Bottom-up Markov Chain Monte Carlo approach for scenario based residential load modelling with publicly available data

Citation for published version (APA):

Nijhuis, M., Gibescu, M., & Cobben, S. (2016). Bottom-up Markov Chain Monte Carlo approach for scenario based residential load modelling with publicly available data. *Energy and Buildings*, 112, 121-129. <https://doi.org/10.1016/j.enbuild.2015.12.004>

DOI:

[10.1016/j.enbuild.2015.12.004](https://doi.org/10.1016/j.enbuild.2015.12.004)

Document status and date:

Published: 15/01/2016

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Bottom-up Markov Chain Monte Carlo approach for scenario based residential load modelling with publicly available data

M. Nijhuis^{a,*}, M. Gibescu^a, J.F.G. Cobben^{a,b}

^a Electrical Energy Systems, Eindhoven University of Technology, Den Dolech 2, Eindhoven, The Netherlands

^b Liander N.V., Utrechtseweg 86, Arnhem, The Netherlands

ARTICLE INFO

Article history:

Received 25 May 2015

Received in revised form 23 October 2015

Accepted 4 December 2015

Available online 9 December 2015

Keywords:

Residential energy consumption modelling

Electricity consumption

Occupancy

Household appliance

Markov Chain

Load curves

ABSTRACT

In the residential sector, with the introduction of electric vehicles and photovoltaics, developments are taking place which have an impact on residential load curves. In order to assess the integration of these new types of technologies on both the generation and load side, as well as to develop mitigation strategies like demand side management, detailed information is required about the load curve of a household. To gain knowledge about this load curve a residential load model is developed based on publicly available data. The model utilises a Markov Chain Monte Carlo method to model the occupancy in a household based on time use surveys, which together with weather variables, neighbourhood characteristics and behavioural data are used to model the switching pattern of appliances. The modelling approach described in this paper is applied for the situation in the Netherlands. The resulting load curve probability distributions are validated with smart meter measurements for 100 Dutch households for a week. The validation shows that the model presented in this paper can be employed for further studies on demand side management approaches and integration issues of new appliances in distribution grids.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The energy transition is changing the loading of the distribution grid especially in residential areas. With distributed generation consumers are becoming producers and the electrification of heat and transportation shifts the energy demand from fossil fuels to electricity. The current practices of grid development employed by network operators needs to be adjusted to ensure efficient integration of new loads and generation technologies and to reap the benefits of demand side management. However the current modelling of residential loads is not accurate enough to assess these problems and opportunities [1], as the energy transition and demand side management alter the diversity and stochastic characteristics of the household load curve. With the installation of smart meters more opportunities arise to validate a residential load modelling approach, while the need for modelling remains present from the perspective of forecasting, long-term planning and due to privacy concerns about smart meter data.

In the literature models have been presented for the estimation of the residential load curve. These models can be divided into roughly two basic categories: Top-down models; which focus

on the loading of MV/LV transformers and generate load curves based on this aggregation level [2] and bottom-up methods which employ statistical energy usage data or time use data to construct load profiles. There are many different approaches when it comes to building the bottom-up models. These bottom-up models can be combined with the top-down models through smart meter data [3,4]. Machine learning approaches are applied to model household load curves based on smart meter data measurements [5,6]. The residential load curve can also only be modelled at the peak times [7,8]. The modelling of household load over multiple decades requires an adjusted modelling approach, for instance based on how typical close all cal households, behaviour and appliances change over scenarios [9]. More in-depth reviews of these different household load curve models have been performed [10,11].

The integration of distributed generation creates a voltage rise which is the main problem in distribution grids [12]. To assess the level of the voltage a more in depth assessment than just the peak load is required. The bottom-up time series approach is the most adequate for the modelling of residential load curves in order to assess the integration of distributed generation. To be able to assess the impacts of DSM, information on the appliance level needs to be known. The changes in the loading because of the shift in appliance usage depend on the actual appliances which is shifted. Demand side management approaches are generally based on information on the appliance level (e.g. [13,14]). A bottom-up household load

* Corresponding author.

E-mail address: m.nijhuis@tue.nl (M. Nijhuis).

Nomenclature

$\mathbb{A}$	set of appliances used in model
$\mathbb{A}_H$	set of appliances present in household
A	appliance
$f_{t,A}$	fraction of time appliance A is on
f_A	fraction of households with appliance A
hr	hour
$I(t)$	irradiance
L_A	average lifetime of appliance A
l_A	remaining lifetime of appliance A
O	occupancy
$P_H(t)$	household power use
$P_A(t)$	power used by appliance A
$P_{RA}(t_A)$	power of appliance A while running during time t_A
P_{SA}	stand-by power of appliance A
r_h	household size adjustment factor
r_w	household wealth adjustment factor
t	time
t_0	initial time shift
t_A	time after which the appliance A became active
T_w	wind speed adjusted temperature
$t_{A,m}$	minimum run time of appliance A
$t_{A,r}$	average run time of appliance A
$T_{h/c}$	heating/cooling temperature
$n[\mu, \sigma]$	random variable from normal distribution $\mathcal{N}(\mu, \sigma)$
$u[a, b]$	random variable from uniform distribution over interval $[a, b]$
$x[\lambda]$	random variable from exponential distribution $\exp(\lambda)$

modelling approach is therefore a logical choice for the assessment of demand side management (since the bottom-up approach allows for the shifting of individual appliances). Different value propositions exist for demand side management of residential loads, based on the self-consumption, electricity price or network loading. Therefore the residential load curve should be adaptable for many different scenarios, for the model to be usable when assessing future possibilities for demand side management. The bottom-up time series approach may be found in a number of references [15–20], however these approaches either focus on a limited aspect of the residential load, use private data or do not allow for the inclusion of scenarios.

The approach shown in [16] is based on a large database of measurements of appliances in a household and as this information is usually not publicly available, this approach cannot be taken. The approach described in [17] focuses on the heating load, which is appropriate for areas with electric heating, however as the Netherlands has a fairly small percentage of electric heating loads, these kind of approaches lack the required level of detail for the non-heating loads. The approach described in [19], is not designed for load modelling over multiple decades and is therefore hard to use in grid development. With the approach used in [18] assumptions have to be made on the switching behaviour of loads. This is not necessary in the approach described in this paper since statistical data from time use surveys are available in many countries [21].

The paper is organised as follows: Section 2 describes the modelling approach, followed by the modelling of the occupancy of a household and the modelling of the appliances in the household. In Section 3 the results of the model for a typical neighbourhood in the Netherlands are presented and the model is validated using smart meter measurements. Conclusions are given in Section 4.

Fig. 1. General schematic of a Markov Chain model with three states.

2. Modelling approach

The proposed modelling approach is based on the occupancy of the household, and to a lesser extent the behaviour of the members of the household. This approach is taken as occupancy is a driving factor for energy usage [22] and the changes in occupancy play an important role in the changes in energy consumption. The occupancy is modelled based on a Markov Chain Monte Carlo method (an explanation on Markov Chain Monte Carlo modelling is given in [23]). In Fig. 1 a schematic of a generic Markov Chain model is given with probabilities P of changing from one state to the other. At each time instance depending on the current state and the associated probabilities the model switches to another state or remains in the current state.

Next to the occupancy of a household the set of electrical appliances in a household and their energy use need to be modelled. This is done by creating a model for the distribution of appliances over the household based on statistical data on appliance ownership and the level of wealth in the neighbourhood and size of the dwelling. These were identified as key drivers for the degree of appliance ownership within a household [24,25].

After the occupancy of a household has been established and appliances assigned to the household the simulation of the electricity usage patterns for the appliances can be performed. This is done by simulating the switching on/off of each appliance individually using another Markov Chain based on the time of day and the weather conditions.

A flow chart of the model is presented in Fig. 2 to illustrate the computation of the load curves. The flow chart starts on the left with the inputs of the model, where the ellipses are national/state-wide inputs and the rounded rectangles are local inputs. In the following subsections the steps are explained in more detail.

2.1. Occupancy modelling

To get a better understanding of the occupancy, the occupancy profile as reported in a time use survey (2042 respondents, reporting their behaviour at a 15 min time interval, available from the Dutch national statistics agency: Statistics Netherlands) has been plotted in Fig. 3 for two consecutive days for three individual persons of different households. From this figure the large differences between the behaviour of the three persons becomes apparent. These differences will translate into differences in energy use, therefore it is important that the variation present in the occupancy is also incorporated into the model.

The first step in creating the occupancy model is determining the characteristics of a reported occupancy time series. As residents can be either active (at home and not sleeping) or inactive

Fig. 5. Comparison of the average level of occupancy of 2042 occupants between the model and the time use survey, for a Sunday and a Monday.

Fig. 6. Comparison of the distribution of the average number of consecutive 15 min intervals of continuous activity or inactivity level of 2042 occupants.

to adjust the occupancy based on correlation with previous day at same time by applying the following formula:

$$O(t) = \frac{O(t - 24 \text{ hr}) \times a + O^*(t) \sqrt{1 - a^2}}{a + \sqrt{1 - a^2}} \text{ if } t \text{ AND } t - 24 \text{ hr are both weekdays} \quad (2)$$

with $O^*(t)$ the initial occupancy as calculated with Eq. (1) and $a = u[0.4, 0.1]$ which is determined from the distribution of individual households of the autocorrelation data at the 24 h time lag. To test the quality of the Markov Chain model, the synthesised occupancy data is compared to the actual data from the time use survey.

In Fig. 5 the average level of occupancy over a set of 2042 occupants over two days for both the model and the time use survey is shown. From the figure it can be seen that there is a large difference in the occupancy during the morning and the middle of the day between the Sunday and the Monday. The modelled and the measured average occupancies are more or less similar. The aggregated modelled occupancy thus closely matches the measured one, however next to the aggregated occupancy the distribution of individual occupancy time series should also match. In order to test this, Figs. 6 and 7 have been created.

In Figs. 6 and 7 the occupancy time series is transformed from a binary series to a series indicating the consecutive number of 15 min intervals of being active as a positive number and the consecutive number of intervals of being inactive as a negative number. The distribution of the mean and standard deviation of the 2042 individual series are shown in the figures. From Fig. 6 it becomes clear that the mean of the distributions is equal for both the model and the time use survey (in accordance with Fig. 5), however the

Fig. 7. Comparison of the distribution of the standard deviation of the number of consecutive 15 min intervals of activity or inactivity level of 2042 occupants.

distribution is flatter for the time use survey than for the model. This indicates that the model does not create enough extreme cases where people on average are staying active or inactive for the majority of a day. For single person households, this has a definite effect on the energy use, however for households consisting of multiple persons the extreme cases of inactivity and activity could be combined to have little effect on the energy use. The standard deviation of the consecutive 15 min intervals as shown in Fig. 7 shows a closer connection, indicating that the variability within the individual times series is similar for the modelled and survey data

2.2. Appliances present in the household

The appliances which are present in the household can be modelled based on statistical data on the level of penetration of a specific appliance in a given population group. As these data are usually only available on the national scale, adjustments have to be made to generate an estimate of the appliances present in a single household. The level of these appliances scales with income and dwelling size (number of rooms in the dwelling) as these are identified as the most important parameters determining the level of appliances [27]. Information on the penetration of different appliances in a household is estimated from the Statistics Netherlands, various commerce organisations and environmental NGO's. Statistics Netherlands has for a number of appliances the penetration grade differentiated with respect to neighbourhood wealth level and dwelling size. The income and dwelling size dependency are modelled through logistic regression analysis on these data and assumed equal for similar appliances for which no statistical data was available. The regression results for these two variables can be seen in Fig. 8 and the resulting formula for the allocation of appliances to a household can be seen in Eq. (3).

$$\text{for all } A \in \mathbb{A} \quad \begin{cases} A \in \mathbb{A}_H, & \text{if } u[0, 1] \leq f_A \times r_h \times r_w \\ A \notin \mathbb{A}_H, & \text{otherwise} \end{cases} \quad (3)$$

2.3. Appliance model

The appliances which are present in the model can be divided into a number of categories. These categories require a different modelling approach. The main modelling of the appliances is done on a 15-min basis. There are however appliances which have an operating time of less than 15-min this requires a shorter time frame as well as to model the starting and stopping of an appliance random within the 15-min interval. The appliances are modelled with 1-min time frame t_A for their own energy usage P_A . For certain

Fig. 8. Regression of the number of appliances versus the normalised dwelling size and income level.

appliance categories the usage is dependent on the time of day. For these appliances the notation includes *hr* which is use for reasons of clarity and interpreted as 60 min in all the calculations. The household load is calculated as the summation of the loads of the individual appliances:

$$P_H(t) = \sum_{A \in A_H} P_A(t) \quad (4)$$

For each category the modelling approach is discussed below. For the appliances the information on the load profile is taken from [28–30].

- Baseload (e.g. fridge, modem) is assumed to be independent of occupancy in terms of energy use. The modelling of these appliances is simply a constant energy use or a switch source with constant switching times evaluated with a variable step size related to the switching of an average appliance of that kind.

$$P_A(t) = \begin{cases} P_{R,A}, & \text{if } (t - t_0) - hr \cdot \left\lfloor \frac{t - t_0}{hr} \right\rfloor < f_{t,A} \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (5)$$

with the initial time shift t_0 randomly chosen from $u[-30, 30]$ to create diversity between appliances. The switching behaviour is based on the percentage of time the appliance is on in one hour by taking the lower bound of the time divided by an hour.

- Night load (e.g. electric boilers, dishwashers) is switched on at an exponentially distributed time interval after 21 h (the beginning of the off-peak tariff).

$$P_A(t) = \begin{cases} P_{R,A}(t_A), & \text{if } x\left(\frac{1}{24 hr}\right) < \frac{\left\lfloor \frac{t - t_0}{24 hr} \right\rfloor - 21 hr}{24 hr} \\ P_{R,A}(t_A), & \text{if } P_A(t-1) \neq P_{S,A} \text{ AND } t_{A,r} \geq t_A \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (6)$$

- Heating and cooling loads (e.g. air-conditioner, electric heating) have a probability to switch on depending on the temperature and the wind speed [31] and are modelled according to a heating degree day approach as illustrated in [32]. The irradiance also has an influence on the required heating, but this effect is for a large part already included in the temperature. For the construction of the daily heating profile the data from the allocation process in the Dutch gas sector [33] is used. This approach is adjusted

for occupancy by the addition of a rule depending on the occupancy and the temperature. If there is an active occupant or the wind speed adjusted temperature $T_w(t)$ drops below zero degrees and at least two hours have passed since the last operation of the central heating system then heating will be used. For a more advanced heating model this rule can be replaced based on the specific proprieties of the dwelling (e.g. orientation, insulation, surface and window area, etc.).

$$P_A(t) = \begin{cases} P_{R,A}(t_A, T_w), & \text{if } T_w(t) \leq T_{h/c}(t) \text{ AND } \sum_{i=t-120}^t O(i) > 0 \\ P_{R,A}(t_A, T_w), & \text{if } t_A < t_{A,m} \text{ AND } P_A(t-1) > P_{S,A} \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (7)$$

- Lighting is modelled based on the solar irradiation and the occupancy, with lights turning on if the irradiance is below 60 W/m² [34]. The level of lighting usage is scaled with the number of occupants present in the dwelling [35].

$$P_A(t) = \begin{cases} P_{R,A}(t_A) \times \sqrt{O(t)}, & \text{if } O(t) > 0 \text{ AND } I(t) < 60 \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (8)$$

- Behavioural loads (e.g. TV, oven) are loads which can have a direct link to the activities reported in the time use survey, like watching TV or preparing food. As the energy related behaviour of the occupants depends on a large number of factors [36], the modelling is simply based on the statistical data. The power of the behavioural loads is modelled in a Markov Chain with a transition matrix which changes over time and these loads can only be activated if at least one person is active in the household. For the creation of these matrices, the same regression approach (based on household size and age) as applied in the modelling of the occupancy is employed.

$$P_A(t) = \begin{cases} P_{R,A}(t_A), & \text{if } u[0, 1] \geq \mathcal{P}(P_A(t)|P_A(t-1)) \text{ AND } O(t) > 0 \\ P_{R,A}(t_A), & \text{if } t_A < t_{A,m} \text{ AND } P_A(t-1) > P_{S,A} \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (9)$$

with $\mathcal{P}(P_A(t)|P_A(t-1))$ the chance of switching the appliance on, if the appliance was off in the previous time instance.

- General loads (e.g. water-cooker, hair dryer) can only be activated if at least one member of the household is active. This category include all loads which are not used frequently enough to link them to activities reported in the time use survey, or have an operating time which is smaller than 15 min.

$$P_A(t) = \begin{cases} P_{R,A}(t_A), & \text{if } u[0, 1] \leq f_{t,A} \text{ AND } O(t) > 0 \\ P_{R,A}(t_A), & \text{if } t_A < t_{A,m} \text{ AND } P_A(t-1) > P_{S,A} \\ P_{S,A}, & \text{otherwise} \end{cases} \quad (10)$$

2.4. Scenario implementation

In order to gain insight in the development of the residential load when the penetration of technologies like heat pumps or EV increases, scenarios can be combined with the residential load model. With the introduction of a certain scenario, the type and number of appliances in the household, the usage of the appliances and the occupancy of the household can change. For most scenarios only large changes in the presence of certain appliances in the household are assessed. Scenarios on the changes in behaviour or the occupancy are not as often discussed in literature and smaller changes are expected. For both types of scenarios the implementation in the model is elaborated upon below.

Fig. 9. Comparison of the active power consumption of a household (averaged over 100 households) for two weekdays, for four different scenario combinations.

As the probability matrices for the occupancy are already divided into groups depending on age and household size, implementing scenarios which change these parameters can be done by changing the utilised probability matrices to the matrices from another group. The changes in behaviour relate to changes in the occupancy or the activities which the persons of the household performs. Scenario implementation involves changing the probability matrices, this is done by updating the probability matrices each half year with ones adjusted to the changes introduced by the scenarios.

To implement the changes in appliances, every appliance has a lifetime l_A based on average reported lifetimes L_A [37]. At a pre-determined number of intervals the appliances remaining lifetime is calculated and the available list of appliances is updated with (depending on the scenario) more efficient appliances. If an appliance remaining lifetime reaches zero, the appliance is removed from the household and based on the penetration level there is a chance that a new appliance of the same type is added to the household.

At each of these intervals there is also a chance that a household gains an additional appliance based on the household characteristics, penetration level and the lifetime of the appliance. If the penetration level of an appliance changes an additional chance of either adding or removing the appliance from the household is applied to each household. This is applied in the following manner:

$$l_A = L_A \times u[0.5, 1.5] \quad (11)$$

if l_A reaches zero, appliance A is removed from the set. After each week $\mathbb{A}_H$ is updated with formula (3) with an additional chance based on the penetration of the appliance in the model compared to the expected penetration with respect to the chosen household size and neighbourhood wealth level. The factor f_A changes based on the scenarios on the penetration of technologies, r_w changes with the economic growth scenario and the factors $P_{RA}(t_A)$ & P_{SA} change with the efficiency scenarios.

The appliance efficiency scenario of the EU [38], the scenario on the Dutch economic growth [39], scenarios on the implementation of electric vehicles, PV, micro-CHP and heat pumps [40,41] are implemented in the model, to illustrate the implementation of the scenarios in the model. To generate a clearer implementation of these scenarios, the scenario studies are consolidated into three scenarios on the implementation of a certain technology; a high, medium or low penetration scenario. In Fig. 9 an example of changes in the aggregated household load for a selected quartet of scenario combinations is given.

From the figure a clear difference can be seen between the different scenarios. The high PV scenario creates a clear energy surplus in

Table 1

RMSE in Watt over one week for seven weeks of transformer data, smart meter data and modelled data.

	Transformer	Smart meter	Model
Transformer	86	78	91
Smart meter	78	72	79
Model	91	79	64

Fig. 10. Autocorrelation of the residuals of the modelled data, with the smart meter data and the transformer data and the residuals of the transformer data with the smart meter data.

the middle of the day, while the heat pump scenario generates more baseload, while the peak is not effected. The introduction of more energy efficient appliances with a low economic growth reduces the overall loading significantly.

3. Results and validation

To assess the performance of the model, a couple of simulations has been performed. First 100 households are simulated for 1 year on an Intel i7-3770 3.4 GHz processor, 8 GB ram computer. The simulation of these 100 households took 27s and 800 MB of RAM memory, expanding the simulation to 1000 households increased the computation time to 249s and memory usage to 2 GB, while lengthening the simulation of the 100 households to 10 years instead of 1 increased the computational time to 214s and the memory usage to 1 GB. Indicating the model can be used to generate large sets of household load curves with no scalability issues. The 1 year of data of these 100 households are compared to smart meter data aggregated to 100 households and scaled (to account for the losses and the number of connected households) transformer data of a transformer of 107 households and one of 94 households. The root mean square error (RMSE) is calculated between the time series of these data sets and shown in Table 1.

The table shows that the RMSE of the model, the transformer data and the smart meter data all are in the same range. The RMSE between the sets of smart meter data and the two set transformer time series are in the same range as the RMSE of these data set with the model. This indicates that the differences between two datasets due to the stochastic nature of the residential load has the same order of magnitude as the differences between the modelled data and the measured transformer and smart meter data. To gain more insight into the errors between the modelled data and the measured data the residuals are computed and the autocorrelation of these residuals is calculated. The results are shown in Fig. 10.

The autocorrelation of the residuals gives an indication as to whether the model constantly generates a positive or negative error at certain time steps. The autocorrelation of the residuals shows a positive peak at the 24h time lag and a negative peak at the 12h

Fig. 11. Comparison of the average active power consumption, for the model and measured data for 100 households over four weekdays.

Fig. 12. Probability distribution of the electricity use of 100 households during a 15 min interval in the night (4 AM) for the model and smart meter data normalised to 1 kW base.

Fig. 13. Probability distribution of the electricity use of 100 households during a 15 min interval in the peak (7 PM) for the model and smart meter data normalised to 1 kW base.

time lag. Both these peaks are larger in magnitude than the confidence bounds, indicating that there is a period biased present in the residuals. The residuals between the transformer and the smart meter data and the modelled and the measured data both all show the same effect. Though there is some bias within the model, this does not exceed the bias between the measured data.

Secondly the model was used to simulate the energy use in 100 households for four days (Wednesday to Saturday) in the winter. The resulting load profiles are compared with data from smart meters and from MV/LV transformer measurements of different neighbourhoods with similar characteristics as can be seen in Fig. 11. As the smart meter data was in 15 min and the transformer measurements in 10 min resolution, a time step of 15 min is employed for the comparison.

For the majority of time instances Fig. 11 shows that the three time series are close together, however the stochastic nature of the residential load, even at an aggregation level of 100 households, remains clearly visible. The modelled household load has on weekdays a morning peak which is slightly higher compared to the smart meter data and the MV/LV transformer data and a minimum load level which is lower compared to the measured data. The evening peak, and the night time energy use is comparable between the three time series, however the peak of the loading measured at the MV/LV transformer level is higher (due to the energy losses in the LV-cables) in comparison to the peak of the smart meter measurements and the model, with the model giving a relatively low peak on the fourth day.

The kernel smoothed probability density function of the set of the energy use of 100 individual households is plotted in Fig. 12 for a single 15 min interval during the night time (at 4 AM as

Table 2

Mean and standard deviation of the night (4 AM) and peak load (7 PM) for both the model and the smart meter data

		Smart meter	Model
Night	Mean [W]	228	237
	Std [W]	195	234
Peak	Mean [W]	820	729
	Std [W]	680	641

indicated in Fig. 11). In Fig. 13 the distribution is plotted for a 15 min interval during the peak (at 7 PM as indicated in Fig. 11). The mean and standard deviation for the model and the smart meter data are given in Table 2. The distributions of the model and the smart meter data have a similar shape, indicating that the model simulates individual household load curves with a similar stochastic properties as measured by the smart meters. The difference in the mean of the distribution of the model and the smart meter data can also be seen in Fig. 11 at the indicated time steps.

The smart meter measurements and the model both show a smaller variance in the night compared to the peak distribution, which is caused by the relatively high share of base load appliances running at night. The peak time shows more variability since at that moment many occupancy/behavioural driven appliances are operated.

4. Conclusions

A model for the creation of residential electrical load curves is presented. The model utilises a bottom-up time series approach which makes it applicable for the assessment of demand side

management through the use of home automation systems. The model is developed in a highly adaptive manner so it can be used to analyse future behavioural changes and the integration of emerging technologies, renewable generation and changes in the energy usage of appliances. The bottom-up approach is based on data publicly available and employed for the specific case of the Netherlands. The model is however general enough to be applied for any country where appliance ownership, time of use survey data is available and can be adjusted to the local conditions based on the specific information about the climate and the age, wealth and household size distribution. Scenario analysis has been introduced in the model so the effectiveness of different home energy management strategies can be evaluated over a range of scenarios. On an aggregated level the model shows behaviour similar to measurements at the MV/LV transformer and to that of smart meter data. The distribution of modelled individual load curves during the daily peak and the night time is comparable to the distribution obtained from smart meter data, validating the hypothesis that the model could be used to compute the loading of a set of aggregated households. The validation with smart meter data on the distribution of individual households shows also that the model generates a residential load profile which is close to the actual measured data.

4.1. Future work

The space heating demand in a household is at the moment simply modelled by using the approximate heat demand profile given by the gas usage, however this could be improved to a more sophisticated method which employs the thermal dynamics of the dwelling as done for instance in [17]. In addition to the heating demand, the tap water was also not modelled in great detail. Methods such as [42] could be implemented in the model to determine the effects of switching to electric instead of natural gas for the heating of tap-water. No scenarios for the human behaviour are currently implemented in the model, as these are at the moment researched to a far smaller extent than scenarios on the changes in appliances. Scenarios can be however be implemented by changing the transition probability matrices over time. As many statistical agencies have been taking time of use surveys for multiple years, trends from these surveys could be deduced and extrapolated as scenarios on the changes in behaviour of occupants (e.g. more prevalence of home office users).

References

- [1] J. Dickert, P. Schegner, Residential load models for network planning purposes, *MEPS* (2010) 1–6.
- [2] C.M. Colson, M.H. Nehrir, An alternative method to load modeling for obtaining end-use load profiles, in: *NAPS*, 2009, pp. 1–5, <http://dx.doi.org/10.1109/NAPS.2009.5484036>.
- [3] K.E. Train, An assessment of the accuracy of statistically adjusted engineering (SAE) models of end-use load curves, *Energy* 17 (7) (1992) 713–723, [http://dx.doi.org/10.1016/0360-5442\(92\)90079-F](http://dx.doi.org/10.1016/0360-5442(92)90079-F).
- [4] F. Wallin, C. Bartusch, E. Thorin, E. Dahlquist, Important parameters for prediction of power loads – a bottom-up approach utilizing measurements from an automatic meter reading system, in: *PowerAfrica*, 2007, pp. 1–7, <http://dx.doi.org/10.1109/PESAfr.2007.4498105>.
- [5] A. Azadeh, Z.S. Faiz, A meta-heuristic framework for forecasting household electricity consumption, *Appl. Soft Comput.* 11 (2011) 614–620, <http://dx.doi.org/10.1016/j.asoc.2009.12.021>.
- [6] T. Zia, D. Bruckner, A. Zaidi, A hidden markov model based procedure for identifying household electric loads, in: *IECON 2011–37th Annual Conference of the IEEE Industrial Electronics Society*, 2011, pp. 3218–3223, <http://dx.doi.org/10.1109/IECON.2011.6119826>.
- [7] M.A. López-Rodríguez, I. Santiago, D. Trillo-Montero, J. Torriti, A. Moreno-Munoz, Analysis and modeling of active occupancy of the residential sector in Spain: an indicator of residential electricity consumption, *Energy Policy* 62 (2013) 742–751, <http://dx.doi.org/10.1016/j.enpol.2013.07.095>.
- [8] O. Ardakanian, S. Keshav, C. Rosenberg, Markovian models for home electricity consumption, in: *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking – GreenNets '11*, 2011, p. 31, <http://dx.doi.org/10.1145/2018536.2018544>.
- [9] A. Grandjean, G. Binet, J. Bieret, J. Adnot, B. Duplessis, A functional analysis of electrical load curve modelling for some households specific electricity end-uses, in: *EEDAL*, Vol. 6th, 2011, p. 1.
- [10] A. Grandjean, J. Adnot, G. Binet, A review and an analysis of the residential electric load curve models, *Renew. Sustain. Energy Rev.* 16 (9) (2012) 6539–6565, <http://dx.doi.org/10.1016/j.rser.2012.08.013>.
- [11] L.G. Swan, V.I. Ugursal, Modeling of end-use energy consumption in the residential sector: a review of modeling techniques, *Renew. Sustain. Energy Rev.* 13 (2009) 1819–1835, <http://dx.doi.org/10.1016/j.rser.2008.09.033>.
- [12] A. Canova, L. Giaccone, F. Spertino, M. Tartaglia, Electrical impact of photovoltaic plant in distributed network, *IEEE Trans. Ind. Appl.* 45 (1) (2009) 341–347, <http://dx.doi.org/10.1109/TIA.2009.2009726>.
- [13] D.G. Infield, J. Short, C. Horne, L.L. Freris, Potential for domestic dynamic demand-side management in the UK, in: *2007 IEEE Power Engineering Society General Meeting*, 2007, pp. 1–6, <http://dx.doi.org/10.1109/PES.2007.385696>.
- [14] D. Caprino, M.L. Della Vedova, T. Facchinetti, Peak shaving through real-time scheduling of household appliances, *Energy Build.* 75 (2014) 133–148, <http://dx.doi.org/10.1016/j.enbuild.2014.02.013>.
- [15] J.V. Paatero, P.D. Lund, A model for generating household electricity load profiles, *Int. J. Energy Res.* 30 (2006) 273–290, <http://dx.doi.org/10.1002/er.1136>.
- [16] J. Widén, E. Wäckelgård, A high-resolution stochastic model of domestic activity patterns and electricity demand, *Appl. Energy* 87 (6) (2010) 1880–1892, <http://dx.doi.org/10.1016/j.apenergy.2009.11.006>.
- [17] K.P. Schneider, J.C. Fuller, Detailed end use load modeling for distribution system analysis, in: *PES General Meeting*, 2010, pp. 1–7, <http://dx.doi.org/10.1109/PES.2010.5588151>.
- [18] J. Dickert, P. Schegner, A time series probabilistic synthetic load curve model for residential customers, in: *PowerTech*, 2011, pp. 1–6, <http://dx.doi.org/10.1109/PTC.2011.6019365>.
- [19] I. Richardson, M. Thomson, D. Infield, C. Clifford, Domestic electricity use: a high-resolution energy demand model, *Energy Build.* 42 (10) (2010) 1878–1887, <http://dx.doi.org/10.1016/j.enbuild.2010.05.023>.
- [20] R. Yao, K. Steemers, A method of formulating energy load profile for domestic buildings in the UK, *Energy Build.* 37 (6) (2005) 663–671, <http://dx.doi.org/10.1016/j.enbuild.2004.09.007>.
- [21] Access time use data – centre for time use research, 2014 <http://www.timeuse.org/information/access-data>.
- [22] S. Firth, K. Lomas, A. Wright, R. Wall, Identifying trends in the use of domestic appliances from household electricity consumption measurements, *Energy Build.* 40 (5) (2008) 926–936, <http://dx.doi.org/10.1016/j.enbuild.2007.07.005>.
- [23] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Springer US, 1996, <http://dx.doi.org/10.1007/978-1-4899-4485-6>.
- [24] Y.G. Yohanis, J.D. Mondol, A. Wright, B. Norton, Real-life energy use in the UK: how occupancy and dwelling characteristics affect domestic electricity use, *Energy Build.* 40 (6) (2008) 1053–1059, <http://dx.doi.org/10.1016/j.enbuild.2007.09.001>.
- [25] A. Druckman, T. Jackson, Household energy consumption in the UK: a highly geographically and socio-economically disaggregated model, *Energy Policy* 36 (8) (2008) 3177–3192, <http://dx.doi.org/10.1016/j.enpol.2008.03.021>.
- [26] A. Pentland, A. Liu, Modeling and prediction of human behavior, *Neural Comput.* 11 (1) (1999) 229–242, <http://dx.doi.org/10.1162/089976699300016890>.
- [27] Nederlands instituut voor budgetvoorlichting, *Energielastenbeschuwing: Verschillen in energielasten tussen huishoudens nader onderzocht*, Tech. rep., NIBUD, Utrecht, 2009.
- [28] R. Stamminger, C. Pakula, H. Jungbecker, M. Braun, I. Rüdenauer, C. Wendker, G. Broil, *Synergy Potential of Smart Appliances*, Tech. rep., Rheinische Friedrich-Wilhelms-Universität Bonn, 2008.
- [29] ACS-F1: database of appliance consumption signatures, 2014, <http://www.wattict.com/web/index.php/databases/acs-f1>.
- [30] Virginia Tech advanced research institute research data, 2014 <http://www.ari.vt.edu/research-data>.
- [31] M. Hart, R. de Dear, Weather sensitivity in household appliance energy end-use, *Energy Build.* 36 (2) (2004) 161–174, <http://dx.doi.org/10.1016/j.enbuild.2003.10.009>.
- [32] R.G. Quayle, H.F. Diaz, Heating degree day data applied to residential heating energy consumption, *J. Clim. Appl. Meteorol.* 19 (1980) 241–246, [http://dx.doi.org/10.1175/1520-0450\(1980\)019<0241:HDDDAT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0450(1980)019<0241:HDDDAT>2.0.CO;2).
- [33] Nederlandse mededingingsautoriteit, *Allocatievoorwaarden gas*, Tech. rep., NMA, 2012.
- [34] I. Richardson, M. Thomson, D. Infield, A. Delahunty, Domestic lighting: a high-resolution energy demand model, *Energy Build.* 41 (7) (2009) 781–789, <http://dx.doi.org/10.1016/j.enbuild.2009.02.010>.
- [35] M. Stokes, M. Rylatt, K. Lomas, A simple model of domestic lighting demand, *Energy Build.* 36 (2) (2004) 103–116, <http://dx.doi.org/10.1016/j.enbuild.2003.10.007>.
- [36] J. Stephenson, B. Barton, G. Carrington, D. Gnoth, R. Lawson, P. Thorsnes, Energy cultures: a framework for understanding energy behaviours, *Energy Policy* 38 (10) (2010) 6120–6129, <http://dx.doi.org/10.1016/j.enpol.2010.05.069>.
- [37] Average life expectancy of major household appliances, 2014 <http://www.statista.com/statistics/220020/average-life-expectancy-of-major-household-appliances/>.

- [38] W. Eichhammer, T. Fleiter, B. Schlomann, S. Faberi, M. Fioretto, N. Piccioni, S. Lechtenböhrer, A. Schüring, G. Resch, Study on the energy savings potentials in EU member states, candidate countries and EEA countries, Tech. rep., Commission, European, 2009.
- [39] L.H.J.M. Janssen, V.R. Okker, J. Schuur, Welvaart en leefomgeving: een scenariostudie voor Nederland in 2040, Tech. rep., Centraal Planbureau, Milieu- en Natuurplanbureau en Ruimtelijk Planbureau, 2006.
- [40] J.J.C. Bruggink, The next 50 years: four European energy futures, Tech. rep., ECN, 2005.
- [41] E. Veldman, Power play impacts of flexibility in future residential electricity demand on distribution network utilisation, Ph.D. thesis, Eindhoven University of Technology, 2013.
- [42] J. Widén, M. Lundh, I. Vassileva, E. Dahlquist, K. Ellegård, E. Wäckelgård, Constructing load profiles for household electricity and hot water from time-use data – modelling approach and validation, Energy Build. 41 (2009) 753–768, <http://dx.doi.org/10.1016/j.enbuild.2009.02.013>.