IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 9, 2020, accepted October 14, 2020, date of publication October 19, 2020, date of current version November 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032228

Predictive Thermal Relation Model for Synthesizing Low Carbon Heating Load Profiles on Distribution Networks

AMY ANDERSON, BRUCE STEPHEN, (Senior Member, IEEE), RORY TELFORD, AND STEPHEN MCARTHUR, (Fellow, IEEE)

Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1RD, U.K.

Corresponding author: Amy Anderson (amy.anderson@strath.ac.uk)

This work was supported by the EPSRC Centre for Doctoral Training in Future Power Networks and Smart Grids (EP/L015471/1), in partnership with SSE.

ABSTRACT The introduction of electric heat pumps as a low carbon option for space heating offers a potential pathway for reducing the carbon emissions resulting from domestic heating demand in the UK. However, the additional power demands of heat pumps over conventional domestic loads have the potential to significantly erode network headroom, particularly at the distribution level. The uptake of this technology within the UK is currently limited and the effects of widespread adoption on distribution networks are not well characterized due to the sparse availability of operational heat pump demand data. This paper outlines a methodology for quantifying the demand impact of heat pumps on Low Voltage networks sensitive to local temperature by deriving fundamental thermal relationships from real heat pump electrical demand data. These relations can then be applied to predict demand for new studies independent of the geographic specifics of the original dataset. The strength of this model is in the ability to predict an aggregated hourly heat pump electrical demand profile that reflects local temperature conditions and intra-day usage as well as population size, thereby also accounting for diversity effects that are difficult to capture in physics based models. This work augments the usability of limited existing data by facilitating demand analysis sensitive to local temperature conditions, rather than blanket rescaling of existing customer data as has been performed in previous studies. This creates future opportunities for examining heat pump demand sensitivity for different geographic locations against existing heat pump assessments, as well as performing studies which incorporate multiple low carbon technologies connected to a Low Voltage network.

INDEX TERMS Load modelling, heat pumps, data analysis.

I. INTRODUCTION

Reducing the contribution of heat to the UK's greenhouse gas emissions presents one of the largest challenges in achieving long-term emissions targets set by government policy. The contribution of domestic heating is estimated to average a third of household emissions [1]. In order to achieve 2050 net-zero goals this must be reduced by a further 95% from 2017 levels [1]. Decarbonisation of the UK's heating sector will require a radical shift in the current status quo, expected to necessitate widespread adoption of low carbon heating with improved efficiency measures.

Electric Heat Pumps (EHP) offer one potential low carbon alternative, reducing CO2 emissions of up to 25% per unit of heat generated [2]. In combination with a fully renewable electricity source, this can reduce the effective household heating CO2 emissions to zero. Advantages include acting as a low-regret option for off-gas grid households [3], and as a low-cost option for newer well-insulated builds.

The growth of heat pump technology has been supported by UK government policy [4] and industry trials [5] but despite this, overall deployment remains low – 76,388 domestic heat pumps were registered with the RHI (Renewable Heat Incentive) scheme in the UK by the end of 2019 [6]. In contrast, the UK advisory body the Committee for Climate Change (CCC) has recommended the installation of one million heat pumps annually by 2030 in order to meet decarbonisation targets [3]. This level of growth presents a major challenge for distribution network operators as heat

The associate editor coordinating the review of this manuscript and approving it for publication was Hui Ma.

pump loads at maximum output are significant both in terms of energy and power compared to existing domestic Low Voltage (LV) network loads.

A. THE CHALLENGE AROUND HEAT PUMP DEMAND MODELLING

The increase in network load contributed by heat pumps presents a serious threat to the existing thermal and voltage limits of LV network assets. In order to maintain quality of service for customers whilst minimising the need to incur costly network reinforcement, network operators require a clear view of how heat pumps will impact networks at a local level.

For this purpose, predictive demand models that use operational demand data have an advantage over conventional physical demand models, in that they can capture individual household behavioral and diversity effects that are difficult to parameterize in physical models. However, there is currently very limited availability of UK-based operational heat pump data to draw on for examining heat pump network effects. Furthermore, heat pump demand is highly sensitive to local temperature conditions; conditions can be highly divergent even within a limited geographical area due to factors such as local topography and level of urbanization. This necessitates a model that captures the full temperature to demand relationship rather than relying on operational maximums.

Traditionally, load prediction at LV level has been limited to modelling peak annual demand and rating physical assets appropriately in order to ensure there is sufficient headroom to meet this modelled peak [7]. The need to decarbonize the energy sector necessitates the further uptake of LV-connected low carbon technologies such as heat pumps, alongside wind, PV and EVs. The inherent stochasticity of these load types with potential for new failure modes demands new prediction approaches beyond the historic method of modelling bulk aggregations.

On this basis, the contribution of this paper is a methodology for quantifying the impact of heat pump demand on LV networks sensitive to local weather conditions by deriving fundamental thermal relationships from real heat pump electrical demand data. These relations can then be applied to predict heat pump electrical demand for new studies independent of the original dataset, thereby maximising the utility of sparsely available heat pump demand data.

This methodology extracts electrical demand versus hour of day and electrical demand versus temperature relationships from individual customers in an operational dataset and translates these relationships into a format that can be then used to probabilistically predict heat pump demand through a black-box type approach. Linear scaling factors are derived from the operational dataset to act as a proxy for the variation in building type, heat pump type and system efficiency that would be seen in a typical UK population. By directly modelling these relationships versus electrical demand, the need to transform a heating demand into an electrical demand is circumvented.

The majority of previous heat pump demand impact studies have focused on predicting demand for extreme cold temperatures rather than fully capturing the temperature/demand relationship [8], [9]. A key strength of this model is the ability to generate a heat pump demand profile that is sensitive to local temperature conditions, hour of day and population size, thereby accounting for diversity effects as well as temperature. Furthermore, by incorporating the full electrical demand versus temperature and time of day relationship, this model facilitates the study of heat pump demand impact alongside other low carbon technologies on an LV network for conditions other than extreme cold days. The mean error and standard deviation of this model are tested versus two heat pump demand datasets, with consistent results versus population size and temperature for both cases. This indicates that the developed approach will be generally applicable for UK based heat pump populations, facilitating analysis of LV networks with demand profiles tailored to local weather conditions.

Section II of this paper describes the existing approaches for modelling heat pump demand in the literature, outlining their benefits and limitations. In section III of this paper, the selected methodology for modelling heat pump demand is described. Section IV presents the validation results as well as primary model results. Section V discusses the results, section VI outlines potential applications and a basic case study and section VII concludes the paper with further possible work.

II. EXISTING HEAT PUMP DEMAND MODELLING APPROACHES

The primary challenge when evaluating the impact of heat pumps on a distribution network is accurately quantifying the magnitude of additional electrical load contributed by the connection of heat pumps. The current low uptake of heat pump technology in the UK results in a general lack of operational demand data that could be used to facilitate heat pump effects analysis and general evaluation of network impacts.

Excessive additional load will result in a significant impact on voltage and reduction in thermal headroom on a network, potentially resulting in a breach of operational limits. The difficulty surrounding heat pumps is that while their heat output is broadly proportional to ambient temperature on a seasonal time frame, at a daily and hourly level the demand profile for a single customer is determined by a broader range of factors. The electrical demand required to meet a target heating demand for a household is influenced by several parameters including building type, heat pump type and building efficiency as well as the behavioural patterns of the individual household. Therefore for a single point in time, the additional electrical load presented to a distribution network due to heat pumps is a function of parameters specific to each household in addition to the common local temperature conditions and time of day. This contrasts significantly with conventional domestic loads on distribution networks which are highly static and predictable in nature.

The instantaneous electrical demand of a typically sized domestic heat pump can be equivalent or in excess of current daily domestic demand peaks [10]. In terms of energy, the average heat pump electricity consumption of 8kWh per day [8] is roughly equivalent to the existing average electricity consumption for a UK household of 8.5kWh per day [11]. Each additional heat pump connected to an LV network is roughly equivalent to the connection of an extra household and therefore is a serious consideration for network headroom at higher levels of penetration.

Analysis and prediction of heat pump electrical demand modelling within the UK is currently constrained in the literature primarily to either small-scale physical models which require a high-level of specific system knowledge to make predictions [12], or methods which rescale existing heat pump trial data to achieve a deterministic outcome [9].

There are currently three fundamental approaches for modelling heat pump load profiles that have been used in the literature:

- Physical model that captures a detailed heat pump/heating system but with limited capture of time of use effects across a population [12], [13]
- Use of existing gas or heating demand data, making assumptions about building type, building insulation and population information, [14], [15]
- Use of electric heat pump trial data; examine and rescale for time periods of interest [8], [9], [16]

These approaches all feature their own specific advantages and disadvantages depending on the specific area of study.

Existing physical approaches are well-suited for simulating highly defined models that clearly characterise one heating system; this makes them ideal for modelling highly-specific behaviours such as fast start-up transients. Underwood *et al.* [12] developed a compressor-based parametric model for capturing seasonal performance of different manufacturer's heat pumps, which was able to achieve good results when comparing actual and modelled heat output. Other works further incorporate building parameters across a population when considering heat pump demand [13], but the approach lacks real demand data to support the full validation of results and it is therefore not possible to quantify the associated error. Heat pump electrical load and therefore its immediate impact on an electrical network is a function of several parameters that will vary from household to household; these include relatively fixed characteristics such heat pump type, building and insulation characteristics but are also strongly linked to ambient temperature conditions and behavioural routines which will vary seasonally. Furthermore it can be expected there will be diversity in heat pump type, building characteristics and behavioural routines even within a local neighbourhood [17]. Fig.1 illustrates sample daily load profiles for nine different households on the same winter's day from the Renewable Heat Premium Payment (RHPP) dataset [18]. All households are based in England and therefore are exposed to similar daily temperature profiles and magnitude. Each customer load profile clearly features a distinctive shape and there is limited commonality from customer to customer. This combination of physical, seasonal and behavioural characteristics makes it very challenging to develop a fully representative physical heat pump model that can translate these population-variable parameters into an aggregated load profile that accurately reflects the energy, power and time of use characteristics of a real heat pump load.

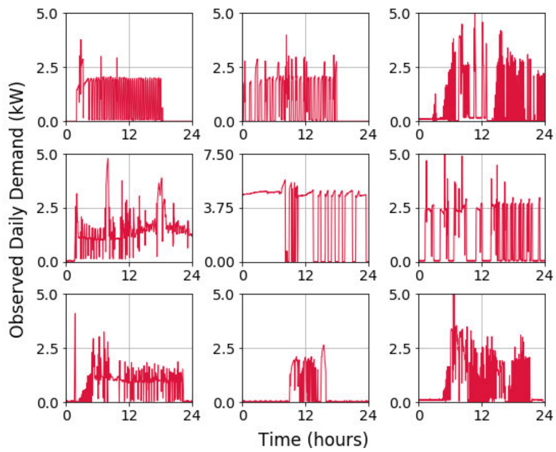

FIGURE 1. Intra-day heat pump electrical demand for nine sample customers on same winter weekday.

A more straightforward approach to modelling heat pump demand is to take existing gas or heating demand data and rescale this the equivalent electric heat pump demand based on an appropriate Coefficient of Performance (COP) figure. This method has been applied for showing large scale effects for the transition to greater levels of heat electrification in the UK [14], whereas elsewhere electrical demand profiles have been derived from ambient temperature and heating demand profiles with an hourly resolution [15]. Whilst strongly linked to the true heating demand characteristics, this method has limitations when applied to LV networks on daily or hourly resolutions. Current gas demand magnitude is partially shaped by equipment type (i.e. combi versus condensing boiler) as well as home characteristics and behaviour. Alteration of the heating system will reshape heating demand according to EHP characteristics. This may potentially also alter pre-existing behavioural thermal routines, such as when occupants choose to enable household heating [17]. Due to lower flow temperatures than conventional boiler based systems the time of use characteristics of heat pumps can be anticipated to be different compared to existing profiles and will potentially be spread more widely across the day.

In contrast to physical models and heating demand based approaches, methods that utilise existing EHP demand data from trials are able to mitigate the requirement to fully characterise the heating system in order to define the electric demand. The primary restriction with this approach is that due to the limited number of heat pumps active within the UK there is sparse operational data from which to draw conclusions about heat pump network impact. Consequently, the level of detailed heat pump analysis involving large populations is limited and generally features linear rescaling or averaging of the aggregated profiles in order to assess heat pump demand magnitude for a given time window.

Models based on real heat pump demand data have been developed in the literature, circumventing the need to fully characterise a physical heating system. A high-resolution probabilistic model drawing on operational data from 72 micro-CHP (combined heat and power) units during field trials in 2011 was developed for electrical heat pump demand prediction [16]. The probabilistic approach of this study enabled the definition of a range of possible demand values with respect to heat pump penetration. However, this study is reliant on the fact that micro-CHP technology represents a good approximation of EHP demand patterns, and does not draw on real EHP operational data.

Recent UK trials have greatly improved the availability of domestic demand EHP data [18], [19], however limited analysis has been performed to date. At present the majority of heat pump demand modelling studies only focus on averaged profiles at operational extremes. As the kinds of loads connected to LV networks become more diverse, with a mix of PV, EV, wind and low carbon heating technologies, there is a strong need for the capability to model realistic heat pump demand profiles alongside the interactions of other technologies.

The methodology described in this paper will define a composite approach between a fully physical demand model that requires detailed inputs and can be difficult to validate, and data-dependent approaches that primarily rescale existing demand data. The concept of synthetically generating demand profiles from real data has been used in other domains as a way of facilitating system analysis for applications where real data may be sparse or difficult to obtain [19], [20]. At present there has so far been limited use of these techniques for heat pump demand applications. Synthetically generated demand profiles derived from real operational data present an opportunity to develop a model that characterises the difficult to capture elements of a physically-defined heat pump model that can be validated against operational data. This maximises the value that can be drawn from limited real world data that is typically costly and practically difficult to obtain.

III. PROBABALISTIC PREDICTION OF LOCALISED HEAT PUMP DEMAND

This paper describes a method for quantifying the demand impact of increased heat pump uptake for population sizes typical of LV networks, sensitive to local ambient temperature. This is performed by extracting the fundamental relationships between heat pump electrical demand, temperature and time of day from a training dataset such that it can be applied to a new target application. A one hour-resolution

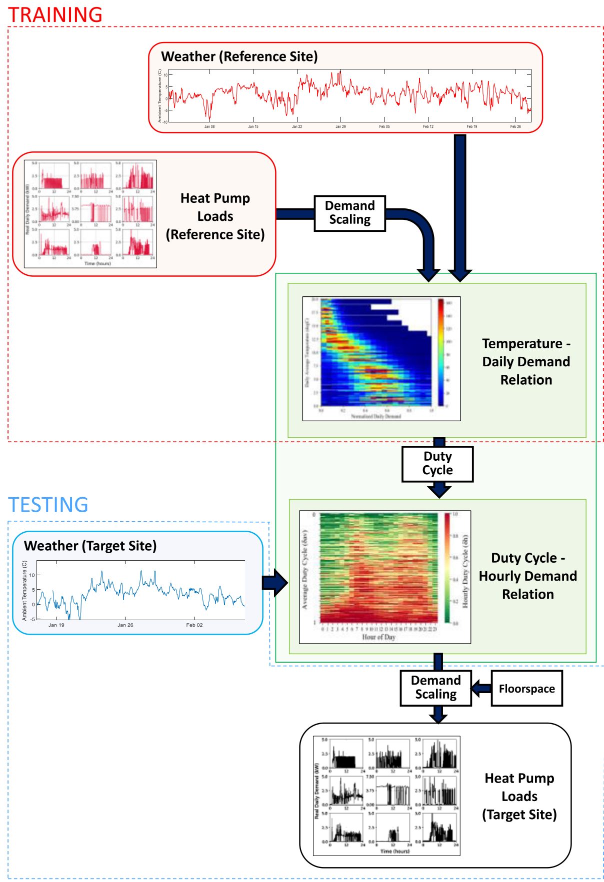

FIGURE 2. Localized Heat Pump Demand Model Overview showing in sample training and out of sample testing/prediction procedures.

is selected for the synthetic demand profiles as a trade-off between achievability and utility.

This model will utilise UK domestic heat pump data coupled with historical weather data in order to characterise the fundamental relationship between daily electrical demand and temperature for heat pumps. This will enable generation of heat pump demand profiles sensitive to local temperature using only the local temperature information and limited inputs to seed the heat pump sizes within the population. The three primary relationships versus ambient temperature this study will characterise are:

- Daily Energy; heat pump electrical demand over the course of a day
- Daily Average Duty Cycle; average heat pump state for a single day
- Hourly Duty Cycle; hourly heat pump state within a single day

Models capturing these relationships are used with local weather observations at a target site to produce a EHP demand for that site; this facilitates modelling hourly heat pump demand for cases independent of the original sample dataset. The interrelationships between model data, characterization and model tests for this work are defined in Fig 2. The datasets used for developing this model are described in more detail in section III.A. The model will be validated with multiple datasets, both to prove the concept but also to quantify the error associated with the model.

A. CASE STUDY DATASETS

This work makes use of three datasets to model the relationships between demand, temperature, and population size. There is presently no large-scale heat pump dataset available that sufficiently captures all of these parameters; therefore different data sources are combined in order to define the relationship between heat pump electrical demand and temperature.

The reference heat pump demand used for model training is the Renewable Heat Premium Payment (RHPP) dataset [19] features 2-minute resolution electrical demand data collected from 418 air source and ground source heat pumps in the UK from October 2013 to March 2015. This dataset does not feature location data or local weather measurements. The electricity usage monitored in this dataset does not include domestic hot water use which will not be modelled as part of this study. Due to the lack of corresponding local weather data available for individual customers in the RHPP dataset, historical weather data from a climatically average location in the UK is paired with the existing RHPP demand data to complete the reference data set. Historical weather data from the Centre for Environmental Data Analysis (CEDA) for the Central England weather station at Pershore is aligned with the RHPP demand data [18]. Pershore weather station is based at an inland, low-lying location in central England and is therefore roughly representative of weather for the majority of the UK population.

Finally, the operational demand data collected during the Low Carbon London (LCL) heat pump trials [19] is used as the target data set for validation purposes. This dataset features electrical heat pump demand and associated local temperature measurements for nine customers; this dataset is used to test that the learned model characteristics produce an accurate predicted heat pump demand from a local temperature measurement. In LCL, customers 1 to 5 and 7 feature two years worth of data; the remaining customers 6, 8 and 9 only feature 60 days of data. All customers have air-source heat pumps installed with heat pump sizes ranging from 8 to 16kW in rating.

B. DAILY DEMAND VERSUS TEMPERATURE CHARACTERISATION

Electric heat pump daily demand is broadly proportional to ambient temperature: lower ambient temperatures translates into higher heat pump daily demand, and vice versa for high ambient temperatures. The influence of parameters such as heat pump rating, efficiency, building insulation type and most importantly occupant routine behavioural parameters result in a range of possible values given a single daily average temperature measurement rather than a single possible value. This is directly observable in the LCL dataset

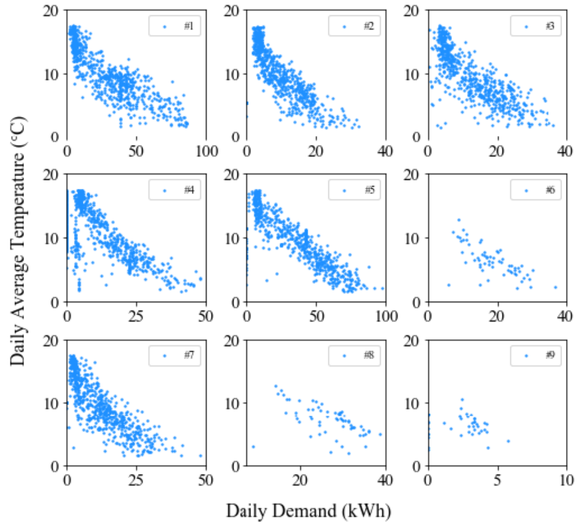

FIGURE 3. Implied joint distributions of daily demand (kWh– x axis) versus daily average temperature (°C – y-axis) for target data set heat pump loads #1 - #9.

shown in Fig.3, with a particularly wide band of possible demand values for 10 °C. Customers 6, 8 and 9 are reduced datasets only featuring 60 days of data and therefore only show a partial illustration of this characteristic – they do not capture the full operational variation due to seasonal changes. This relationship is presumed to exist in the RHPP dataset, however is masked by the lack of available corresponding temperature data. This range of possible demand values for a single daily average temperature is the basis for taking a probabilistic approach in this study. In order to create the basis for the model, the hourly demand measurements for each RHPP customer are converted from an hourly advance to daily total in kWh. In order to allow for comparison across the entire dataset, the total daily demand for each customer is normalised with respect to a reference population maxima using the formula:

$$D_{normalised} = \frac{D_{kWh}}{Dmax} \tag{1}$$

where D_{kWh} is the daily demand for a specific day and $Dmax$ is the maximum daily demand for the customer dataset being normalised. This scales all customer data on a range from 0 to 1; 0 representing zero demand and 1 representing maximum demand. The normalised customer demand is mapped to the CEDA Pershore weather station ambient temperature data for the same time period as the training dataset. The training customer demand data and weather data is then unified and plotted in the heatmap shown in Fig.4. The aggregated demand data has been split into 1 °C intervals and the distribution plotted in Fig.4. This clearly illustrates the same characteristic shape as the Low Carbon London data in Fig.3; a narrow tail for ambient temperatures above 15 °C and a widening band of higher demand for lower temperatures. Datapoints for lower temperatures are sparse in the overall

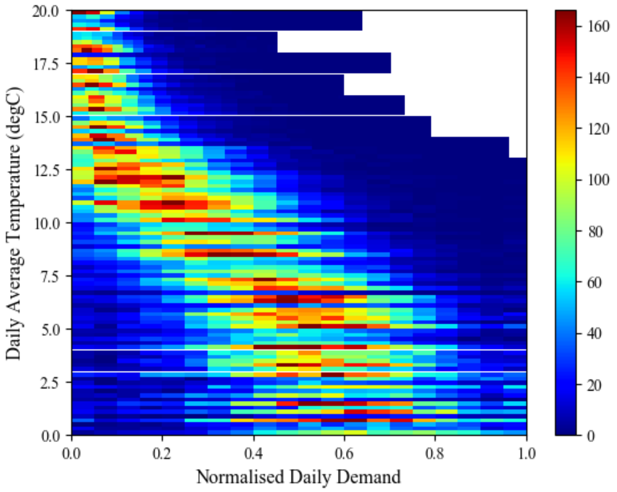

FIGURE 4. Joint histogram showing implied dependency structure of normalized daily heat pump demand against daily average temperature (°C) from training dataset.

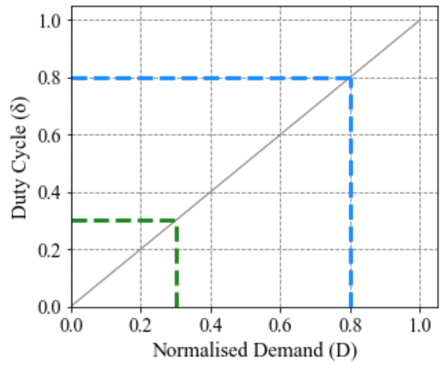

FIGURE 5. Conceptual operating envelope for daily average demand versus daily average duty cycle.

dataset and the increased granularity of the distribution at very low temperatures is visible.

The standard deviation and mean is calculated for each of the 1 °C dataset intervals, translating the raw data into a simplified Gaussian distribution for each temperature band. This provides a plausible range of normalised demand values for each 1 °C interval. This is represented as:

$$f_t\,(x) = \frac{1}{\sigma_t\sqrt{\hat{I}2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu_t}{\sigma_t}\right)^2} \tag{2}$$

Therefore when making a demand prediction the daily average temperature is first identified, and then the demand value is generated based on a distribution determined by the corresponding standard deviation and mean for that temperature that has been derived from the data in Fig.4. This allows for the creation of a look up table of standard deviation and mean values for each temperature. The normalised probability density is scaled into kWh using the formula shown in (3).

$$D_{Predicted(kWh)} = f_t\,(x) \times \frac{D_{kWh}}{D_{Normalised}} \tag{3}$$

C. DAILY DUTY CYCLE

A model for predicting daily heat pump energy has been derived however it is still necessary to characterise how this energy is distributed throughout the day. For this purpose the modelled duty cycle will be disassociated from time of day characteristics. The purpose of this section is to derive the proportion of on and off durations for a particular day, not how heat pump activity is distributed throughout the day.

Heat pump outputs can be one of two types: fixed output or inverter based variable output. Fixed output heat pumps operate by cycling on and off between maximum power during the required heating period. Inverter based heat pumps can modulate their output to any intermediate point between zero and full power as required to meet heating demand. Typical fixed output heat pumps operating periods range from 9 to 40 minutes [22], therefore over a time period of one hour the power characteristics of a fixed output versus inverter based output will average to the same waveform in order to meet the same heating demand for the majority of cases if conversion efficiencies are assumed to be identical.

Conceptually, therefore the daily average demand is directly proportional to the amount of time the heat pump spends in the ''on''-state. The ratio of time spent in the on-state versus the off-state will be represented as a duty cycle measure. On this basis this work will develop a two-state model for linking the previously derived daily heat pump energy in section III.B to an hourly demand figure.

The modelled duty cycle δ_{AV} is calculated from the ratio of the predicted daily demand $D_{Predicted(kWh)}$ over the real maximum demand $Dmax$ for the customer profile. This is further scaled by duty cycle population data through the value δmax as shown in (4). Each customer profile has their own fixed value of δmax, representing the maximum time the heat pump spends on at the cold operational extreme. The theoretical upper limit for δmax is 1 (representing always on) however the mean δmax obtained from the training population is 0.68, representing a heat pump that is on 68% of the time at its upper operational extreme. The values of δmax are derived from the RHPP population dataset for validation purposes.

For an instantaneous sample period heat pumps of any output type can be assumed to be in one of two states: on or off. Under steady state conditions the on-state can be assumed to be fixed, although ramping to steady state will introduce intermediate values of demand. Therefore the daily heat pump demand is directly proportional to the amount of time a heat pump spends on the on-state. Fig.5 and (4). illustrate the conceptual relationship between the maximum daily demand $Dmax$, the maximum daily duty cycle δmax scaling factors, the predicted daily demand $D_{Predicted(kWh)}$ and the derived duty cycle δ_{AV}.

$$\delta_{AV} = \frac{D_{Predicted(kWh)}}{Dmax} \times \delta max \tag{4}$$

The duty cycle δ_{AV} therefore represents the average heat pump state for the daily time period. Once the daily demand $D_{Predicted(kWh)}$ and the daily average duty cycle δ_{AV} is known the on-time t_{on}, off-time t_{off} and on-power P_{on} can be determined. For a single day this is then used to derive the t_{off}, t_{on} and P_{on} for the heat pump as shown in (5), (6) and (7). This makes the assumption that the heat pump is either fully on or fully off with no intermediate values.

$$t_{on} = (t_{on} + t_{off}) \times (\delta_{AV}) \tag{5}$$

$$t_{off} = \frac{t_{on} \times (1 - \delta_{AV})}{\delta_{AV}} \tag{6}$$

$$P_{on} = \frac{D_{Predicted(kWh)}}{\delta_{AV}} \tag{7}$$

Fig.6 illustrates the relationship between the original observed demand and the modelled demand P_{on} derived from (5), (6) and (7). The left hand figure shows the original observed demand over the course of a day; the right hand-figure shows the same observed demand dataset but sorted by magnitude rather than plotted versus time. This is plotted alongside the modelled demand P_{on}, illustrating how derivation of δ_{AV} from $D_{Predicted(kWh)}$ can be used to model a good approximation for the proportion of time a heat pump spends in either the on or off state. The distribution of the modelled on/off states for a single day are derived in the next section.

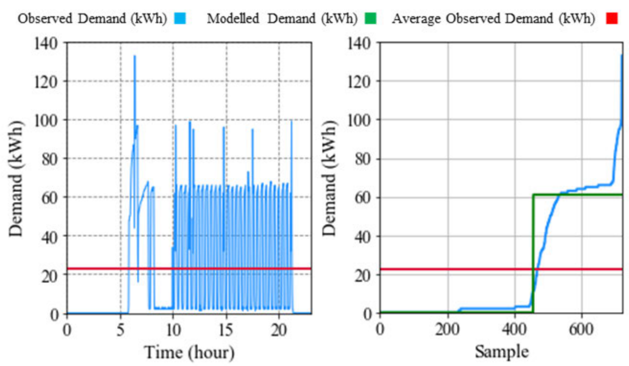

FIGURE 6. **(left) Observed Demand (kWh/2mins), (right) Modelled Demand (kWh/2mins) and Observed Demand (kWh/2mins) for single day.**

D. HOURLY DUTY CYCLE

The remaining aspect of the model is to develop a method for determining heat pump hourly demand profile from the daily average energy and daily average duty cycle. As illustrated in Fig.1 the shape of a daily electrical demand profile can vary significantly from household to household due to heat pump and building parameters, as well as behavioural routines. The primary concern around heat pump type loads is that without a storage element to buffer or shift time of use, heat pump loads can be anticipated to feature low levels of temporal diversity in combination with their high energy and high power characteristics. Therefore the specific time of day for heat pump activity becomes of critical interest – a load that is distributed evenly throughout the day will not pose the same risks to voltage and thermal limits as a load that tends to be clustered around existing the domestic load peaks in the morning and evening. The time of use patterns across the entire RHPP dataset will be characterised and then fed back into a validation model.

Ideally power would be modelled as an instantaneous value, however due to the variability in heat pump type and operation, it is not possible to develop a high-resolution predictive model that is fully applicable for the entire dataset. The approach here is to therefore develop a practical method for characterising sub-daily demand magnitudes to a reasonable resolution for network based analysis and validation. The aim of this model is not to detect or

characterise fast transients (which are better predicted by physics based models), but rather steady state network conditions and how they contribute to network limits. This work will therefore model hourly demand magnitude rather than instantaneous power. Further work would be possible to reduce this time resolution further for specific applications.

It has been shown in section 3.3 that modelling a linear relationship between the daily demand and daily average duty cycle δ_{AV} through (4) allows for derivation of the t_{on}, t_{off} and P_{on} values for a particular day. This section will outline a framework for linking the derived daily average duty cycle to a set of hourly duty cycle values determined by ambient temperature. This set of hourly duty cycles will retain the overall predicted t_{on}, t_{off} and P_{on} values determined by the δ_{AV} for a particular day. The individual time of day versus temperature relationships for all RHPP customers will be aggregated into a single framework that can be used to generate synthetic demand profiles. Fig.7 illustrates a sample raw demand profile translated into daily and hourly duty cycle for a single customer.

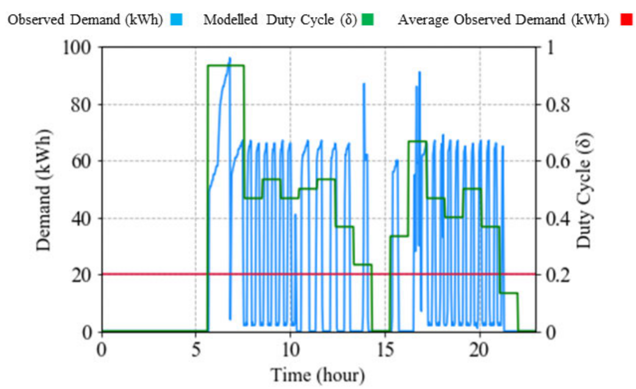

FIGURE 7. **Real demand (kWh/2 mins, blue), modelled duty cycle (green) and average real demand (red).**

E. HEAT PUMP DEMAND FROM TEMPERATURE TRANSLATION MODEL

This section describes the process for translating the daily average duty cycle δ_{AV} into a corresponding set of hourly duty cycle values, allowing for shaping of an overall daily demand profile. Building on the relationship between daily demand and daily duty cycle defined in section III.D, for all daily customers the raw load profile is translated into an hourly duty cycle and temperature profile relation. Due to the direct relationship between heat pump energy consumption and ambient temperature, the time of use characteristics for a particular heat pump will also vary with temperature – heat pump activity during the day reduces with warmer temperatures and vice versa for cold temperatures. The output of this conditioning stage is that for each unique customer there exists an aggregated duty cycle profile for each hour of the day and temperature combination dataset The data conditioning process to transform the raw demand data into a temperature versus hourly duty cycle profile is outlined graphically in Fig.8. For each customer, the raw demand profile for each 1°C slice is converted to instantaneous heat pump state as shown in (8), where *Emax* represents the maximum instantaneous demand magnitude for the day. This is then converted to hourly duty cycle. Finally, all profiles belonging to the same customer and 1°C temperature set are averaged to create an aggregated profile of hourly duty cycle versus temperature through (10), where n represents the number of datasets available for a particular customer and 1°C temperature combination. This process does not make any distinction between weekend, weekday or any exceptional days such as holidays or bank weekends. The impact of the weekday/weekend distinction on modelling strategies has been assessed in [23], which identifies clear changes in load routines during weekdays versus weekends. The aim of this work however is to first develop a generalised model that can then be tailored to suit specific analysis tasks.

$$E(t) = \begin{cases} 1 & t > 0.1(Emax) \\ 0 & t < 0.1(Emax) \end{cases} \tag{8}$$

$$\delta_{h\_n} = \frac{1}{30} \sum_{t}^{t+30} E(t) \tag{9}$$

$$\delta_h = \frac{1}{n} \sum_{1}^{n} \delta_{h\_n} \tag{10}$$

Each customer therefore has a 24 x 20 array where there is a row associated with each hour of the day and a column associated with each 1°C temperature slice, with each cell representing an hourly duty cycle δ_{h} that reflects the heat pump activity for those conditions. The array is defined as follows for each customer:

$$\delta_h = f(temperature, hourofday) \tag{11}$$

The final output of this process is shown at the bottom of Fig.8 as the hour of day versus temperature plot. This shows an example of a time versus temperature relationship for a single customer. Fig.9 shows a further selection of temperature versus time of use profiles for nine additional customers: this clearly illustrates the diversity in time versus temperature relationships for multiple customers. From this limited selection it can be seen that customers tend to retain heat pump

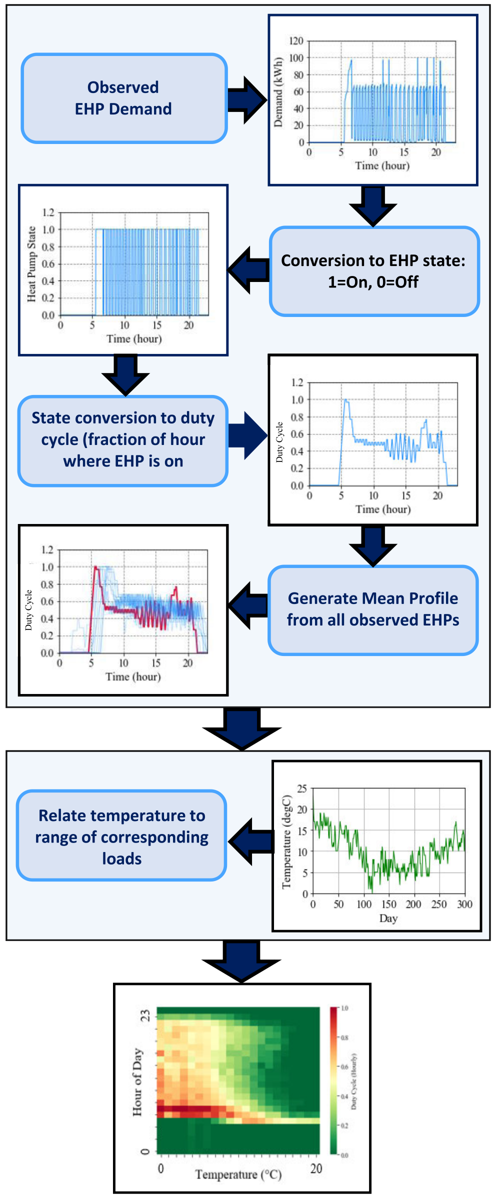

FIGURE 8. Conditioning of raw heat pump electrical demand profile into time of use profile versus temperature.

behaviours across the temperature range. Customers that do not enable heating during the day for cold extremes tend not to enable heating for any other temperature. Similarly, customers that have heating operating continuously at cold extremes still exhibit this behaviour at warmer temperatures.

F. DAILY TO HOURLY DEMAND RELATION LEARNING

This section will further condition the data in order to link sets of hourly duty cycles to a single daily average duty cycle figure and therefore shape a demand curve based on a single daily average duty cycle value. Each of the 418 temperature versus time of day profiles derived from the RHPP dataset and defined in (11) are combined into a three-dimensional array

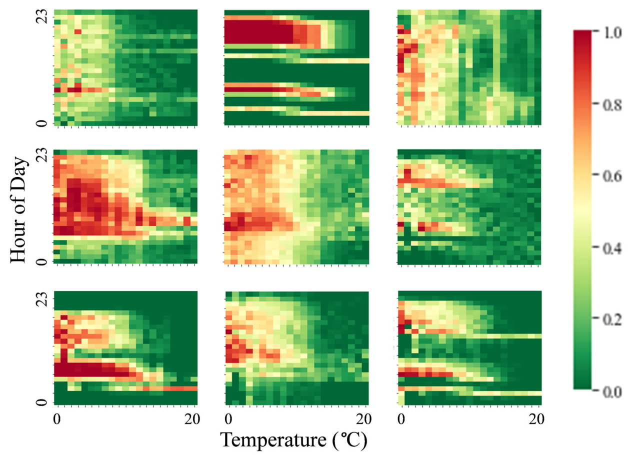

FIGURE 9. Selection of nine customer temperature versus time of use profiles from training dataset.

defined as:

$$\delta_h = f(temperature, hourofday, customer) \tag{12}$$

The customer axis is transformed into a numeric δ_{av} value by computing the average daily duty cycle for each customer and temperature set of 24 δ_h values:

$$\delta_{av} = \frac{1}{24}\sum_{0}^{23}\delta_h \tag{13}$$

The array in (12) is now modified with the customer axis being replaced with the δ_{av} corresponding figure calculated in:

$$\delta_h = f(temperature, hourofday, \delta_{av}) \tag{14}$$

Finally, this array sorted by average daily duty cycle. This sorts the array by heat pump activity for the entire population sorted by most active to least active.

Fig.10 shows the 0°C slice of this array. This array shows the distribution of heat pump activity across the entire RHPP population versus time of day. Clear morning and evening peaks are visible, but there are also customers with very high and very low heat pump activity at either extreme. This array clearly illustrates that heat pump activity exists on a continuum rather than there being clearly defined repeating profiles. It is theorised that this characteristic will be true for all UK based heat pump populations over a certain size that feature a certain level of diversity. The subsequent temperature slices in Fig.11 clearly show the reduction in heat pump activity with temperature. The onset of the morning demand peak at 6am and drop off at 10pm correlates closely with the December averaged profiles obtained from the Customer Led Network Revolution heat pump study, which consisted of 89 customers [8].

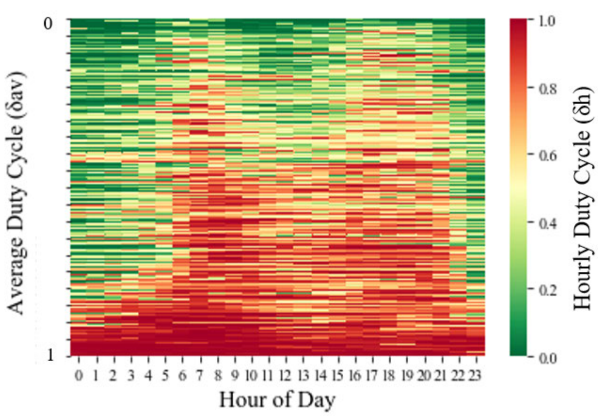

FIGURE 10. Heat Pump Hourly Duty Cycle (0 to 24 hours, x-axis) versus Average Duty Cycle (y-axis).

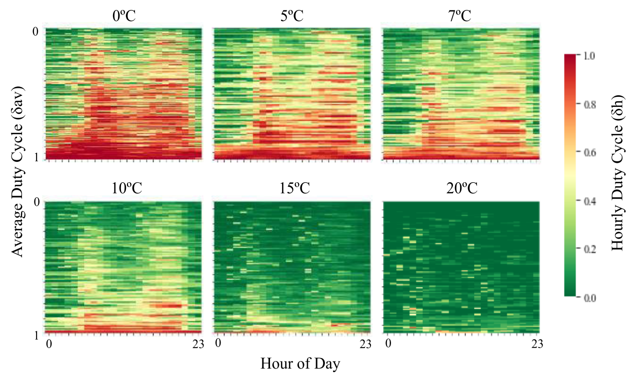

FIGURE 11. RHPP Heat Pump Hourly Duty Cycle Distribution - (0 to 24 hours, x-axis) versus Average Duty Cycle (y-axis) for 0°C, 5°C, 7°C, 10°C, 15°C and 19°C.

It is now possible to generate an hourly demand estimate using only a daily average temperature input combined with the demand and duty cycle linear scaling factors. The daily demand is generated as per the relationship shown in Fig.4. From this a daily average duty cycle can be obtained through the relationship between daily demand and daily average duty cycle shown in (4). Finally the daily average duty cycle is paired with the closet matching set of hourly duty cycles for the appropriate temperature in (14). The predicted hourly demand $D_{predicted\_h}$ for each hour of the day being calculated is obtained though (15):

$$D_{predicted\_h} = \delta_h \times \frac{D_{predicted(kWh)}}{24} \tag{15}$$

where $D_{predicted(kWh)}$ is the predicted daily power and δh is the predicted hourly duty cycle. Fig.12 illustrates the final load profile output of the process for a single theoretical customer and range of temperature values.

G. TRAINING DATASET – CUSTOMER PROFILES

The RHPP training dataset acts as the source of scaling factors for daily demand *Dmax* and duty cycle *δmax* in this study in (1) and (4). The structure of this model is such that the model may be seeded with scaling factors from other populations or datasets as required for specific analysis. This section

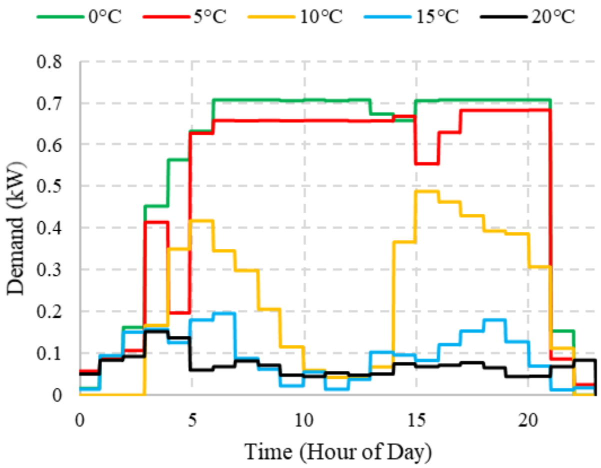

FIGURE 12. Daily demand profiles for a single customer for 0°C, 5°C, 10°C, 15°C and 20°C.

describes the general population characteristics of the scaling factors used within this study.

The limitations of using the RHPP dataset method have been identified in previous works – customers in the RHPP dataset are predominately local authority landlords rather than private tenants [9]. Additionally, heat pump technology has advanced since the installation of the sample population hardware in 2013 and therefore this may not fully be representative of a modern population due to improvements in achievable COPs and building efficiencies.

The normalised demand generated in section 3.2 is multiplied from the scaling factor $Dmax$ derived from the RHPP population demand magnitude. Fig.13 illustrates the mean daily demand for each customer across the RHPP dataset. The population predominately features customers at the lower end of the mean daily demand.

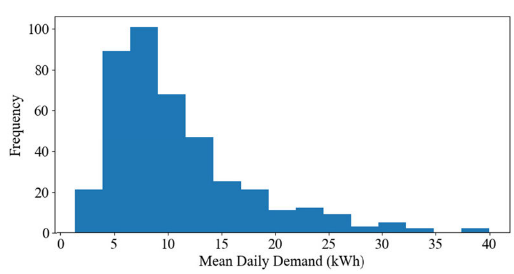

FIGURE 13. Distribution of RHPP Customers by mean heat pump daily demand (kWh).

The maximum duty cycle δmax per customer in the RHPP dataset is shown in Fig.14. This shows that apart from customers with very low heat pump demand, maximum duty cycle is not strongly correlated to heat pump demand and is distributed normally throughout the dataset. The right hand side of Fig.14 shows that maximum duty cycle is approximately normally distributed throughout the RHPP dataset.

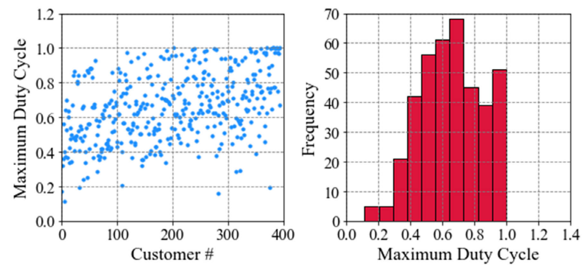

FIGURE 14. (left) Maximum duty cycle versus heat pump maximum demand (kWh), (right) Distribution of maximum duty cycle for RHPP customer dataset (kWh).

IV. VALIDATION OF PREDICTED DEMAND PROFILES

The developed model is validated against the observed demand data in the RHPP training dataset and the LCL target dataset. The model is fundamentally derived from training dataset characteristics and therefore the tests versus the LCL dataset validate how well the method works versus unseen data. Large scale population testing is performed with the training dataset, using the user profiles and demand shapes generated from the training dataset. The inputs used to generate a demand prediction are for a single day are the daily average demand, and the customer scaling factors: maximum daily demand $Dmax$ and maximum daily duty cycle δmax. These inputs are then fed into the thermal relations derived in this paper in order to generate daily and hourly demand predictions. Random populations of customers ranging from 1 to 160 are used; with the error being assessed as a simple mean absolute percentage error from 0°C to 15°C. Heat pump behaviour below 0°C is beyond the scope of this model due to the lack data for this condition in the used datasets; beyond 15°C is not examined as beyond this point heat pump activity becomes minimal in the real data in terms of energy. Further testing is performed to assess the quality of the predicted demand shape with respect to the actual demand shape; this is assessed as the percentage overlap between the real demand and predicted demand.

A. DAILY DEMAND TESTING

The real versus predicted daily demand is calculated for randomly selected groups of 1, 5, 10, 20, 40, 80 and 160 customers. For each size group, a random selection of customers is selected from the training dataset and the real daily demand versus predicted daily demand calculated. This process is replicated 100 times for each group size in order to obtain a mean and standard deviation for error, with a new random selection of customers generated each time. 100 runs per customer group is chosen as a trade-off between computing time and accuracy. From this the mean absolute percentage error for each customer in each group is calculated as shown in, where D_{real_d} is the real daily demand and $D_{predicted\_d}$ is the predicted daily demand as derived from (3). This process is then repeated for the temperature points 0°C, 5°C, 10°C, and 15°C. The final MAPE for each temperature/group number

TABLE 1. Daily demand MAPE and its standard deviation for 0, 5, 10, 15°C and aggregations of 1, 10, 20, 40, 80, 160 customers using training dataset.

	0ºC		5ºC	
#	MAPE (%)	σ_d (%)	MAPE (%)	σ_d (%)
1	38.6	43.3	82.9	285.9
5	23.7	23.1	21.4	18.6
10	17.0	13.8	18.7	13.7
20	15.1	11.4	15.1	8.9
40	12.1	7.2	13.9	8.8
80	12.6	4.6	13.2	4.9
160	12.6	3.2	12.3	3.2

	10ºC		15ºC	
#	MAPE (%)	σ_d (%)	MAPE (%)	σ_d (%)
1	82.8	230.0	68.3	109.4
5	20.8	18.6	24.6	20.6
10	16.3	10.8	18.5	13.3
20	13.2	8.9	12.5	8.8
40	13.0	8.1	11.0	7.9
80	12.3	4.5	8.4	4.8
160	12.8	3.4	7.6	5.5

combination is simply the average MAPE obtained for each set. The percentage standard deviation σ is calculated as per (17). The results for this process are shown in Table 1.

$$MAPE = \sum \left| \frac{D_{real\_d} - D_{predicted\_d}}{D_{real\_d}} \right| \times 100 \quad (16)$$

$$\sigma_d = \sqrt{\frac{\sum (D_{predicted\_d} - \mu)^2}{n}} \quad (17)$$

For the target dataset there are nine customers and the maximum overlap in time for the overall dataset is a 60 day period ranging from 17/01/2014 to 20/03/2014. The aggregated customer demand from this period is used to evaluate the training dataset derived model for predicting demand for other customer datasets. Two temperature profiles are used to synthesise the daily demand: one is the Pershore weather dataset used for the training dataset, the second is the averaged local temperature data from the target dataset. Fig.15 shows the real versus predicted daily demand for the aggregated target data profile over a two month winter period. This shows a smaller than anticipated error between real and predicted demand when compared to the RHPP findings. This can be attributed to the fact that the Pershore weather data does not reflect the local ambient temperature conditions for the target dataset customers, which are distributed throughout the south-east of England. However, this result does illustrate that the normalised demand versus temperature relationship derived from the training dataset shown in Fig.4, combined with a simple scaling factor is able to achieve good results for daily demand for small heat pump population even with

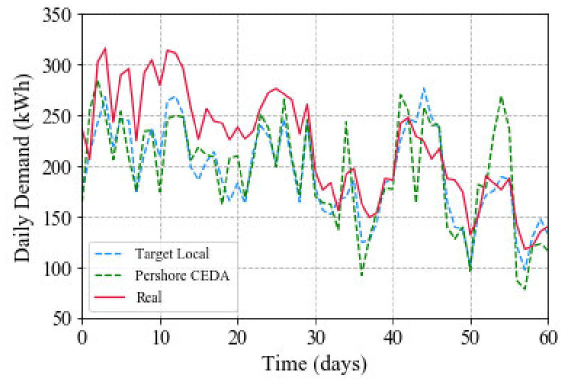

FIGURE 15. Real Demand, Predicted Demand (Target Local Weather) and Predicted Demand (Pershore CEDA Weather) for 9 aggregated LCL customers from 17/01/2014 to 20/03/2014.

a non-local temperature series. The MAPE and σ modelled using the Pershore weather dataset in Table 2 align well with the corresponding training dataset error values in Table 1 for population sizes of 10.

TABLE 2. Daily demand MAPE and its standard deviation for target customer dataset using target local weather and Pershore CEDA weather observations.

	Target Local Weather		Pershore CEDA Weather	
#	MAPE (%)	σ_d (%)	MAPE (%)	σ_d (%)
9	14.4	8.1	18.6	11.1

B. HOURLY DEMAND TESTING

There are several features of interest when examining the shape of a daily demand profile for network design and operations. The magnitude of the demand peak, the time of the demand peak, in addition to maximum rates of change are all of interest when assessing network impact, however this list is not exhaustive. The MAPE method used in the previous section is not suitable for measuring the shape quality; small values feature heavily in the dataset at higher temperatures and large errors in small hourly demands which can skew the whole figure. In order to overcome these issues a weighted MAPE is used. This MAPE is weighted by the sum of total real demand for a given day and is calculated as shown in (18), where D_{real_h} is real hourly demand, and $D_{predicted\_h}$ is the corresponding predicted value of demand for the same hour generated from (15).

$$WMAPE = \frac{\sum_{h=0}^{23} \left| D_{real\_h} - D_{predicted\_h} \right|}{\sum D_{real\_h}} \times 100 \quad (18)$$

$$\sigma_h = \sqrt{\frac{\sum (D_{predicted\_h} - \mu)^2}{n}} \quad (19)$$

V. MODEL PERFORMANCE EVALUATION

The proposed model in this study has outlined a simple approach for predicting daily and hourly heat pump demand profiles for a user-defined sample population, using only daily average temperature and linear scaling factors as inputs. The performance of the model against two independent datasets have been examined in order to evaluate the wider applicability of this model for UK based heat pump population, with consistent results across both the training RHPP and target LCL datasets. Whilst constrained by the lack of further available heat pump demand data to examine this point further, these initial results do indicate that the derived model will be applicable in general for UK customers.

As illustrated in Fig.12 and Fig.15 the model offers good capability for predicting the magnitude of daily demand for heat pump populations for operating conditions ranging from 0°C to 15°C. The mean daily demand error reduces as temperature increases; this is in line with the dependency relationship plotted in Fig.4 which features a smaller band of possible values for higher temperatures when compared to cold temperatures. In contrast, the mean error for hourly demand increases with temperature. This can be attributed to the greater diversity in demand shape at warm temperatures as users are more likely to transition to only operating heating for limited time windows.

The general error characteristics for both the daily and hourly demand tests reflect findings of previous LV studies which observed strong scaling relationships between the number of households and MAPE of a forecast method [24]. Whilst there are no directly comparable heat pump studies using MAPE that the results of this work can be compared against, it can be broadly compared with existing LV studies forecasting other load types. Previous works forecasting LV substation loads using have achieved MAPE's in the region of 11%–16% utilising ARIMA methods [25] [26]. While the model contributed here is a predictor of load from temperature, it could offer a forecasting capability if used in conjunction with a temperature forecast from a numerical weather prediction model. Typically forecasting temperature yields lower errors than demand so the anticipated heat pump demand forecast error would be broadly aligned with this figure. The demand activity peaks shown in Fig.8 and Fig.9 are in agreement with the demand peaks of averaged heat pump demand data from a comparable but geographically separate trial [8].

Whilst limited in size compared to the overall training dataset, the results for the target population show consistent error results when compared with the RHPP error for groups of a similar size. This does suggest that the RHPP derived characteristics for the demand and demand shape model will be widely applicable for UK households.

As has been shown, heat pump electrical demand magnitude is highly sensitive to temperature. Whereas existing works tend to focus on the demand impact of heat pumps for the extreme cold case [4] [15] [16], this work facilitates the generation of representative heat pump demand profiles ranging from 0°C up to 20°C. Given the increased penetration of low carbon technologies (LCTs) on LV networks, it becomes of increased importance to model the combined effects of LCTs alongside conventional loads rather than study the extreme case for one technology type in isolation. The future LV power system will need to be safely rated to incorporate the effects of photovoltaics, wind, and electrical vehicles in addition to low-carbon heating. The temperature sensitivity of this model allows for generation of demand profiles for any seasonal condition rather than the extreme case, enabling study of heat pump effects alongside other technologies.

Table 2, which uses local weather data, illustrates a noticeable effect on the final error of a demand forecast. A typical winter day is therefore expected to be different in shape and magnitude depending on the local climate extremes – the model presented can generate locally specific demand profiles alongside a quantified error, rather than using an extreme winter case not tailored to local conditions.

Whilst this model captures the behavioural time of use relationship that is typically absent from physical models, there are certain pre-requisites to consider when using this method to predict electric heat pump demand. In particular, this model is dependent on source user profiles in order to seed appropriate scaling factors when performing demand normalisation and a proxy for these on de-normalisation. The scaling factors used in this study are contemporary to the capabilities of heat pump technology at the time of the original study. In order to revise these scaling factors for future generations of hardware, these values would have to be adjusted in taking into account typical COP's and critical physical parameters for new hardware. Scaling factors will inevitably be a function of building parameters such as floor area, building layout and insulation efficiency as well as heat pump rating, itself related to the latter parameters. Building floorspace has been shown to have the greatest influence on household heating demand [27]; a potential opportunity for further work would be to derive building characteristics including floorspace from remote imagery or aerial lidar data in order to automatically define scaling factors tailored to a local area [28].

It has not been possible to model the effects of heat pump demand below 0°C due to the very limited availability of data from this extreme operating region. Below 0°C the COP of conventional EHP's drops off significantly, greatly reducing conversion efficiency. The typical mitigation strategy to counter this behaviour is to install secondary resistive heating to supplement the heat pump output for extreme cold cases. This raises the threat of yet higher demand peaks that are driven by temperature and would require a second model to incorporate the load characteristics of this behaviour.

VI. APPLICATIONS AND CASE STUDY

Through the development of a EHP-specific load model, this work facilitates the further analysis of EHP impact on LV networks both in isolation, and alongside the effects of other

TABLE 3. Hourly demand WMAPE and its standard deviation (σ) for 0, 5, 10, 15°C using training dataset.

	0ºC		5ºC	
#	WMAPE (%)	σ_h (%)	WMAPE (%)	σ_h (%)
1	66.1	14.7	68.5	16.8
5	37.4	7.7	37.7	6.9
10	27.0	5.8	31.2	5.3
20	20.1	3.7	29.7	4.1
40	15.6	3.1	17.4	3.9
80	13.1	2.2	14.6	4.7
160	11.6	2.0	13.6	2.1

	10ºC		15ºC	
#	WMAPE (%)	σ_h (%)	WMAPE (%)	σ_h (%)
1	81.3	13.7	87.1	9.0
5	50.2	7.3	69.9	7.0
10	38.3	5.6	58.1	6.6
20	31.2	4.1	47.7	5.1
40	24.0	4.0	38.2	3.7
80	17.9	4.7	32.3	2.9
160	14.6	2.2	27.7	2.2

TABLE 4. Hourly demand WMAPE and its standard deviation for target customer dataset using target local weather and Pershore CEDA weather.

	Target Local Weather		Pershore CEDA Weather	
#	WMAPE (%)	σ_h (%)	WMAPE (%)	σ_h (%)
9	35	13.1	37	15.6

low-carbon technologies. The method could be used for a range of applications including; examining EHP penetration network impact, as part of a demand response analysis or as part of mixed-energy network studies. The probabilistic approach allows for confidence intervals to be defined alongside a load prediction; these could be derived from the relevant MAPE and σ in Table 1 and Table 3. This complements existing probabilistic approaches for other LV-connected low carbon technologies [29], [30]. It is therefore envisioned that this heat pump specific model could be used alongside other probabilistic load types in order to thoroughly examine possible network conditions in the presence of load and generation uncertainty.

The main challenge EHPs present is the effect at scale at the last mile of distribution networks. Underground cables at this part of the network accounts for a significant volume of the assets of a network owner and replacement or reinforcement of these to accommodate EHP load may require an investment beyond their capabilities.

As an example application, the predictive model is used to model a simple power flow scenario for a single LV feeder with 40 connected households and varying levels of EHP penetration. Smart meter data from the Low Carbon London trials [31] is used to create a base domestic load and combined with increasing levels of EHP penetration on the feeder. Heat pump electrical demand is then predicted for an artificial five day winter period; Fig.16 illustrates the output of this study.

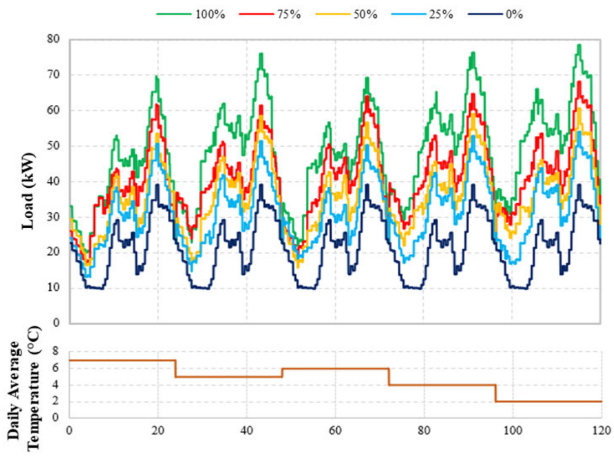

FIGURE 16. Distribution network feeder load for 0%, 25%, 50%, 75% and 100% EHP penetration on 40 customer residential network during five day winter period (a-top) and corresponding daily average temperature (b-bottom).

The fundamental shape of the overall load profile does not significantly change between the 0% and 100% penetration cases; the morning and afternoon peaks are roughly concurrent for all levels of penetration however the morning peaks additionally get wider. However the magnitude of the daily load peaks can be seen to significantly increase in value, with the evening peak approximately doubling in value for the 100% case. This agrees with the expectation outlined at the outset of this paper that full heat pump penetration is roughly equivalent to doubling the number of households on a network.

VII. CONCLUSION

This paper has defined a model for quantifying the demand impact of increased uptake of electric heat pumps for population sizes representative of typical LV network applications using demand relationships derived from existing operational datasets and sensitive to local weather conditions. A generic relationship between heat pump electrical demand and temperature has been identified from real customer data and validated on two independent datasets. This model facilitates the analysis of heat pump demand that is sensitive to local temperature conditions, rather than blanket rescaling of existing customer data as has been performed for previous studies, augmenting the utility of sparsely available demand data. By using a probabilistic approach, the distribution of prediction error has been quantified. This creates future opportunities for examining heat pump demand sensitivity for different geographical locations against existing heat pump assessments, as well as performing studies which incorporate multiple low carbon technologies connected to a LV network.

The main priority for further work would be to relate the magnitude of electrical demand to an estimated COP and nameplate heat pump rating, such that the scaling factors used for the model could be modified to accommodate improvements in heat pump efficiency. It would additionally be of interest to examine the variation in weekday, weekend and exceptional events.

ACKNOWLEDGMENT

Thanks to Frank Clifton of SSE for his advice and cooperation. Underlying heat pump demand data can be accessed via [18] and [19], respectively.

REFERENCES

[1] Catapult Energy Systems. (2019). *Living Carbon Free—Exploring With a Net-Zero Target Means for Households*. Accessed: Feb. 2, 2020. [Online]. Available: https://www.theccc.org.uk/wp-content/uploads/2019/06/ESC-Living-Carbon-Free-CCC.pdf

[2] M. Ahmadi, M. Ahmadi, M. Sadaghiani, M. Ghazvini, S. Shahriar, and M. A. Nazari, "Ground source heat pump carbon emissions and ground-source heat pump systems for heating and cooling of buildings: A review," *Environ. Prog. Sustain. Energy*, vol. 37, no. 4, pp. 1241–1265, 2017.

[3] Committee on Climate Change. (Oct. 2016). *Next Steps for UK Heat Policy*. Accessed: Feb. 2, 2020. [Online]. Available: https://www.theccc.org.uk/wp-content/uploads/2016/10/Next-steps-for-UK-heat-policy-Committee-on-Climate-Change-October-2016.pdf

[4] Department for Business, Energy & Industrial Strategy. (Dec. 2016). *Renewable Heat Incentive: A Reformed Sheme*. Accessed: Feb. 22, 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/577024/RHI_Reform_Government_response_FINAL.pdf

[5] UCL Energy Institute. (Mar. 2017). *Final Report on Analysis of Heat Pump Data from the Rewnewable Heat Premium Payment Scheme*. Accessed: Feb. 22, 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/606818/DECC_RHPP_161214_Final_Report_v1-13.pdf

[6] National Statistics. (Jan. 23, 2020). *RHI Monthly Deployment Data: December 2019 (Annual Edition)*. Accessed: Feb. 2, 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/859989/RHI_monthly_official_stats_tables_Dec_19_final.xlsx

[7] S. Haben, G. Giasemidis, F. Ziel, and S. Arora, "Short term load forecasting and the effect of temperature at the low voltage level," *Int. J. Forecasting*, vol. 35, no. 4, pp. 1469–1484, Oct. 2019.

[8] Durham Energy Intstitute and Element Energy. (Jan. 23, 2015). *CLNR—Insight Report: Domestic Heat Pumps*. Accessed: Feb. 2, 2020. [Online]. Available: http://www.networkrevolution.co.uk/wp-content/uploads/2015/01/CLNR-L091-Insight-Report-Domestic-Heat-Pumps.pdf

[9] J. Love, A. Z. Smith, S. Watson, E. Oikonomoua, A. Summerfield, C. Gleeson, P. Biddulpha, L. F. Chiu, J. Wingfield, C. Martin, A. Stone, and R. Lowe, "The addition of heat pump electricity load profiles to GB electricity," *Appl. Energy*, vol. 204, pp. 332–342, Oct. 2017.

[10] B. Asare-Bediako, W. L. Kling, and P. F. Ribeiro, "Future residential load profiles: Scenario-based analysis of high penetration of heavy loads and distributed generation," *Energy Buildings*, vol. 75, pp. 228–238, Jun. 2014.

[11] Ofgem. (Jun. 22, 2017). *Typical Domestic Consumption Values for Gas and Electricity*. Accessed: Apr. 13, 2020. [Online]. Available: https://www.ofgem.gov.uk/publications-and-updates/typical-domestic-consumption-values-gas-and-electricity-0

[12] C. P. Underwood, M. Royapoor, and B. Sturm, "Parametric modelling of domestic air-source heat pumps," *Energy Buildings*, vol. 139, pp. 578–589, Mar. 2017.

[13] C. Protopapadaki and D. Saelens, "Heat pump and PV impact on residential low-voltage distribution grids," *Appl. Energy*, vol. 192, pp. 268–281, Apr. 2017.

[14] I. A. G. Wilson, A. J. R. Rennie, Y. Ding, P. C. Eames, P. J. Hall, and N. J. Kelly, "Historical daily gas and electrical energy flows through great Britain's transmission networks and the decarbonisation of domestic heat," *Energy Policy*, vol. 61, pp. 301–305, Oct. 2013.

[15] J. Conrad and S. Greif, "Modelling load profiles of heat pumps," *Energy Efficiency Plants Buildings*, vol. 12, no. 4, p. 766, 2019.

[16] A. Navarro-Espinosa and P. Mancarella, "Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks," *Appl. Energy*, vol. 127, pp. 249–266, Aug. 2014.

[17] C. Hanmer, M. Shipworth, D. Shipworth, and E. Carter, "How household thermal routines shape UK home heating demand patterns," *Energy Efficiency*, vol. 12, no. 1, pp. 5–17, Jan. 2019.

[18] R. Lowe, "Renewable heat premium payment scheme: Heat pump monitoring: Cleaned data, 2013–2015," Dept. Energy Climate Change, Data Collection, U.K. Data Service, Colchester, U.K., Tech. Rep., 2017, doi: 10.5255/UKDA-SN-8151-1.

[19] *Low Carbon London Heat Pump Load Profiles*, UKPN, UK Power Netw., London, U.K., 2017.

[20] Y. Fujimoto, S. Furuya, Y. Hayashi, and T. Osaka, "Generating synthetic profiles of onshore wind power for power flow simulation on power system," *J. Energy Eng.*, vol. 143, no. 3, Jun. 2017, Art. no. F4016001.

[21] S. Hosseini, S. Kelouwani, K. Agbossou, A. Cardenas, and N. Henao, "A semi-synthetic dataset development tool for household energy consumption analysis," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Toronto, ON, Canada, Mar. 2017, pp. 564–569.

[22] Department of Energy and Climate Change. (Nov. 2012). *The Effects of Cycling on Heat Pump Performance*. Accessed: Apr. 13, 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/65695/7389-effects-cycling-heat-pump-performance.pdf

[23] B. Stephen, X. Tang, P. R. Harvey, S. Galloway, and K. I. Jennett, "Incorporating practice theory in sub-profile models for short term aggregated residential load forecasting," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1591–1598, Jul. 2017.

[24] R. Sevlian and R. Rajagopal, "A scaling law for short term load forecasting on varying levels of aggregation," *Int. J. Electr. Power Energy Syst.*, vol. 98, pp. 350–361, Jun. 2018.

[25] D. Alberg and M. Last, "Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms," *Vietnam J. Comput. Sci.*, vol. 5, nos. 3–4, pp. 241–249, Sep. 2018.

[26] N. Ding, C. Benoit, G. Foggia, Y. Besanger, and F. Wurtz, "Neural network-based model design for short-term load forecast in distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 72–81, Jan. 2016.

[27] Department of Energy & Climate Change. (Jun. 2013). *National Energy Efficiency Data—Framework*. Accessed: Aug. 11, 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/209089/National_Energy_Efficiency_Data-framework_June_2013_Part_I.pdf

[28] N. L. Gavankar and S. K. Ghosh, "Automatic building footprint extraction from high-resolution satellite image using mathematical morphology," *Eur. J. Remote Sens.*, vol. 51, no. 1, pp. 182–193, Jan. 2018.

[29] A. T. Procopiou, J. Quirós-Tortós, and L. F. Ochoa, "HPC-based probabilistic analysis of LV networks with EVs: Impacts and control," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1479–1487, May 2017.

[30] Y. Ma, D. Azuatalam, T. Power, A. C. Chapman, and G. Verbič, "A novel probabilistic framework to study the impact of photovoltaic-battery systems on low-voltage distribution networks," *Appl. Energy*, vol. 254, Nov. 2019, Art. no. 113669.

[31] UK Power Networks. (Feb. 2015). *SmartMeter Energy Consumption Data in London Households*. [Online]. Available: https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households.

AMY ANDERSON received the B.Eng. degree in electronics and electrical engineering from the University of Glasgow, U.K., in 2012. She is currently pursuing the Ph.D. degree with the Department of Electronic and Electrical Engineering, University of Strathclyde. Her research interests include smart grids, energy management in power systems, energy storage, distributed generation, and utilizing machine learning techniques to support the transition of power networks to low-carbon operation.

BRUCE STEPHEN (Senior Member, IEEE) received the B.Sc. degree in aeronautical engineering from the University of Glasgow, U.K., in 1997, and the M.Sc. degree in computer science and the Ph.D. degree from the University of Strathclyde, U.K., in 1998 and 2005, respectively. He holds a Senior Research Fellow position with the Institute for Energy and Environment, University of Strathclyde. His research interests include power system condition monitoring, renewable integration, characterizing low-voltage network behavior, and characterization of demand in electrical distribution networks. He is a Chartered Engineer and a Fellow of the Higher Education Academy.

RORY TELFORD received the B.Eng., M.Sc., and Ph.D. degrees in electronic and electrical engineering from the University of Strathclyde, Glasgow, U.K., in 2008, 2011, and 2017, respectively. He is currently a Research Associate with the Institute for Energy and Environment, University of Strathclyde. His research interests include application of AI techniques, data-driven fault diagnostics, and power system modelling and analysis.

STEPHEN MCARTHUR (Fellow, IEEE) received the B.Eng. (Hons.) and Ph.D. degrees from the University of Strathclyde, Glasgow, U.K., in 1992 and 1996, respectively. He is currently a Deputy Associate Principal for Research, Knowledge Exchange and Innovation and a Co-Director with the Institute for Energy and Environment, University of Strathclyde. His research interests include intelligent system applications in power engineering, covering condition monitoring, diagnostics and prognostics, active network management, and wider smart grid applications.