existence of heating/cooling technologies requires further research – and this is likely to include both a more nuanced approach to sample selection and the re-weighting of the generated profiles to represent a wider area of energy demand. Other factors include the impact of disruptions to the power supply (highly variable between different countries) and the relative importance of non-residential loads, which differ significantly between communities. This study, by framing these challenges more completely, will allow for further improvements in the modelling approach as the named projects progress.

The future evolution of this process will therefore require more data, and further investigation into the upper limits of scalability – itself linked to the data (and sample size) populating those models. Moving forward, the CEDRI project will develop the modelling framework to incorporate discrete thermal modelling (including other energy vectors).
to provide a more complete picture of community energy. Furthermore, empirical correlations with input parameters (weather, occupancy, etc) will be determined such that modelled demand profiles can be morphed for different futures. This will then be able to inform design of community energy networks, reflecting the need to manage infrastructure and (potentially) demand response strategies to run these networks more efficiently.

Acknowledgements

The CEDRI project (Grant reference EP/R008655/1) is funded by the Engineering and Physical Sciences Research Council (EPSRC) and Indian Department of Science and Technology as part of the Newton-Bhabha programme. The CESI project (EP/P001173/1) is also funded by EPSRC. The support of these funders, and related organisations, is greatly appreciated.

References

Auroville (2018). Auroville organisation. Available at: https://www.auroville.org/


D.P. Jenkins, et al.