

## Article

# A Cross-Country Model for End-Use Specific Aggregated Household Load Profiles

Marlon Schlemminger <sup>1,\*</sup> , Raphael Niepelt <sup>1,2</sup> and Rolf Brendel <sup>1,2</sup>

<sup>1</sup> Department Solar Energy, Leibniz University Hannover, Appelstr. 2, 30167 Hannover, Germany; niepelt@isfh.de (R.N.); brendel@isfh.de (R.B.)

<sup>2</sup> Institute for Solar Energy Research Hamelin (ISFH), Am Ohrberg 1, 31860 Emmerthal, Germany

\* Correspondence: schlemminger@solar.uni-hannover.de

**Abstract:** End-use specific residential electricity load profiles are of interest for energy system modelling that requires future load curves or demand-side management. We present a model that is applicable across countries to predict consumption on a regional and national scale, using openly available data. The model uses neural networks (NNs) to correlate measured consumption from one country (United Kingdom) with weather data and daily profiles of a mix of human activity and device specific power profiles. We then use region-specific weather data and time-use surveys as input for the trained NNs to predict unscaled electric load profiles. The total power profile consists of the end-use household load profiles scaled with real consumption. We compare the model's results with measured and independently simulated profiles of various European countries. The NNs achieve a mean absolute error compared with the average load of 6.5 to 33% for the test set. For Germany, the standard deviation between the simulation, the standard load profile H0, and measurements from the University of Applied Sciences Berlin is 26.5%. Our approach reduces the amount of input data required compared with existing models for modelling region-specific electricity load profiles considering end-uses and seasonality based on weather parameters. Hourly load profiles for 29 European countries based on four historical weather years are distributed under an open license.

**Keywords:** energy system modelling; household load profile; neural network; end-uses; consumer behavior; cross-country; open data



**Citation:** Schlemminger, M.; Niepelt, R.; Brendel, R. A Cross-Country Model for End-Use Specific Aggregated Household Load Profiles. *Energies* **2021**, *14*, 2167. <https://doi.org/10.3390/en14082167>

Academic Editor: Gianfranco Chicco

Received: 10 March 2021

Accepted: 11 April 2021

Published: 13 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Electrical load profiles are of significant interest to estimate time-dependent energy consumption if real measurement data are scarce or predictions for the future are required. This includes profiles on different spatial scopes such as single buildings, municipalities, or whole countries and for different sectors such as the residential, the commercial, and the industrial sector. The residential sector currently contributes 25% to the final energy consumption in the European Union (EU) [1].

End-use specific profiles are also of interest for optimizing future demand-side management (DSM), because a change of end-uses is likely with the introduction of new technologies and only particular end-uses are suitable for DSM. The adequate prediction of future household load profiles is thus of importance for energy system modelling. Owing to the strong coupling between the European countries, a suitable profile generation method should be capable of considering country-specific differences in consumer behavior, and thus household load profiles.

Most existing models for residential load profile prediction can be grouped into two categories: bottom-up and top-down models. Bottom-up models generally model the household in detail by looking at appliances and occupants and aggregate them to build the total consumption profile. Generating profiles on a regional or national level with bottom-up models then requires an extrapolation of the results for single buildings using socio-

economic or building stock data. In contrast, top-down models use aggregated data and correlate them to macro-variables such as weather conditions or consumer behavior [2,3].

The literature shows extensive studies using both methods. An early stochastic bottom-up model comes from Richardson [4], who implemented a simulation for single households in 1-min resolution. The authors modelled the probability for switching on appliances based on daily activity profiles from the United Kingdom (UK). The activity profiles are generated from time-use surveys (TUSs) [5]. Such a comprehensive TUS is comparable to diaries that characterize how people spend their time. TUS-based models require country-specific TUS data. Appliances are assigned to households based on national ownership data and each appliance assumes a specific load profile once switched on. McKenna and Thomson [6] extended the model by a thermal module six years later.

Ardakanian et al. [7] recognized the main weakness of Richardson's model—its requirement for many input datasets such as the appliances in each home and their respective load profile, the number of occupants, and their behavior. Therefore, the authors developed Markov chains based on real measurements from 20 buildings over a timespan of four months, which reduced the amount of input data. They were able to predict the residential load on the level of local power transformers, but the model is not able to represent appliances or to consider the influence of weather data. This conflict between a detailed representation of the load curve requiring much input data and minimizing input data while keeping a certain level of detail is still ongoing in the literature.

Multiple detailed bottom-up models exist. Gottwalt et al. [8] scheduled appliances based on stochastic start times, runtimes, and consumer occupancy. Tsagarakis et al. [9] also used TUS data from the United Kingdom and constructed Markov chains to develop activity profiles of single inhabitants of residential buildings. These authors combined the activity profiles with ownership databases and appliance profiles and added the use of ambient weather parameters to influence the generated load profiles. A similar approach comes from Lombardi et al. [10], who developed an open-source stochastic bottom-up model on the appliance level for multiple sectors, which was primarily designed for remote areas, but is able to predict larger regions as well. Fischer et al. [11] used stochastic appliance modelling as well. They determine the appliance ownership for a household from socio-economic data such as the number of occupants and working patterns. One of the most sophisticated load profile models is that by Pflugradt [12]. The author chose to model desires of people such as being hungry or having to wash laundry and used time series of the probability to have a certain desire to model consumer behavior. The model is then able to construct load curves up to a spatial resolution of quarters. A more recent model comes from Gao et al. [13]. The authors predicted the appliance consumption of households from evaluating similar historic days. Factors that are considered for identifying similar days are weather parameters, weekdays, employment status, major events such as holidays, and the time gap between the prediction and the historical date.

Top-down models are widespread as well. Singh and Yassine [14] used an unsupervised clustering analysis on large measurement datasets to extract frequent patterns and trained a Bayesian network on their dataset. The network then used seven information streams such as the hour of the day, the weekday, or the season, as well as dependencies between appliances to predict load profiles. Ge et al. [15] aimed at a model with significantly lower data input and fitted multiple Gaussian functions on measured profiles from 200 households. The only parameters they considered to predict the load curve were the number of homes, the number of bedrooms per home, and the number of occupants per home. Rahman et al. [16] used recurrent neural networks to correlate weather variables such as the solar irradiation and schedule-dependent variables such as the hour of the day to the aggregated electricity consumption of residential buildings. Arens et al. [17] extracted typical daily load profiles from measurements of a single household using neural networks. Anvari et al. [18] used empirical mode decomposition to extract both short-term and long-term trends from measurements over four consecutive weeks. A widespread method in Germany to predict the consumption of households is using the standard load

profile (SLP) H0 [19], which is a formula with fixed parameters that takes into account the weekday and the season. Stokes [20] extracted average daily and seasonal profiles for a number of devices from a dataset measured across 1200 households in the United Kingdom during the 1990s. The author provided formulas with fixed parameters per appliance that describe both the daily and seasonal trends.

Most of the above-mentioned models are capable of predicting the consumption on the level of single persons or households. The selection of an appropriate model is influenced by its functionality, computational expenditure, and the validity of the results. However, in the context of energy system analysis and technology impact assessment, a desirable combination of features exists that none of the models so far provides:

- Seasonality based on weather data: Modern energy system models focus on the transition to renewable, weather-dependent energies. These models naturally incorporate fluctuating renewable electricity supply. Using fluctuating profiles of both supply and demand based on the same datasets allows researchers to analyze their interaction. Only six load profile models from above are able to use weather data for at least some appliances [4,9,11–13,16].
- Consider end-uses/appliances: Most bottom-up models naturally provide appliance-specific load curves. However, most top-down models do not consider appliances, but are only built for aggregated load curves. Energy system models often consider time horizons until 2050 or later and use assumptions on the development of electric consumption per end-use (e.g., decreasing consumption for space heating due to better building insulation). Out of the models analyzed, two top-down models [14,20] and eight bottom-up models [4,8–13] were end-use specific.
- Cross-country: Energy system models spanning multiple countries require country-specific load profiles. All models besides [10] were either not able to extrapolate their results to other countries or did not demonstrate it.
- We define six categories of input data that are necessary to extrapolate a model to larger regions: Formulas with fixed coefficients, time-use surveys, appliance ownership, socio-economic data, weather data, and load profile measurements. The effort to use and acquire the data differs between the categories, which means that the amount of data categories alone does not define the total input data requirements. Formulas with fixed coefficients require little to no effort, because they are provided by the author. Extrapolating bottom-up models from single households to regions or countries often requires large appliance ownership and socio-economic data and makes the application computationally expensive. The complexity of load profile measurements depends on its level of detail such as the number of households and the temporal resolution. Out of the models analyzed, we only estimated a small demand for input data for five top-down models [7,15,18–20] for an extrapolation of the model to a national level.

Additionally, providing the model itself or the input and output data under an open license helps with reusability. We found seven models [6,9,10,12,18–20] doing this.

In Table 1, we present an overview of the reviewed models. None of them were able to meet all of the aforementioned combination of features. In this paper, we develop a model that meets those criteria at manageable computational expenditure and input data requirements. We only use openly available information. Our approach is capable of generating load curves for specific regions in Europe. We use a neural network (NN) to correlate the extensively measured consumption in the United Kingdom from [20] with region-specific input data that are also available for other countries. This allows the derivation of load profiles for any region where this kind of data are available without the need for measured consumption data. The generated load profiles shall be used in energy system analysis and enable researchers to model future shifts in consumption. This classifies our model as long-term load forecasting according to [21], although it shows characteristics of short-term forecasting such as the hourly resolution as well. Region-specific input data we use include the following:

- Residents' activities based on TUS data. TUS data result from surveys that extract the time spent in various activities over the day from diaries.
- Typical daily power consumption profiles of devices that depend little on region and human activities (e.g., fridges).
- The outside temperature and global irradiation.
- The German standard load profile gas (SLPG) to include energy consumption for heating.

The key contributions of this paper are as follows:

- Correlate empirical consumption data with weather conditions, device power consumption profiles, and residents' activities.
- Use the correlations found to develop regionalized hourly end-use specific residential load profiles for 29 European countries for four historical weather years.
- Distribute the resulting load profiles under an open license.

**Table 1.** Literature overview of related models. The models are sorted by year. The table provides information about the criteria necessary for further use in energy system models: Is the model able to be applied in multiple countries? Does weather data influence the seasonality? Is it end-use specific? Are input or output data or the model itself available under an open license? How much data are required to model larger regions? We define six categories of data: Formulas with fixed coefficients (F), time-use surveys (TUSs), appliance ownership (A), socio-economic data (SE), weather data (W), and load profile measurements (M). Please note that the amount of data categories alone does not define the total input data requirements, because the complexity between categories differs.

Authors	Year	Method	Data Requirements	Cross-Country	Seasonality Based on Weather Data	End-Use Specific	Open License
Meier et al. [19]	1999	Top-Down	Small (F)	No	No	No	Yes
Stokes [20]	2005	Top-Down	Small (F)	No	No	Yes	Yes
Richardson et al. [4]	2010	Bottom-Up	Large (TUS, A, SE, W)	Possibly	Limited (Only lighting)	Yes	No
Ardakanian et al. [7]	2011	Hybrid	Small (M)	No	No	No	No
Gottwalt et al. [8]	2011	Bottom-Up	Large (A, TUS, M)	Possibly	No	Yes	No
Collin et al. [9]	2014	Bottom-Up	Large (TUS, A, SE, W)	Possibly	Mostly (lighting and heat)	Yes	Yes
Fischer et al. [11]	2015	Bottom-Up	Large (A, SE, M)	No	Mostly (lighting and heat pump)	Yes	No
Ge et al. [15]	2016	Top-Down	Small (SE, M)	Possibly	No	No	No
McKenna et al. [6]	2016	Bottom-Up	Large (TUS, A, SE, W)	Possibly	Mostly (lighting and heat)	Yes	Yes
Pflugradt [12]	2016	Bottom-Up	Large (A, SE)	Possibly	Yes	Yes	Yes
Singh & Yassine [14]	2018	Top-Down	Large (M)	No	No	Yes	No
Gao et al. [13]	2018	Bottom-Up	Large (A, SE, M, W)	No	Yes	Yes	No
Rahman et al. [16]	2018	Top-Down	Large (M, W, A)	No	Yes	No	No
Arens et al. [17]	2018	Top-Down	Large (M, SE)	No	No	No	No
Lombardi et al. [10]	2019	Bottom-Up	Large (SE, A)	Yes	No	Yes	Yes
Anvari et al. [18]	2020	Top-Down	Small (M)	Possibly	No	No	In parts
This work	2021	Hybrid	Small (TUS, W, M)	Yes	Yes	Yes	In parts

The paper is organized as follows. Section 2 presents our methodology, divided into the pre-processing of input data and training and prediction of the neural networks in Sections 2.1 and 2.2 and the implementation of space heating and hot water in Section 2.3.

Section 3 presents, validates, and discusses the results and Section 4 gives a summary and conclusion.

## 2. Materials and Methods

The total electrical load curve of a multitude of households consists of different end-use load curves. In order to regionalize the modelled total curve using region-specific TUS data, we also have to consider a region-specific share of consumption for the different load contributions. To model the impact of residents' behavior on the load curve, we have to align deterministic activity data and statistical consumption data. For that reason, we classify all modelled load contributions according to the German Working Group on Energy Balances (AGEB) [22] into the Eurostat-compatible end-use categories:

- Space Heating;
- Hot Water;
- Process Heat;
- Lighting;
- Cooling;
- Mechanical Energy;
- Information and Communications Technology (ICT).

The AGEB originally defines two separate end-uses for cooling (air conditioning and process cooling). We merge them into a single end-use because of the currently limited relevance of air conditioning in Europe. According to Eurostat, space cooling has a share of less than 1% on the final energy consumption in households for all European countries except for Greece and Malta [23].

With the load curves for each end-use and region-specific data on time use, temperature, and solar irradiation, we are able to build a regionalized total electrical load curve by scaling each end-use with its respective annual consumption. Our synthetic load curves thus combine aspects of a bottom-up approach via modelling of various end-uses with aspects of a top-down approach via scaling of the various loads. For validation, we compare the results of our approach to measurements of total electrical loads, which are more common than end-use specific measurements. The lowest spatial resolution that we consider for our methodology is NUTS3 regions. The NUTS classification (nomenclature of territorial units for statistics) is a hierarchical system for dividing the territory in Europe for regional statistics, and there are currently 1348 NUTS3 regions within the EU and the United Kingdom.

### 2.1. Neural Network-Assisted Linking of Household Load and Human Behaviour

As a starting point, we use Stokes' work [20] on load profiles extracted from measurement data across 1200 households in the United Kingdom during the 1990s. Stokes showed that, for most devices in a household, the yearly power consumption profile can be described by a linear combination of daily power consumption profile  $S_{hh,device}$  and a sinusoidal seasonal trend. The resulting time series  $d_{hh,device}$  describes the devices' contribution to the end-use load profile with half-hourly ( $hh$ ) values over a year. We consider the devices cooking, lighting, dishwasher, washing machine, fridge-freezer, and miscellaneous for our model. The majority of real devices that are assigned to miscellaneous by Stokes consist of ICT such as television or home office equipment.

$S_{hh,device}$  (in kW) is a mix of device-typical weather-independent profiles and the impact of human behavior with varying shares of both depending on the device. For some devices, it is different for working days (Monday to Friday), Saturdays, and Sundays.

For each device, Stokes then identified a single weather parameter that controls the yearly trend and fitted a sine wave function to this trend. Irradiation is the weather parameter for lighting and the outside temperature is used for every other device.

The power consumption of a device  $d_{hh,device}$  is defined as a single sine function:

$$d_{hh, device} = S_{hh,device} * \sin(n + x_{day, device}) + k_{hh, device} \quad (1)$$



and as the sum of two sine functions in the case of lighting:

$$d_{hh, lighting} = S_{1, hh, lighting} * \sin(n + x_{1, day, lighting}) + S_{2, hh, lighting} * \sin(n + x_{2, day, lighting}) + k_{hh, lighting} \quad (2)$$

where  $n$  (in rad; 1 year  $\hat{=}$   $2\pi$ ) describes the day of the year moved by a phase angle  $x$ ,  $x_1$ , and  $x_2$  (in rad) and  $k$  (in kW) is a constant depending on the device. Further explanation of the methodology and tables with values regarding  $S$ ,  $k$ , and  $x$  can be found in Chapter 4 and Appendix B of Stokes' work [20].

While these fixed daily profiles with a seasonal sine function are sufficient to model the United Kingdom [20] (p. 96), they were not meant to offer an appropriate parameterization to account for different weather and human behavior in other countries. In order to generate country-specific profiles, we quantify the dependence of the total profile  $d_{hh, device}$  for each device on (i) the average U.K. weather data  $w_{avg}$  denoting the irradiation or the temperature and (ii) Stokes' daily device usage profile  $S_{hh, device}$ . If this dependence is known, we can replace the inputs with country-specific weather data  $w'$  and country- and device-specific, but weather-independent device usage profiles  $S'_{device}$  to generate country-specific individualized device profiles  $d'_{device}$ .

#### Training and Network Characteristics

Neural networks (NNs) are commonly applied for modelling unknown and possibly non-linear relationships [24,25]. Hence, NNs are an obvious choice when it comes to studying the aforementioned dependence of the total profile  $d_{hh, device}$  on the weather data  $w_{avg}$  and Stokes' daily device usage profile  $S_{hh, device}$ .

NNs are typically trained on large datasets with millions of entries and hundreds of input parameters, which in return requires the network to consist of high numbers of neurons. In our case, we only have a training dataset of 8760 h of which we separate random 20% as a test set and keep the remaining 80% as the training set. We have two input parameters:  $w_{avg}$  denoting the irradiation or the temperature, and thus the yearly trend and the daily usage profile  $S_{hh, device}$ . For  $w$ , we use the reconstruction dataset coastDat-2 COSMO-CLM [26], which we remap to a grid of weather cells of approximately  $30 \times 25$  km, covering the whole European continent. We use hourly averages of the datasets global irradiation and temperature of all weather cells intersecting with the United Kingdom over the timespan from 1994 to 2014 to build the training dataset  $w_{avg}$ .

The second part of the training input vector consists of a repetition of the daily usage profile  $S_{hh, device}$  that is specific for working days (Monday to Friday), Saturdays, and Sundays in case of the human behavior-dependent devices, but has only one weekday profile for constantly running devices like fridges (see Table 2). We train the model to predict the season-dependent total load profile  $d_{hh, device}$  ( $w_{avg}$ ,  $S_{norm, device}$ ). To prevent the NN from overfitting, we use two hidden layers with only 16 neurons each and a strong regularization parameter of 0.01 [25]. Both hidden layers use the ReLU activation function, as it showed the lowest error out of three different functions on our problem, and minimize the mean squared error (MSE). In a dataset with  $n$  samples, where  $Y$  are the observed values and  $Y'$  are the predicted values, the MSE is defined as

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - Y'_i)^2 \quad (3)$$

**Table 2.** Modelling end-uses: The columns contain six devices and the rows show the dependence of the end-use specific power consumption profiles on different types of parameters. MAE, mean absolute error; ICT, information and communications technology.

Parameter	Lighting	Cooking	Miscellaneous	Fridge-Freezers	Dishwasher	Washing Machine
No. of weekday profiles	3	3	3	1	1	3
Weather parameter	Global irradiation	Temperature	Temperature	Temperature	Temperature	Temperature
Test set MAE/Mean	33.0%	30.1%	8.4%	6.5%	7.3%	13.9%
Source of basic profile $S'_{device}$	TUS	TUS	Stokes	Stokes	Stokes	Stokes
Weighting factor for end-use $wde$	100% of Lighting	55% of Process Heat	100% of ICT	100% of Cooling	23% of Process Heat	22% of Process Heat, 100% of Mechanical Energy

This makes our NNs a supervised learning approach, because we classify input and desired output data for the model in advance. We evaluate the performance of our model by the relation of the mean absolute error to the mean load (MAE/Mean) on our test set, with the MAE defined as

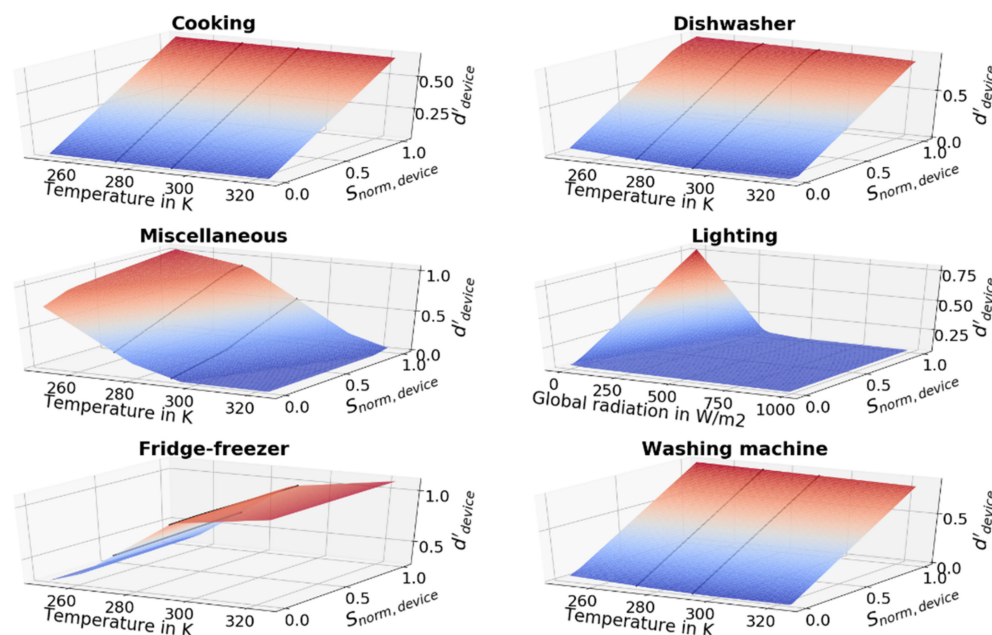
$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - Y'_i| \quad (4)$$

Another common error metric for neural networks is the mean absolute percentage error (MAPE), defined as

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - Y'_i}{Y_i} \right| \quad (5)$$

However, the MAPE has shortcomings that prevented us from using it. It is obvious from Equation (5) that the MAPE is not defined for observed values  $Y$  that are zero. Zeros are included in our training dataset though. Dropping zeros from the MAPE calculation is not an option, because this would exclude these time steps from the error metric. Therefore, we use the MAE/Mean in line with [27], which ranges from 6.5% to 33% in our test set for the different devices. We show examples of the training datasets in Figures A1–A4 in Appendix A.

Figure 1 shows the relationship between input and output quantified by the NNs with the weather parameter  $w$  on the X-axis, the normalized device usage  $S_{norm,device}$  on the Y-axis, and the normalized load  $d'_{device}$  on the Z-axis. The dependence on the daily usage profile  $S_{norm,device}$  is stronger for the devices cooking, dishwasher, and washing machine, while the fridge-freezer depends more on the ambient temperature. The miscellaneous and lighting profiles show a connection to both, while lighting is strongly influenced by the absence and only to a small degree by the intensity of daylight. A desired side effect of the strong regularization and small network design is that the modeled relations in Figure 1 do not have sharp edges or local extrema. This may allow us to apply the model to temperature ranges that the training dataset did not cover. Generally, the design of NNs does not allow such extrapolation, because their output can be random outside of their training range. Owing to the strong averaging of input data over multiple years and weather cells, we only train the network on temperatures between 273 and 292 K. However, we expect that the simplicity of the model allows an extrapolation to temperatures outside the training range.



**Figure 1.** Output of the trained neural networks  $d_{device}$  showing the dependence of the total profile  $d_{hh,device}$  on the weather data and Stokes' daily device profile  $S_{hh,device}$ . The training range is shown as dotted black lines.

## 2.2. Prediction of Household Load Profiles Based on Weather Data and Human Behaviour

After the training phase, we apply region-specific input parameters and let the trained network predict the load profile. For this, we exchange the original weather data  $w_{avg}$  by data  $w'$  of individual years and the respective regions to eliminate both temporal and spatial harmonization. In addition, we substitute Stokes' daily usage profiles  $S_{norm,device}$  for some devices by region-specific activity profiles  $TUS_{device}$  to generate the basic profiles  $S'_{device}$ . The basic profiles take into account the results of TUS for lighting and cooking to include differences in human behavior between various countries:

- The basic profile  $TUS_{lighting}$  for lighting represents the share of people that are home and awake.
- The basic profile  $TUS_{cooking}$  for cooking represents the share of people cooking.
- As shown in Figure 1, process cooling has little dependence on the daily profile and does not need an individualized profile. Therefore, we use Stokes' device profile of fridge-freezers  $S_{norm,fridge-freezer}$ .
- Dishwasher and washing machine would use a combination of human behavior and a device profile, because the activity of starting the machine itself only takes minutes, but entails a process of 1–2 h of electricity consumption. This requires precise information about the device used, which the TUS data do not deliver: they only specify a general category of 'housework'. Therefore, we use Stokes' profiles  $S_{norm,washing machine}$  and  $S_{norm,dishwasher}$ .
- ICT should use TUS-based data, too. However, the age of the available data surveys prevents a reasonable consideration: The data for four out of the seven countries are more than 30 years old, when ICT consisted mainly of a few hours of television in the evening. A comparison between TUS-based and Stokes-based ICT profiles showed that Stokes' profile  $S_{norm,miscellaneous}$  performs better across countries.

TUS data are available in harmonized formats across 30 countries on a central database, but only Bulgaria, Germany, Spain, France, Hungary, and Italy show adequate data quality [28]. We present an excerpt of German TUS data in Table A1 in the Appendix A. Data in the same format are also available for the United Kingdom [29]. Additional proprietary TUS data are available, e.g., by the German federal statistics office, but we decided to only work with open data for this paper. We analyze these TUS data to develop shares for each



hour of the day of how many persons of a group perform a certain activity. We call the result activity profiles  $TUS_{device}$ . The diary entries contain information about the month, weekday, sex, and age. From this, we further differentiate the data by weekday (Monday to Friday, Saturday, Sunday/public holiday) and season to achieve an extra level of detail. We apply weighting by sex and age, because the surveys are not representative of the country's average. We do this by grouping the interviewees into groups of sex and age (younger than 20 years; 20 to 29 years; 30 to 49 years; 50 to 64 years; older than 64 years), building activity profiles for each group, and weighting each group via the census 2011 population data available on Eurostat [30]. We show an example of an activity profile for cooking in Figure A5 in the Appendix A.

Feeding  $S'_{device}$  and  $w'$  to the trained NNs generates hourly resolved load profiles  $d'_{device}$  for each device. We assign every considered device load to the AGEb end-use classification to achieve end-use profiles  $d_{end-use}$  that are compatible with common energy system analysis tools. This is straightforward for most considered devices (see Table 2). The only exception is process heat, where we assume a weighting factor  $wde_{device,process\_heat}$  of the devices cooking, washing machine, and dishwasher that are not separately listed in the AGEb data according to their annual consumption in Germany in 2015 [31].

$$d_{end-use} = \sum_{devices} d'_{device} * wde_{device, end-use} \quad (6)$$

Table 2 summarizes some key information about our model and the differences between devices. We use the TUS data from neighboring countries to extend the coverage of our model to countries with missing free TUS data: Bulgaria and Hungary to cover Eastern Europe, France and Italy to cover Central-Western Europe, Portugal and Ireland are assigned to Spain and the United Kingdom and Germany to cover Central-Eastern Europe and Scandinavia. With these adaptations, we are able to model regionalized household load profiles for EU-27 (excluding Cyprus), Norway, Switzerland, and the United Kingdom.

### 2.3. Implementing Space Heating and Hot Water

Stokes' data are not convenient for the end-uses space heating and hot water, because they only covered price-driven profiles for electric heating using off-peak electricity tariffs, whereas we aim for price-independent load profiles. With the German standard load profile gas (SLPG), a scientifically verified [32] and industrially applied method exists to predict the consumption of natural gas for the end-uses space heating and hot water. Furthermore, the SLPG uses the ambient temperature to determine the consumption, which allows us to create individual profiles for different regions. Originally developed in 2003 [20], there are two guidelines by the German Association of Energy and Water Industries (BDEW) from 2018 [33] and its predecessor organization BGW from 2006 [34] improving the approach. We combine both approaches similar to [35]: the partially linearized sigmoid function of the BDEW, which improves the profiles' performance in extreme temperature ranges, with the hour-scaling values and the option to choose parameter values based on the average wind speed of the BGW. The hourly resolved profile is defined as

$$Q_h = hF_h \quad (7)$$

with  $h$  being the daily consumption factor (no unit) and  $F_h$  the hour scaling values (no unit).  $F_h$  is given in tabular form and depends on the building type, the hour of the day, and the outside temperature in steps of 5 °C from −15 °C to +25 °C. The partially linearized sigmoid function to determine  $h$  is defined as

$$h = \left[ \frac{A}{1 + \left( \frac{B - \theta}{\theta - \theta_0} \right)^C} + D \right] + \max \left( m_h \frac{\theta}{\theta_0} + b_h, m_w \frac{\theta}{\theta_0} + b_w \right) \quad (8)$$

with  $\theta_0 = 40\text{ }^\circ\text{C}$ ,  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $m_h$ ,  $b_h$ ,  $m_w$ , and  $b_w$  (all without unit) determined by the building type and wind site. The subscripts  $h$  and  $w$  denote heating and warm water, respectively.  $\theta$  is defined as

$$\theta = \frac{\theta_D + 0.5\theta_{D-1} + 0.25\theta_{D-2} + 0.125\theta_{D-3}}{1 + 0.5 + 0.25 + 0.125} \quad (9)$$

with  $\theta_{D-n}$  being the average temperature of the day  $n$  days before. The BDEW gives values for the parameters depending on building type and wind location. For building types, we use values for single and multi-family houses and weight them by a ratio of 50:50. The wind location has an influence on the temperature dependency. The higher the average wind speed, the lower the share of gas for cooking. We classify NUTS regions into two subclasses with a threshold of 4.4 m/s to determine the share of cooking gas.

The BDEW assumes that consumption for hot water is close to constant over the year, which Stokes confirms as well. This allows us to model its load curve as the sum of the temperature-independent parameter  $D$  and the linearized part for water heating, which is larger than the linearized part of space heating for ambient temperatures  $\geq 15\text{ }^\circ\text{C}$ .

$$h_{hot\ water} = D + m_w \max\left(\frac{\theta}{^\circ\text{C}}, 15\text{ }^\circ\text{C}\right) + b_w \quad (10)$$

The hourly scaling values of hot water  $F_{h, hot\ water}$  are given by the hourly scaling values of the highest ambient temperature range  $\geq 25\text{ }^\circ\text{C}$ . The difference between the total curve and the hot water curve is the load profile for space heating.

The SLPG does not account for the conversion of final energy to useful energy, because conventional gas boilers have an almost constant efficiency over the year. The coefficient of performance (COP) of a heat pump, which we assume to supply the total electrical space heating and hot water demand, does however vary with the outside temperature. This effect impacts the shape of the load curve. We model hourly COP values by a parameterized empirical formula given in [36] for air source and ground source heat pumps (ASHPs and GSHPs):

$$\begin{aligned} COP_{ASHP} &= 6.81 - 0.121 \frac{\Delta T}{^\circ\text{C}} + 0.00063 \left(\frac{\Delta T}{^\circ\text{C}}\right)^2 \\ COP_{GSHP} &= 8.77 - 0.15 \frac{\Delta T}{^\circ\text{C}} + 0.000734 \left(\frac{\Delta T}{^\circ\text{C}}\right)^2 \end{aligned} \quad (11)$$

with  $\Delta T$  being the difference between the source and the heat sink temperature. We use a constant of  $10\text{ }^\circ\text{C}$  as the source for GSHPs and the outside air temperature for ASHPs. Concerning the heat sink temperature, we assume a ratio  $s_{58} = 67\%$  for  $58\text{ }^\circ\text{C}$  and  $s_{40} = 33\%$  for  $40\text{ }^\circ\text{C}$ . We mix ASHPs and GSHPs by their market share in Germany over the last years, which is  $f_{ASHP} = 72\%$  and  $f_{GSHP} = 28\%$  [37]. The COP of the combined heat pump is the inverse average:

$$\begin{aligned} \frac{1}{COP} &= f_{ASHP} \left\{ s_{58} \frac{1}{COP_{ASHP}(58\text{ }^\circ\text{C}-T_L)} + s_{40} \frac{1}{COP_{ASHP}(40\text{ }^\circ\text{C}-T_L)} \right\} \\ &+ f_{GSHP} \left\{ s_{58} \frac{1}{COP_{GSHP}(58\text{ }^\circ\text{C}-10\text{ }^\circ\text{C})} + s_{40} \frac{1}{COP_{GSHP}(40\text{ }^\circ\text{C}-10\text{ }^\circ\text{C})} \right\} \end{aligned} \quad (12)$$

This results in COPs for Germany between 2.1 and 5 and a seasonal performance factor of 2.8.

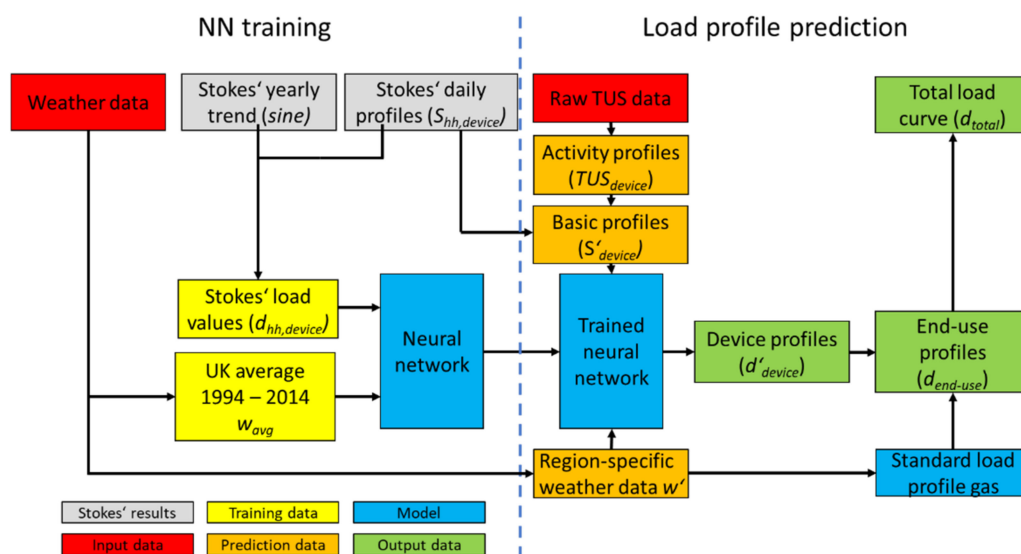
Using profiles for the consumption of gas to predict temporal electricity consumption implies that none of them include a time shift of more than an hour between the time of final energy consumption and useful energy consumption caused by the replacement of a gas boiler with a heat pump.

#### 2.4. Annual Consumption

We use annual electricity consumption specific per NUTS3 region and end-use in order to scale the load shapes developed in Sections 2.2 and 2.3 to present or future consumption. This allows us to compare the results of our approach to measurements of total electrical

loads, which are more common than end-use specific measurements. Eurostat provides present country aggregated sums [38] of electricity consumption in households per end-use, which need to be disaggregated to NUTS3 regions. Spatial disaggregation is performed similar to [39] or [40] via the census 2011 population data available on Eurostat [30], assuming a direct correlation between population size and electricity consumption. We present the consumption per end-use and country used in this work in Table A2 in the Appendix A.

Figure 2 summarizes the whole structure of our model. Depending on the end-use, we generate profiles from either a trained neural network or the standard load profile gas. The neural networks are trained on measured electricity consumption and average weather data from the United Kingdom and use region-specific activity profiles from time-use surveys and weather data to make predictions.



**Figure 2.** Structure of the model. The left side shows the training of the neural networks and the right side shows how we use the trained neural networks to predict the end-use load profiles. We scale each end-use profile by annual consumption to achieve the total load curve. NN, neural network; TUS, time-use survey.

### 3. Results and Discussion

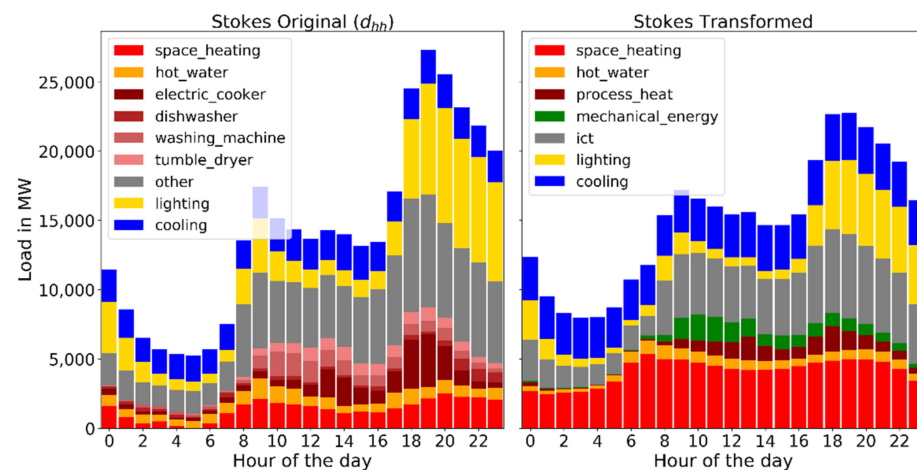
In this section, we present results on a country-aggregated level for the weather year 2012. Available data disaggregated into end-uses for comparison with our results are rare, which is why we can predominantly only make plausibility checks. Load profiles of different origins are always scaled to the same annual consumption to ensure comparability and all timestamps are translated to Central European (Summer) Time for comparison. We evaluated the approach also for the weather years 2011 to 2014 and all results are similar to 2012.

Figure 3 compares Stokes' original electrical load curve to the transformation to AGEB end-uses for the average winter weekday in the United Kingdom and shows values grouped by the hour of the day. The transformation includes three major steps:

1. Replacing the load curves of space heating and hot water by the SLPG. This smoothens both end-uses by shifting consumption from the morning and afternoon peaks to the night and noon troughs.
2. Merging the electric cooker and the wet appliances into process heat and mechanical energy. This exchanges their evening peak with the noon peak.
3. Rescaling each end-use with its respective Eurostat consumption, which has an impact on the contribution of each end-use to the total load curve. Cooling and most notably space heating grow in relevance, while especially lighting has less total consumption. Explanations for these shifts are either that the original measurements were not

representative for the average U.K. household or that changes in behavior occurred during the last 20 years.

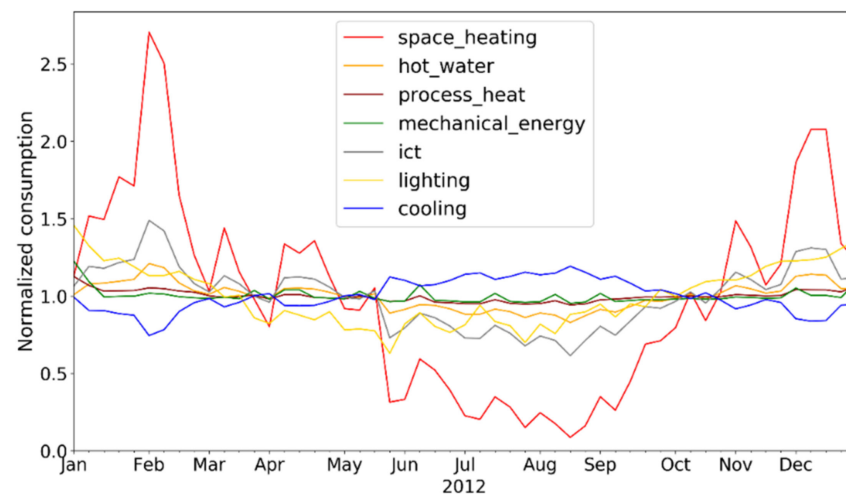
The following transition from the transformed load curve according to Stokes to our simulation via neural networks has comparably small effects. Some consumption shifts from the evening peak to the morning peak, most notably at the end-uses lighting and ICT, but the overall shape remains unchanged with a high consistency ( $R^2 = 0.94$ ). We show a visualization of that transition in Figure A6 of the Appendix A.



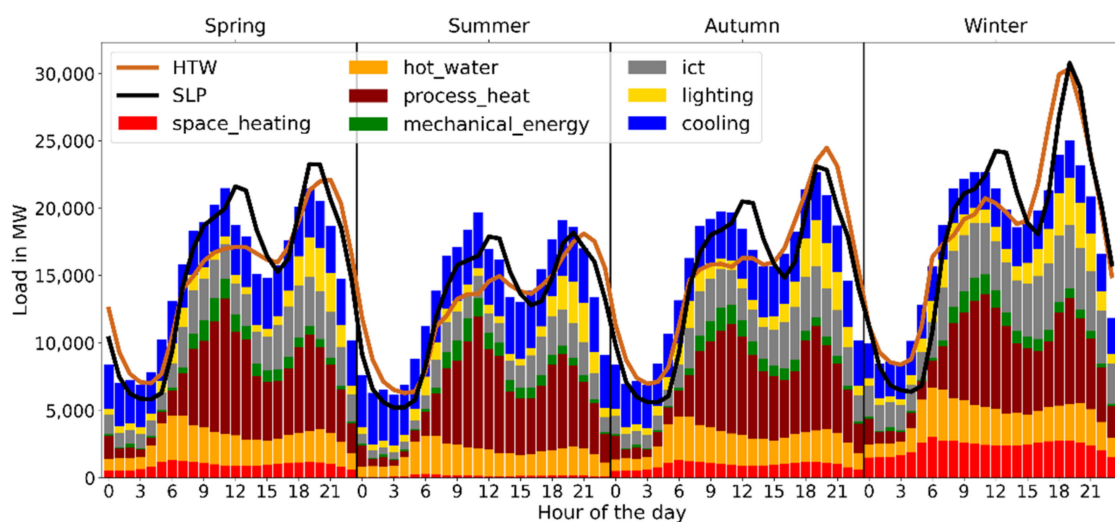
**Figure 3.** Stokes' original load curve (left) vs. transformation to German Working Group on Energy Balances (AGEB) end-uses (right) for the average winter weekday (December–February) in the United Kingdom. The sum under both graphs is identical, but the contributions of some end-uses to the total load curve differ. Possible explanations for the shifts of consumption between end-uses might be that consumers' behavior changed during the last 20 years or that the original measurements (left) were not representative for the whole United Kingdom.

Figure 4 shows the seasonality coming from our simulation approach plotted together with the seasonality coming from the SLPG for space heating and hot water, normalized to a mean of one for better visualization. The plot shows the weekly mean consumption, and is thus not designed to show the load profile within a week. Although the seasonality of space heating is most dominant, one should not neglect the one of the end-uses lighting, ICT, and cooling. Different shares of each end-use on the total consumption between countries will lead to different seasonality as well. Furthermore, the advantage of using consumption profiles directly correlated to fluctuating weather data instead of smoothed profiles becomes apparent. Although a general yearly trend is visible, peaks (e.g., during late January and early February) occur and considerably deviate from the yearly trend. Large parts of Europe experienced an extreme cold wave from the end of January to mid-February 2012, which we are able to represent in the load profile of each end-use.

Figure 5 shows the simulation results for Germany and compares them to two other datasets for validation. The first one is the SLP H0, which is commonly used by German utilities to estimate the load of non-metered customers [19], and the second one is a dataset measured by the University of Applied Sciences (HTW) Berlin in 74 single-family households in spatial proximity to one another [41]. The figure shows the average load grouped by season and the hour of the day. Each seasons consists of three months starting with spring in March.



**Figure 4.** Seasonality of weekly mean between end-uses in the United Kingdom for the weather year 2012. A general yearly trend exists for each end-use, but deviations from that trend such as in late January and early February exist as a result of extreme weather conditions.



**Figure 5.** Seasonally average load curve of this work compared to German standard load profile (SLP) and HTW. Each season consists of three months starting with spring in March. Values are grouped by the hour of the day.

All three load curves share similar troughs at night and peaks in the evening with the exception of the winter evening peak, which our simulation underestimates. The noon peak is close in height between the SLP and the simulation, while the HTW data only have a visible noon peak in winter. A weekday-specific analysis shows that the HTW profiles possess a morning peak rather than a noon peak on working days and only show typical noon peaks on Saturdays and Sundays. The noon peak of the simulation takes place one to two hours early, which stems predominantly from process heat. TUS data suggest a peak for cooking at 11:00 and both the washing machine and dishwasher have their early peak at 10:00, while the SLP assumes its peak around 12:00. Generally, HTW, SLP, and our simulation are equally close to each other with a standard deviation of 3 to 3.8 GW or 26.5% (see Figure A7 in Appendix A). Table 3 lists selected indicators for a quantitative comparison. Both the peak load and load factor, which is the relation of average load to peak load, are closer between SLP and HTW, whereas the minimum load is closer between HTW and this work. We have to consider the possibility that both comparison profiles are currently not entirely representative for Germany either, because of the age of the SLP and the limited scope of the HTW measurements (only single-family houses that are all in



spatial proximity). The SLP is estimated to show a deviation to the actual consumption of more than 20% [3], which is similar to our results.

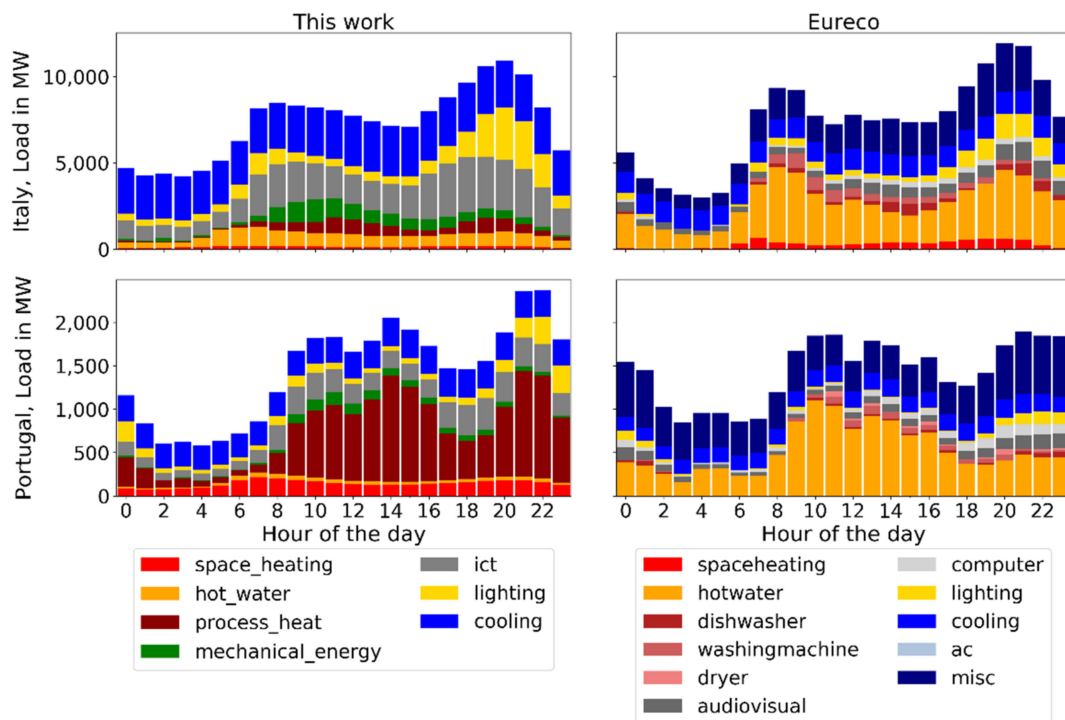
**Table 3.** Comparison of selected indicators over a full year between simulation, German standard load profile (SLP), and the University of Applied Sciences Berlin (HTW). The load factor describes the relation of average load to peak load.

Indicator	This Work	SLP	HTW
Peak load (MW)	30,072	35,779	38,107
Minimum load (MW)	6350	4981	5688
Load factor (no unit)	0.510	0.427	0.403

Analyzing the simulation results of Germany and the United Kingdom already shows a strength of our approach, because we are able to represent very different shares of end-uses in the overall load profile. To extend this analysis, we compare the results of this work to one model for Germany [8] and three models for the United Kingdom [4,9,15]. We calculate the coefficient of determination  $R^2$  between the daily load curves from this work and the existing models. We compare the average daily load curve, because full annual load curves are not available for the existing models. We scale all profiles to the same daily consumption and shift them to the same timezone.

The comparison shows that our approach is capable of generating cross-country load curves that are comparable to existing models built and tested for one country specifically. The German load curve from this work is closely related to the results from [8] with an  $R^2$  of 0.92 and the U.K. load curve is closest to the typical load curve from [4] and [15] with an  $R^2$  of 0.94 and 0.91, respectively. Only *UK Collin* that models the average winter load curve and *UK Ri Sy* that does not contain thermal loads show larger deviations to all other profiles including our work. Figure A8 of the Appendix A shows  $R^2$  and Figures A9 and A10 show the corresponding average daily load curves.

To evaluate our model for other countries than Germany and the United Kingdom, we use a dataset that originates from the EURECO project [42]. During the project, end-use specific load curves of 400 households in four countries were measured. Figure 6 shows a comparison of simulated yearly average load data for Italy and Portugal compared with measured load data from 100 households each. For both countries, the selection of households is not representative for the country's average, as households with best possibilities for measurements were chosen instead. Nevertheless, this work and EURECO load curves are surprisingly close to each other and match both in point of time and height of their peaks. An influence of the selection of households might be visible in the comparably high measured consumption for hot water in both countries. The EURECO project data report a larger share of heating system consumption for both countries than displayed with the statistical data from Eurostat. The shape of EURECO's hot water is, however, close to the simulated process heat load. The AGEBA states that load for devices such as the washing machine or dishwasher is split into mechanical energy used by the electric motors for rotation and process heat used to heat the water [43]. Thus, differences between EURECO and our work could also be a sign of unequal labelling of the measurements. The consistency of the daily averaged curves for both countries is high with an  $R^2$  of 0.8 for Portugal and 0.87 for Italy. Unfortunately, only the averaged load curves are available from EURECO, which is why a quantitative comparison such as Table 3 is not possible.



**Figure 6.** Yearly average load curve of this work compared with EURECO measurements for Italy (above) and Portugal (below). Values are grouped by the hour of the day.

Despite the overall good performance, our model has limitations and uncertainties coming mainly from the small data basis for this study. The training of the model is based on a dataset originating from only the United Kingdom and we have to assume that the dependency on ambient weather conditions is consistent between European countries. In addition, freely available TUS data were found for seven countries in a harmonized format, which requires an arbitrary assignment of profiles for countries where data are not available. For certain countries such as Germany, freely available TUS data are over 40 years old and behavior is very likely to have changed, which prevented us from using country-specific profiles for the end-use ICT. The model would benefit from an update of TUS data in both quantity and quality.

Furthermore, the potential to improve spatial differentiation on resolutions smaller than countries is significant. Information about heating technologies used (heat pumps/electric boilers) or the end-use shares is only available on a national level so far and prohibits an automated detailed examination on a regional level.

#### 4. Conclusions

In this paper, we propose a deterministic approach to model end-use specific regional average household load profiles. We base our model on weather data combined with device profiles and time-use surveys, which are openly available for various European countries. Using historic consumption data, we scaled each end-use profile to achieve a total profile, which we then compared to real-world measurements and other modelling techniques. Our model was trained on modelled U.K. data, where we achieved a MAD/Mean on the test set of 6.5 to 33%. We validated it versus measurements from Germany, Italy, and Portugal, where we could reproduce the corresponding country-specific characteristic overall load shape including night trough and noon and evening peak. We achieved a precision of 26.5% on the annual load curve in Germany.

In addition to the ability to represent country-specific characteristics stemming from end-use shares, our model provides an advantage over the usage of standard load profiles, because it creates real seasonality based on weather data. In combination with generation

profiles of variable renewable energy sources, it allows to use fluctuating profiles of both supply and demand based on the same datasets to analyze their interaction. Because the purpose of many energy system models is to predict future, largely electrified energy systems, we enable modelers to individually adjust the consumption of each end-use based on predictions and represent these changes in the combined load curve. We see a trend of using weather data as the central input to predict regional or national load profiles [44,45]. These models reduce the amount of input data required with an improved methodology, which is motivated by the lack of detailed, cross-country measurement datasets [3]. Although some load profiles developed specifically for one single country might perform better, we offer an approach that can be adopted to all regions where TUS data are available. Our model does not aim to replace existing models, but fills the research gap of an end-use specific, weather-dependent model that is applicable across countries without detailed country-specific appliance ownership and socio-economic data.

**Author Contributions:** Conceptualization, M.S. and R.N.; Data curation, M.S.; Formal analysis, M.S.; Funding acquisition, R.B.; Investigation, M.S.; Methodology, M.S.; Project administration, R.N.; Resources, M.S.; Software, M.S.; Supervision, R.N. and R.B.; Validation, M.S.; Visualization, M.S.; Writing—original draft, M.S.; Writing—review & editing, R.N. and R.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Ministry of Science and Culture of Lower-Saxony, which we gratefully acknowledge. The publication of this article was funded by the Open Access fund of Leibniz Universität Hannover.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data produced in this work are available at the following URL under an open license at the Research Data Repository of the Leibniz Universität Hannover: <https://doi.org/10.25835/0043305>. The repository contains the normalized end-use profiles, the assumed COPs of heat pumps, and real annual consumption to scale the normalized profiles.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

AGEB	German Working Group on Energy Balances
ASHP	Air-source heat pump
BDEW/BGW	German Association of Energy and Water Industries
DSM	Demand-side management
GSHP	Ground-source heat pump
HTW	University of Applied Sciences Berlin
ICT	Information and communication technologies
MAE	Mean absolute error
NN	Neural network
NUTS	Nomenclature of territorial units for statistics
ReLU	Rectified linear unit
SLP(G)	Standard load profile (gas)
TUS	Time-use survey
UK	United Kingdom
$A, B, C, D, b_h, b_w, m_h, m_w$	Building type dependent parameters of the SLPG
$d_{hh,device}$	Stokes' electricity load profile values per device in half-hourly resolution
$d'_{device}$	Unscaled electricity consumption profile per device, generated by our methodology
$d_{end-use}$	Unscaled electricity consumption profile per AGEB end-use
$d_{total}$	Total residential electricity consumption profile

$F_h$	Hour scaling values of SLPG
$f_{ASHP/GSHP}$	Market share of ASHP or GSHP
$h$	Daily consumption factor of SLPG
$k_{hh,device}$	Stokes' constant per device in half-hourly resolution
$S_{hh,device}$	Stokes' daily profile per device in half-hourly resolution
$S'_{device}$	Basic profile per device
$S_{norm,device}$	The normalized basic profile per device
$TUS_{device}$	Activity profile generated from TUS data and assigned to a device
$\theta$	Temperature
$Q_h$	Hourly heat profile from SLPG
$w_{avg}$	Average U.K. weather data between 1994 and 2014
$wde_{device,end-use}$	Weighting factor per device and end-use
$w'$	Region-specific weather data
$x_{day,device}$	Stokes' phase angle per day and device

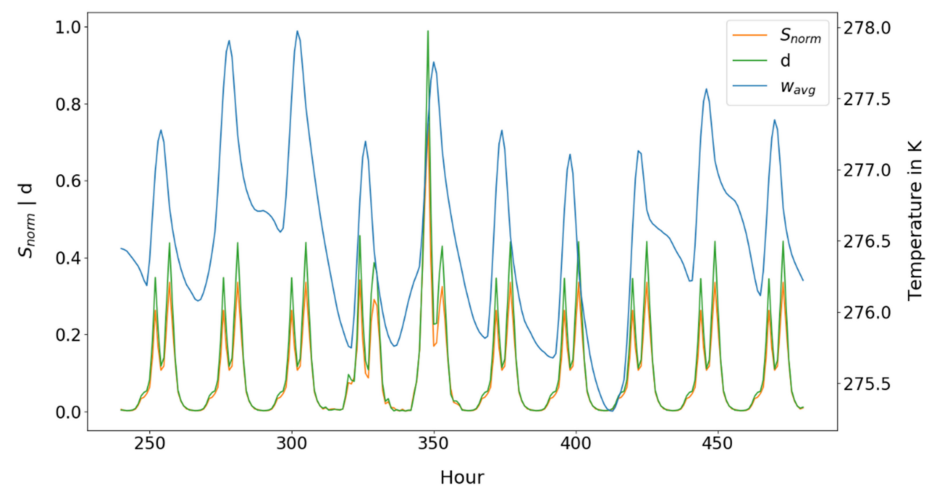
## Appendix A

**Table A1.** Example of time-use survey (TUS) data for one woman of a single day during May in Germany. A TUS is comparable to diaries that characterize how people spend their time. Data are published anonymized, but contain personal information such as age and gender. The activities build harmonized groups to allow easier comparison. The total dataset for Germany consists of 105,000 rows.

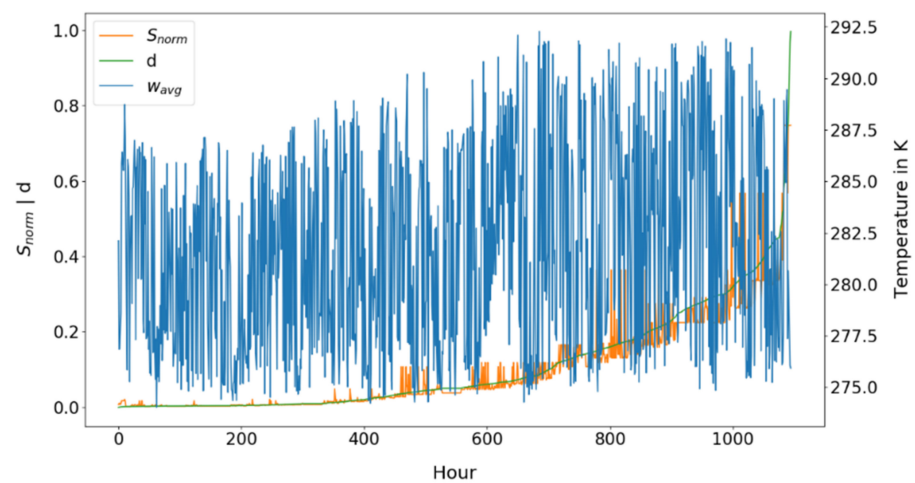
Starting Time	Duration	Location	Activity
00:00	05:45	At own home	sleep
05:45	00:15	At own home	dress/personal care
06:00	01:10	At own home	odd jobs
07:10	01:00	At own home	childcare
08:10	00:30	At own home	childcare
08:40	01:20	At services or shops	shopping
10:00	01:00	At own home	cook, wash up
11:00	01:00	At own home	odd jobs
12:00	00:30	At own home	childcare
12:30	00:30	At own home	cook, wash up
13:00	00:20	At own home	cook, wash up
13:20	00:20	At own home	cook, wash up
13:40	00:20	At own home	cook, wash up
14:00	03:00	At own home	housework
17:00	00:30	At own home	housework
17:30	00:30	At own home	cook, wash up
18:00	00:30	At own home	meals and snacks
18:30	01:00	At own home	cook, wash up
19:30	01:00	At own home	childcare
20:30	01:30	At own home	other leisure
22:00	01:00	At own home	dress/personal care

**Table A2.** Annual consumption per end-use for the countries presented in this work. We use annual consumption per end-use to scale the normalized profiles of our methodology to real load curves.

End-Use	Germany	Italy	Portugal	UK
Space Heating	9.21 TWh	1.13 TWh	1.19 TWh	18.69 TWh
Hot Water	18.03 TWh	5.97 TWh	0.31 TWh	5.07 TWh
Process Heat	37.80 TWh	3.66 TWh	5.30 TWh	6.07 TWh
ICT	21.82 TWh	16.80 TWh	1.97 TWh	25.30 TWh
Cooling	30.08 TWh	23.90 TWh	2.81 TWh	33.49 TWh
Lighting	10.91 TWh	8.40 TWh	0.98 TWh	12.65 TWh
Mechanical Energy	5.78 TWh	4.45 TWh	0.52 TWh	6.70 TWh

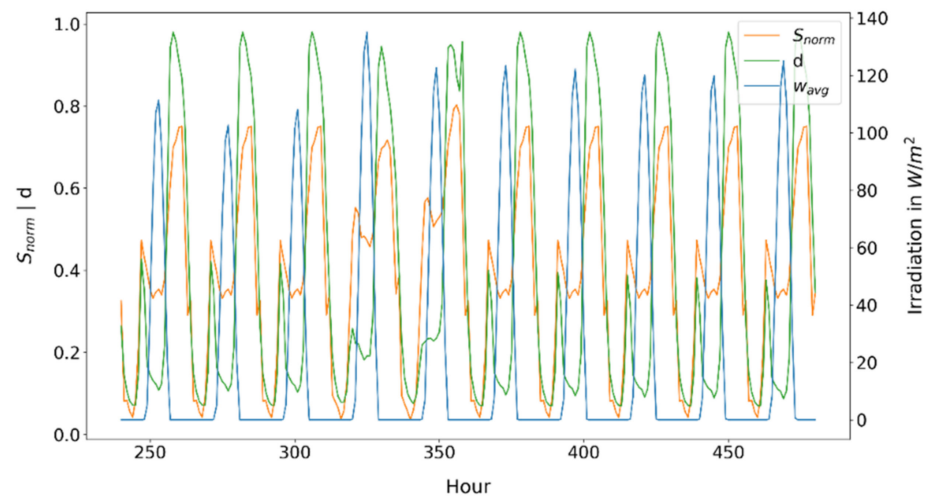


**Figure A1.** Example of training data for the device cooking. The plot shows an excerpt of 10 days in January. We train the neural networks to predict  $d$  from  $S_{norm}$  and  $w_{avg}$  and, therefore, correlate consumption data with input data that are comprehensively available. In this case of cooking,  $d$  is strongly dependent on  $S_{norm}$  and less dependent on  $w_{avg}$ .

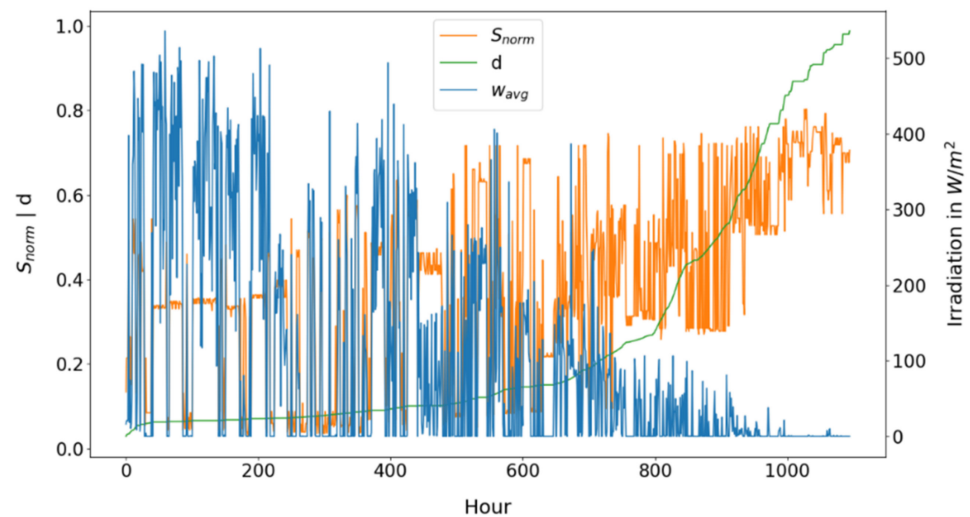


**Figure A2.** Example of the training data of our neural network for the device cooking, ordered by  $d$ . Only every eighth hour is shown to increase the visibility. We train the neural networks to predict  $d$  from  $S_{norm}$  and  $w_{avg}$  and, therefore, correlate consumption data with input data that are available comprehensively. In this case of cooking,  $d$  is strongly dependent on  $S_{norm}$  and less dependent on  $w_{avg}$ .

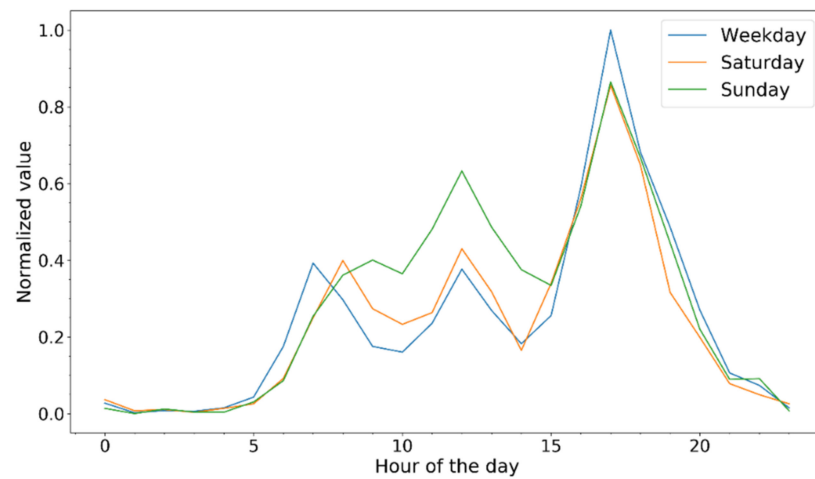




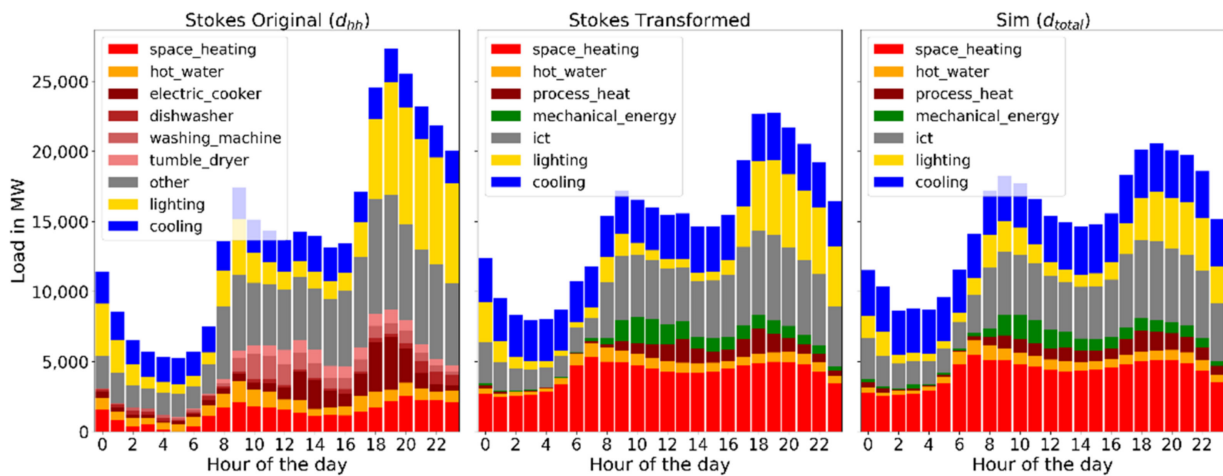
**Figure A3.** Example of training data for the device lighting. The plot shows an excerpt of 10 days in January. We train the neural networks to predict  $d$  from  $S_{norm}$  and  $w_{avg}$  and, therefore, correlate consumption data with input data that are comprehensively available. In this case of lighting,  $d$  is less dependent on  $S_{norm}$  and strongly dependent on  $w_{avg}$ .



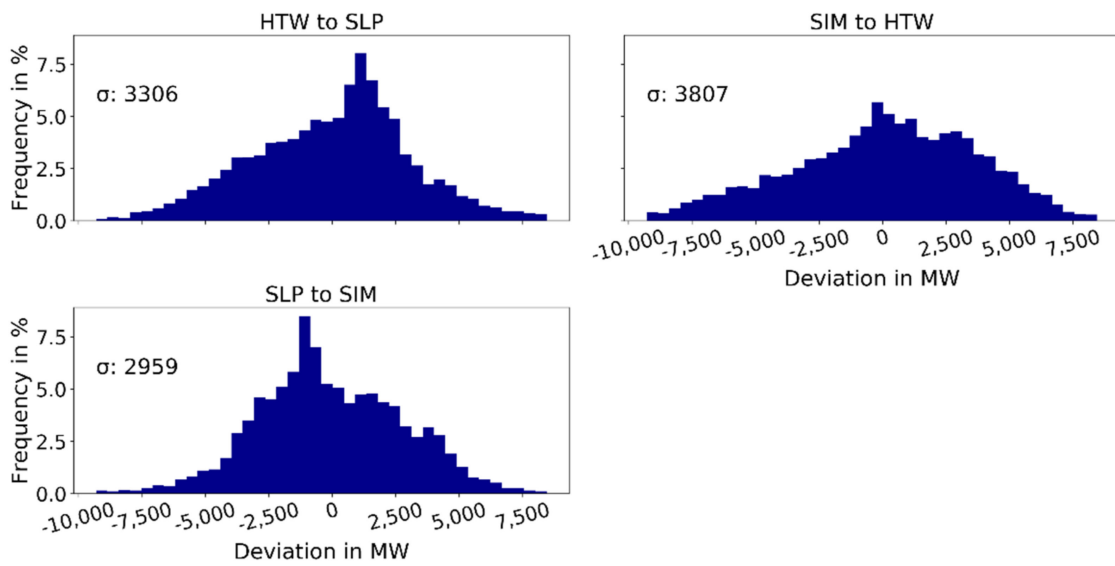
**Figure A4.** Example of the training data of our neural network for the device lighting, ordered by  $d$ . Only every eighth hour is shown to increase the visibility. We train the neural networks to predict  $d$  from  $S_{norm}$  and  $w_{avg}$  and, therefore, correlate consumption data with input data that are available comprehensively. In this case of lighting,  $d$  is less dependent on  $S_{norm}$  and strongly dependent on  $w_{avg}$ .



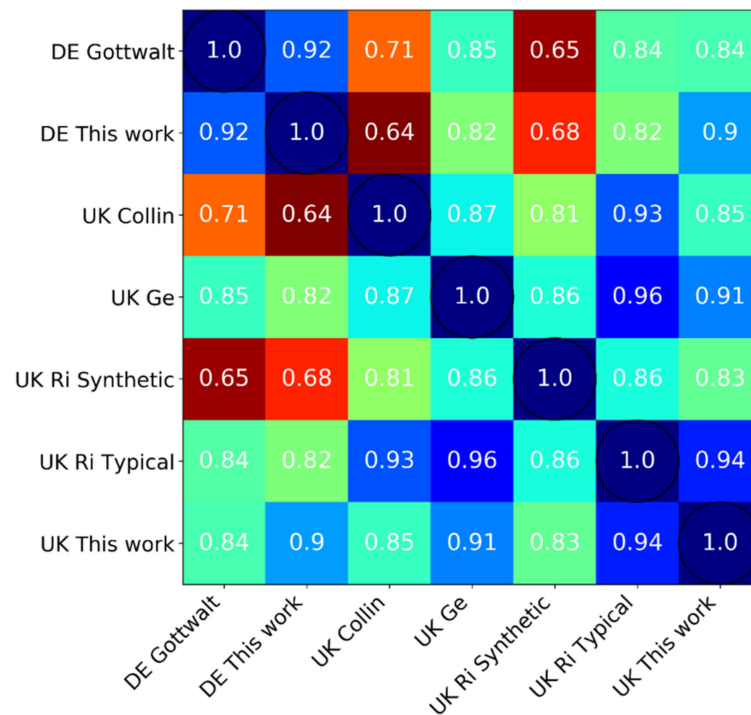
**Figure A5.** Average activity profile from time-use survey (TUS) data for the activity cooking in the United Kingdom in spring season. To build activity profiles, we analyze TUS data to develop shares for each hour of the day of how many persons perform a certain activity. From this, we differentiate the data by age, gender, weekday, and season and scale them with real population data to represent a country's average. The cooking profile shows its peak in the evening between 17:00 and 18:00 and a noon peak at 12:00. The noon peak is more pronounced on Sunday, while the evening peak is more pronounced on weekdays.



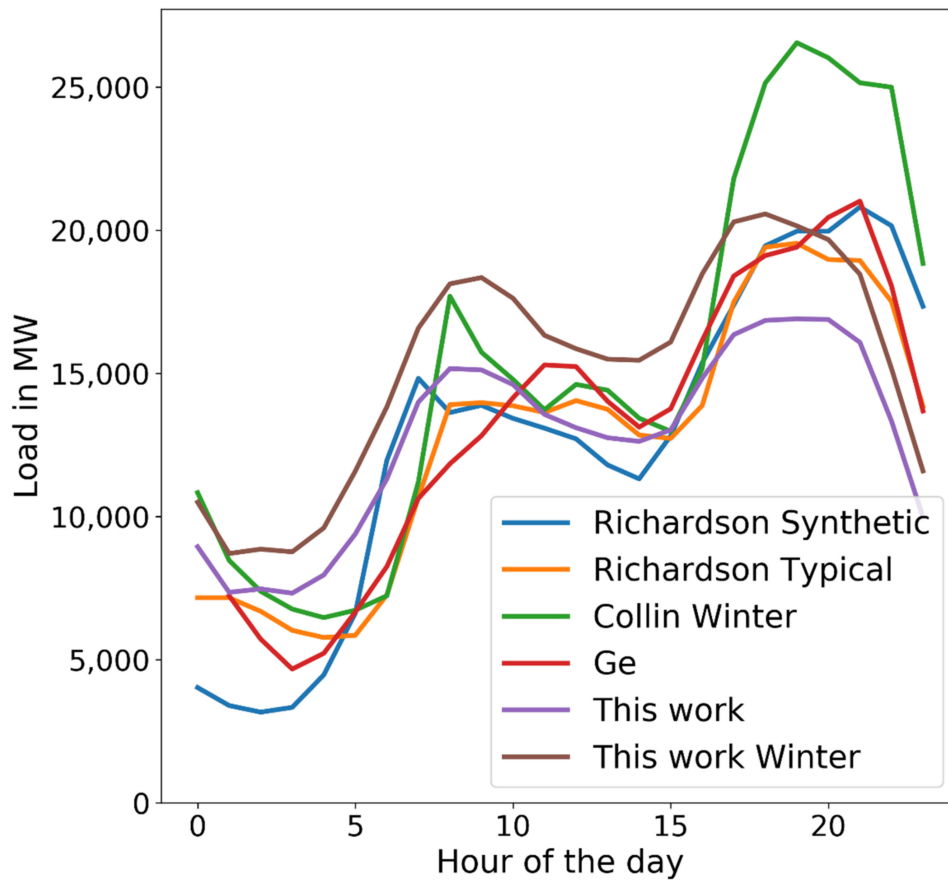
**Figure A6.** Stokes' original load curve (left) vs. transformation to AGEb end-uses (middle) vs. our simulation results using NNs (right) for the average winter weekday (December–February) in the United Kingdom. The first step from Stokes original to Stokes transformed replaces the load curves of space heating and hot water by the SLPG, merges the electric cooker and the wet appliances into process heat and mechanical energy, and rescales each end-use with its respective Eurostat consumption. The step from Stokes transformed to our simulation uses the trained neural networks and has comparably small effects.



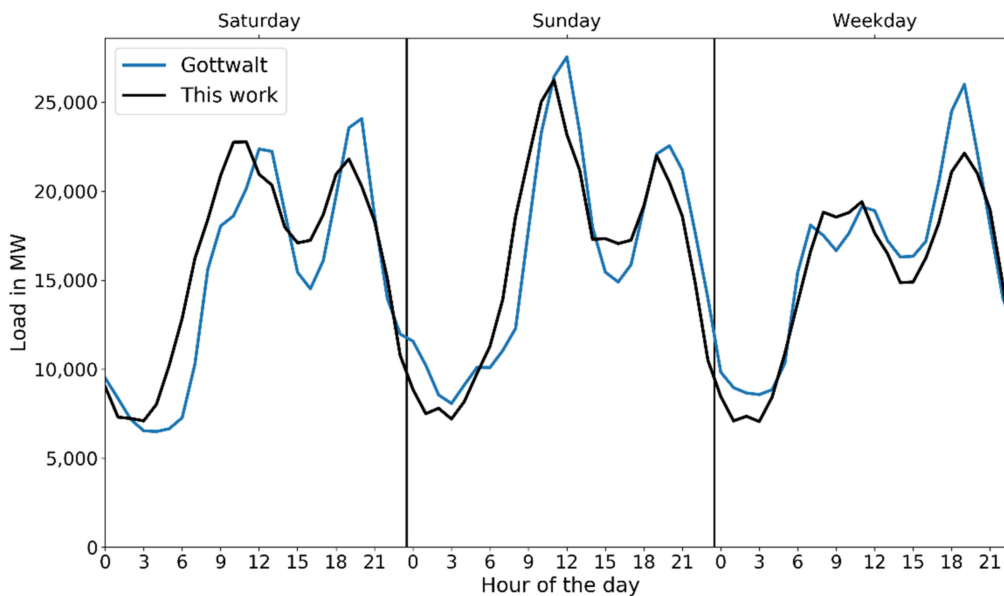
**Figure A7.** Deviation in load between our simulation, measurements from the HTW Berlin, and the German standard load profile H0 for Germany. Values shown are for 8760 h in the weather year 2012. The annual consumption is 134 TWh (average load of 15.3 GW) and is identical for all three profiles.



**Figure A8.** The coefficient of determination  $R^2$  between this work and data reproduced from [8], Gottwalt et al.: 2011; [9], Collin et al.: 2012, [4], Richardson et al.: 2010 and [15], Ge et al.: 2016. All profiles are scaled to the same daily consumption and shifted to the same timezone. *UK Collin* only contains the load curve during winter months and *UK Ri Synthetic* does not contain thermal loads.



**Figure A9.** Comparison of the daily average load curve in the United Kingdom between this work and data reproduced from [9], Collin et al.: 2012, [4], Richardson et al.: 2010 and [15], Ge et al.: 2016. *Collin Winter* and *This work Winter* only show the load curve during winter months and are scaled to the same average consumption. All other profiles are scaled to the same average annual consumption. *Richardson Synthetic* does not contain thermal loads. The X-axis shows local time.



**Figure A10.** Comparison of the daily average load curve in Germany between this work and data reproduced from [8], Gottwalt et al.: 2011. Values are grouped by the hour of the day and the type of weekday and are scaled to the same annual consumption. The X-axis shows local time.

## References

1. Eurostat. Complete Energy Balances. Available online: [https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/nrg\\_110a.tsv.gz](https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/nrg_110a.tsv.gz) (accessed on 2 December 2019).
2. Swan, L.G.; Ugursal, V.I. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1819–1835. [[CrossRef](#)]
3. Proedrou, E. A Comprehensive Review of Residential Electricity Load Profile Models. *IEEE Access* **2021**, *9*, 12114–12133. [[CrossRef](#)]
4. Richardson, I.; Thomson, M.; Infield, D.; Clifford, C. Domestic electricity use: A high-resolution energy demand model. *Energy Build.* **2010**, *42*, 1878–1887. [[CrossRef](#)]
5. European Union. *Harmonised European Time Use Surveys (HETUS) 2018 Guidelines: 2019 Edition*; European Union: Luxembourg, 2019; Available online: <https://ec.europa.eu/eurostat/documents/3859598/9710775/KS-GQ-19-003-EN-N.pdf/ee48c0bd-7287-411a-86b6-fb0f6d5068cc> (accessed on 3 September 2020).
6. McKenna, E.; Thomson, M. High-resolution stochastic integrated thermal–electrical domestic demand model. *Appl. Energy* **2016**, *165*, 445–461. [[CrossRef](#)]
7. Ardakanian, O.; Keshav, S.; Rosenberg, C. Markovian models for home electricity consumption. In Proceedings of the 2nd ACM SIGCOMM Workshop on Green Networking, Toronto, ON, Canada, August, 2011; pp. 31–36. Available online: <https://dl.acm.org/doi/pdf/10.1145/2018536.2018544> (accessed on 25 March 2021).
8. Gottwalt, S.; Ketter, W.; Block, C.; Collins, J.; Weinhardt, C. Demand side management—A simulation of household behavior under variable prices. *Energy Policy* **2011**, *39*, 8163–8174. [[CrossRef](#)]
9. Tsagarakis, G.; Collin, A.J.; Kiprakis, A.E. Modelling the electrical loads of UK residential energy users. In Proceedings of the 2012 47th International Universities Power Engineering Conference (UPEC), Uxbridge, UK, 4–7 September 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 1–6, ISBN 978-1-4673-2856-2.
10. Lombardi, F.; Balderrama, S.; Quoilin, S.; Colombo, E. Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model. *Energy* **2019**, *177*, 433–444. [[CrossRef](#)]
11. Fischer, D.; Härtl, A.; Wille-Haussmann, B. Model for electric load profiles with high time resolution for German households. *Energy Build.* **2015**, *92*, 170–179. [[CrossRef](#)]
12. Pflugradt, N. Modellierung von Wasser- und Energieverbräuchen in Haushalten. Ph.D. Thesis, Technische Universität Chemnitz, Chemnitz, Germany, 2016.
13. Gao, B.; Liu, X.; Zhu, Z. A Bottom-Up Model for Household Load Profile Based on the Consumption Behavior of Residents. *Energies* **2018**, *11*, 2112. [[CrossRef](#)]
14. Singh, S.; Yassine, A. Big Data Mining of Energy Time Series for Behavioral Analytics and Energy Consumption Forecasting. *Energies* **2018**, *11*, 452. [[CrossRef](#)]
15. Ge, Y.; Zhou, C.; Hepburn, D.M. Domestic electricity load modelling by multiple Gaussian functions. *Energy Build.* **2016**, *126*, 455–462. [[CrossRef](#)]
16. Rahman, A.; Srikumar, V.; Smith, A.D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **2018**, *212*, 372–385. [[CrossRef](#)]
17. Arens, S.; Derendorf, K.; Schuldt, F.; Maydell, K.V.; Agert, C. Effect of EV Movement Schedule and Machine Learning-Based Load Forecasting on Electricity Cost of a Single Household. *Energies* **2018**, *11*, 2913. [[CrossRef](#)]
18. Anvari, M.; Proedrou, E.; Schäfer, B.; Beck, C.; Kanzt, H.; Timme, M. Data-Driven Load Profiles and the Dynamics of Residential Electric Power Consumption. 2020. Available online: <https://arxiv.org/pdf/2009.09287.pdf> (accessed on 25 March 2021).
19. Meier, H.; Fünfgeld, C.; Adam, T.; Schieferdecker, B. *Repräsentative VDEW-Lastprofile*, Frankfurt am Main. 1999. Available online: [https://www.bdew.de/media/documents/1999\\_Repraesentative-VDEW-Lastprofile.pdf](https://www.bdew.de/media/documents/1999_Repraesentative-VDEW-Lastprofile.pdf) (accessed on 23 January 2020).
20. Stokes, M. Removing Barriers to Embedded Generation: A Fine-Grained Load Model to Support Low Voltage Network Performance Analysis. Ph.D. Thesis, De Montfort University Leicester, Leicester, UK, 2005.
21. Hong, T.; Fan, S. Probabilistic electric load forecasting: A tutorial review. *Int. J. Forecast.* **2016**, *32*, 914–938. [[CrossRef](#)]
22. Working Group on Energy Balances. Zusammenfassung Anwendungsbilanzen für die Endenergiesektoren 2013–2017. Available online: <https://ag-energiebilanzen.de/8-0-Anwendungsbilanzen.html> (accessed on 5 December 2019).
23. Eurostat. Energy Consumption in Households. Available online: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy\\_consumption\\_in\\_households#Energy\\_products\\_used\\_in\\_the\\_residential\\_sector](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households#Energy_products_used_in_the_residential_sector) (accessed on 16 April 2019).
24. Dörn, S. *Programmieren für Ingenieure und Naturwissenschaftler: Intelligente Algorithmen und Digitale Technologien*; Springer: Berlin/Heidelberg, Germany, 2018; ISBN 978-3-662-54304-7.
25. Shalev-Shwartz, S.; Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms*; Cambridge University Press: Cambridge, UK, 2014; ISBN 078-1-107-05713-5.
26. Geyer, B.; Rockel, B. coastdat-2 COSMO-CLM Atmospheric Reconstruction. Available online: [https://cera-www.dkrz.de/WDCC/ui/ceraresearch/q?query=coastDat\\_2\\_COSMO-CLM&page=0&rows=15](https://cera-www.dkrz.de/WDCC/ui/ceraresearch/q?query=coastDat_2_COSMO-CLM&page=0&rows=15) (accessed on 25 March 2019).
27. Kolassa, S.; Schütz, W. Advantages of the MAD/Mean Ratio over the MAPE. *Foresight Int. J. Appl. Forecast.* **2007**, *40*–43.
28. Gershuny, J.; Fisher, K. Multinational Time Use Study. Available online: <https://www.timeuse.org/mtus> (accessed on 2 April 2019).
29. Gershuny, J.; Sullivan, O. United Kingdom Time Use Survey. Available online: <https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8128> (accessed on 21 May 2019).



30. Eurostat. Population by Single Year of Age and NUTS 3 Region. Available online: [https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/cens\\_11ag\\_r3.tsv.gz](https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/cens_11ag_r3.tsv.gz) (accessed on 3 December 2019).
31. Scharp, M.; Wilke, A.; Magro, M. Bestandsaufnahme: Energieeffizienz von Haushaltsgeräten. 2016. Available online: [https://www.izt.de/fileadmin/publikationen/SLHw\\_Energieeffizienz\\_Haushaltsgeraete.pdf](https://www.izt.de/fileadmin/publikationen/SLHw_Energieeffizienz_Haushaltsgeraete.pdf) (accessed on 16 December 2019).
32. Hellwig, M. Entwicklung und Anwendung Parametrisierter Standard-Lastprofile [Development and Application of Parameterized Standard Load Profiles]. Ph.D. Thesis, Technische Universität Chemnitz, München, Germany, 2003.
33. German Association of Energy and Water Industries; German Association of Local Utilities; European Association of Local Energy Distributors. Abwicklung von Standardlastprofilen Gas [Execution of Gas Standard Load Profiles]. 2018. Available online: [https://www.vku.de/fileadmin/user\\_upload/Verbandsseite/Sparten/Energiewirtschaft/Gasthemen/nach\\_Leitfaden\\_SLP.pdf](https://www.vku.de/fileadmin/user_upload/Verbandsseite/Sparten/Energiewirtschaft/Gasthemen/nach_Leitfaden_SLP.pdf) (accessed on 8 March 2021).
34. German Association of Energy and Water Industries. Anwendung von Standardlastprofilen zur Belieferung Nichtleistungsgemessener Kunden [Execution of Gas Standard Load Profiles for the Supply of Non-Metered Customers]: Praxisinformation P 2006/8. 2006. Available online: [http://www.gwb-netz.de/wa\\_files/05\\_bgw\\_leitfaden\\_lastprofile\\_56550.pdf](http://www.gwb-netz.de/wa_files/05_bgw_leitfaden_lastprofile_56550.pdf) (accessed on 16 December 2019).
35. Ruhnau, O.; Hirth, L.; Praktiknjo, A. Time series of heat demand and heat pump efficiency for energy system modeling. *Sci. Data* **2019**, *6*, 189. [[CrossRef](#)] [[PubMed](#)]
36. Staffell, I.; Brett, D.; Brandon, N.; Hawkes, A. A review of domestic heat pumps. *Energy Environ. Sci.* **2012**, *5*, 9291. [[CrossRef](#)]
37. Born, H.; Schimpf-Willenbrink, S.; Lange, H.; Bussmann, G.; Bracke, R. Analyse Des Deutschen Wärmepumpenmarktes: Bestandsaufnahme und Trends. Available online: [http://www.geothermie-zentrum.de/fileadmin/media/geothermiezentrum/Projekte/WP-Studie/Web\\_WaermepumpenstudieIII\\_01.pdf](http://www.geothermie-zentrum.de/fileadmin/media/geothermiezentrum/Projekte/WP-Studie/Web_WaermepumpenstudieIII_01.pdf) (accessed on 14 July 2020).
38. Eurostat. Disaggregated Final Energy Consumption in Households-Quantities. Available online: [https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/nrg\\_d\\_hhq.tsv.gz](https://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?file=data/nrg_d_hhq.tsv.gz) (accessed on 4 September 2020).
39. Hülk, L.; Wienholt, L.; Cußmann, I.; Müller, U.P.; Matke, C.; Kötter, E. Allocation of annual electricity consumption and power generation capacities across multiple voltage levels in a high spatial resolution. *Int. J. Sustain. Energy Plan. Manag.* **2017**, *79–92*. [[CrossRef](#)]
40. Robinius, M.; Stein, F.t.; Schwane, A.; Stolten, D. A Top-Down Spatially Resolved Electrical Load Model. *Energies* **2017**, *10*, 361. [[CrossRef](#)]
41. Tjaden, T.; Bergner, J.; Weniger, J.; Quaschnig, V. Repräsentative Lastprofile für Wohngebäude in Deutschland auf 1-sekündiger Datenbasis [Data file]. Available online: <https://pvspeicher.htw-berlin.de/veroeffentlichungen/daten/lastprofile/> (accessed on 23 January 2020).
42. Project EURECO. End-Use Metering Campaign in 400 Households of the European Community: Assessment of the Potential Electricity Savings. Available online: [http://www.eerg.it/resource/pages/it/Progetti\\_-\\_MICENE/finalreporteureco2002.pdf](http://www.eerg.it/resource/pages/it/Progetti_-_MICENE/finalreporteureco2002.pdf) (accessed on 23 January 2020).
43. Schmidt, C.; Frondel, M.; Janßen-Timmen, R.; Sommer, S. Erstellung der Anwendungsbilanzen 2016 und 2017 für den Sektor der Privaten Haushalte und den Verkehrssektor in Deutschland: Endbericht-Juni 2019. 2019. Available online: [https://ag-energiebilanzen.de/index.php?article\\_id=29&fileName=rwi\\_anwendungsbilanz\\_2019\\_priv\\_hh\\_und\\_verkehr.pdf](https://ag-energiebilanzen.de/index.php?article_id=29&fileName=rwi_anwendungsbilanz_2019_priv_hh_und_verkehr.pdf) (accessed on 30 March 2021).
44. Ali, D.; Yohanna, M.; Puwu, M.I.; Garkida, B.M. Long-term load forecast modelling using a fuzzy logic approach. *Pac. Sci. Rev. A Nat. Sci. Eng.* **2016**, *18*, 123–127. [[CrossRef](#)]
45. Behm, C.; Nolting, L.; Praktiknjo, A. How to model European electricity load profiles using artificial neural networks. *Appl. Energy* **2020**, *277*, 115564. [[CrossRef](#)]