ASSESSMENT OF PROBABILISTIC METHODS FOR SIMULATING HOUSEHOLD LOAD PATTERNS IN DISTRIBUTION GRIDS

Michiel Nijhuis
Eindhoven University of Technology
The Netherlands
m.nijhuis@tue.nl

Bram Vonk
Enexis - The Netherlands
Eindhoven University of Technology
bram.vonk@enexis.nl

Madeleine Gibescu
Eindhoven University of Technology
The Netherlands
m.gibescu@tue.nl

Sjef Cobben
Liander N.V. – The Netherlands
Eindhoven University of Technology
sjef.cobben@alliander.com

Han Slootweg
Enexis - The Netherlands
Eindhoven University of Technology
han.slootweg@enexis.nl

ABSTRACT

Upcoming intermittent generation technologies such as wind and photovoltaics, and the integration of energy intensive loads such electric vehicles, are creating new challenges for DNOs in planning and operation of distribution grids. Probabilistic load profile models of households are often used to assess the adequacy of the distribution networks and the possible mitigation strategies. These models are required to be accurate, allow for the implementation of local conditions and scenarios while remaining computationally feasible. Two modelling approaches for generating household load profiles are compared. The first is a top-down model based on field measurements and the second is a bottom-up model based on the energy usage of appliances within a household. The assessment which an asset engineer needs to perform depends which model is most appropriate to use. For most studies the quicker top-down method generates the best results, while for detailed scenario assessments the bottom-up model is preferable.

INTRODUCTION

The transition towards a more sustainable and more electricity driven energy system is changing the household load. With the introduction of technologies like electric vehicles (EVs) and heat pumps the household load is expected to rise, while with the introduction of rooftop photovoltaics (PV) households are starting to generate their own electricity. For a DNO (Distribution Network Operator) this results in an increased risk of asset overloading and more difficult voltage control at multiple voltage levels. Next to these developments, the rise in ICT allows for new mitigation options. Demand response and active network reconfiguration become options which the DNO can apply to enhance the capabilities of the network and to deal with changes in load profiles.

Most of the connections of a DNO consist of residential customers. Assessing the load of households is thus of vital importance for the DNO to accurately assess the adequacy of the distribution grid. The modelling of loads is currently done by using standard profiles on the MV-level, and applying a maximum load by using a coincidence factor dependent on the number of households for the peak loading on the LV-level.

Neither of these models are suitable for the changes introduced by the energy transition, as both the coincidence factor and the peak loading become highly dependent on the penetration rate of electric vehicles and heat pumps, while the use of generic load profiles is no longer adequate as the penetration of different technologies and therefore the load curve of the combined households can differ significantly from one neighbourhood to another. Next to this, more advanced probabilistic methods are necessary to be able to fully evaluate the benefits of 'smart grid' mitigation options, like demand response. As the models should be used by asset engineers responsible for the long-term planning of the grid (where many future scenarios must be evaluated), next to the accuracy of the proposed household load model, a reasonable computational time is an important requirement.

For the creation of a household load model most approaches can be classified as either a top-down or a bottom-up approach. Assessments on the general strengths and weaknesses of such models have been examined in [1], but not for the specific requirements of network planning. To assess which approach is more suitable for the purpose of network planning, a top-down and a bottom-up model are compared in this work. Assessing their strengths and weaknesses allows for the creation of more advanced models in the future.

First, the top-down model will be discussed in detail, followed by a discussion of the bottom-up model. After this, the assessment criteria for the models will be proposed. Hereafter the models are compared based on these criteria and conclusions are drawn on the suitability of these models for the network planning process.

THE TOP-DOWN MODEL

The top down model is developed to generate load patterns for hundreds of MV-LV transformers in a MV grid, to enable impact analysis of PV and EV penetration.

Such grids contain tens of thousands households which again individually have multiple appliances as described further for the bottom-up model. Incorporating the individual contribution of each appliance would be a significant computational burden for practical integration into a network planning tool. Therefore, a top-down model is suggested as shown in Figure 1.

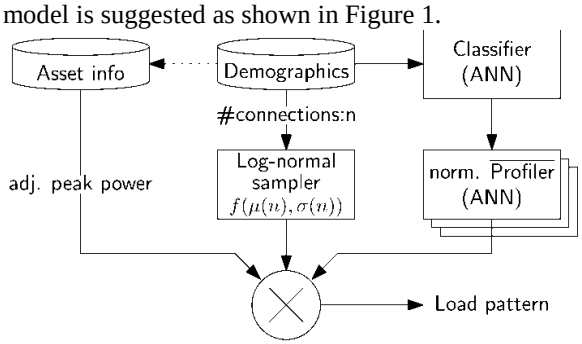

Figure 1: Flowchart of the top-down determination of the household load.

The general thought behind the top-down model is to split the probabilistic load behaviour and the expectation of the load per moment in time. It is assumed that the averaging time window is large enough to consider the load as an independent stochastic variable as stated in [2][3]. The approach is based on measurements from transformers and smart meters.

Classifier: More than 30 reference MV/LV transformers were measured for a period of two years. Relative demographics such as rental, building usage, and income level shares, and population density are used to classify and link all transformers in a MV grid to this set of reference transformers. This was done with an artificial neural network (ANN)-based classifier. Poorly classified transformers (classifying factor < 0.8 of range [0,1]) are linked to an extra transformer class. The classifier determines which profiler model is used in the next step.

Profiler: The measurements of the reference transformers are used as target data to train several ANNs. The output is a profile of a day with a time resolution of fifteen minutes. The input of the ANN consists of the daily temperature, separate dummy variables for Saturdays and Sundays, and the (co-)sine components of the day of the year and the time of the day. The ANN profiler smoother output profile tham the target data it is trained with due to the performance evaluation (MSE) in the training process. In post-processing the profiler's output is therefore normalised such that the maximum yearly value equals one. The resulting profile is the expectation of the load per time of day.

Log-normal sampler: The log-normal sampler adds the stochastic component to the top-down model. The idea behind this is explained step-by-step as follows. Smart meter data of 27 household connections from the months January and June are used to make average household patterns of combinations ranging up to 27 households. In Figure 2 this is shown per time of day (tod).

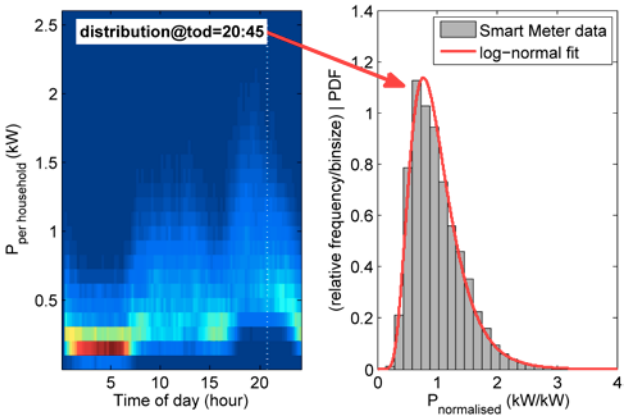

Figure 2: Per household load pattern from a combination of five households (January, workdays); Bivariate histogram per tod (left) and a normalised distribution (right).

As stated by [4], the distribution per tod fits a log-normal distribution with parameters a,b and density function:

$$f(x|a,b) = \frac{1}{xb\sqrt[2]{2\pi}} e^{\frac{-(\ln x - a)^2}{2b^2}}$$

When the per tod distributions are normalised by their monthly average values (Figure 2, right inset), the parameters are mainly dependent on the number of households of a combination n. For a certain aggregation level the log-normal distribution of the time window resembles a normal distribution as [2] states. Figure 3 shows with a scatter plot and fitted curves, how both parameters depend on the number of households for different type of days and months. The latter factors do not seem to affect the fitted curve, although, there are some variations in R^2. The fitted curves do, however, deviate from the theoretical curve for the sum of log-normal distributions given by [5], especially in the range of 10 to 100 households when also the transformer measurements are included. The following fit function for a and b is used:

$$f(n|p_1,p_2,p_3,p_4,p_5) = \frac{p_1}{p_2+p_3 n^{p_4}} + p_5 \text{ and}$$

$$a = f(n|\text{-}20.14, 30.46, 10.85, 1.477,\ \text{-}3.191e^{-3}),$$

$$b = f(n|\ 229.9, 206.6, 49.47, 1.102, 72.04e^{-3}).$$

Depending on n a log-normal distribution is initially created for each transformer. Then, for every day, 96 (15-minute) values are sampled from this distribution.

Adjusted peak power: The DNO generally measures the yearly peak power of a transformer. If not, this could also be deducted from demographics. However, the yearly peak will be higher than the maximum of the expectation profile generated by the profiler ANN. Therefore, the peaks are adjusted by a factor consistent with the peak value probability. This factor is therefore distribution dependent and determined by the number of households.

Element-wise multiplication of the profiler and sampler results, times the adjusted peak power, results in a probabilistic load pattern. By omitting the log-normal sampler in this last step, one gets a smooth (average) profile which could be considered as a load forecast.

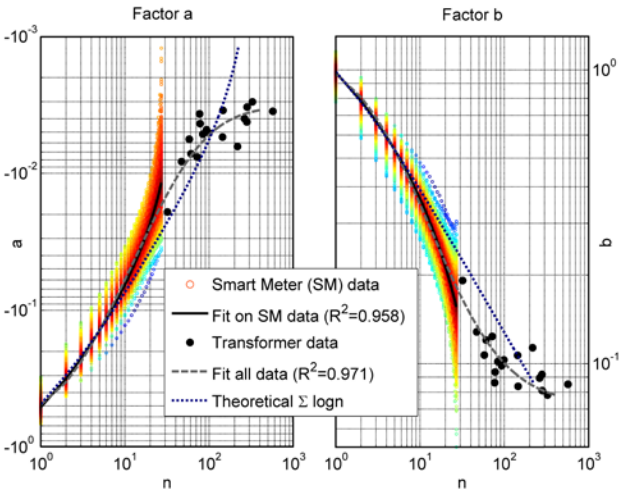

Figure 3: The log-normal parameters a (left) and b (right) versus the number of households per transformer n.

THE BOTTOM-UP MODEL

The household load is a result of the interactions between the household members and the appliances in the household. To model the household load, both these aspects must be taken into account. A modelling approach based on the occupancy of a household, and (depending on data availability) on the behaviour of the members of the household is employed. An overview of the modelling approach is given in Figure 4.

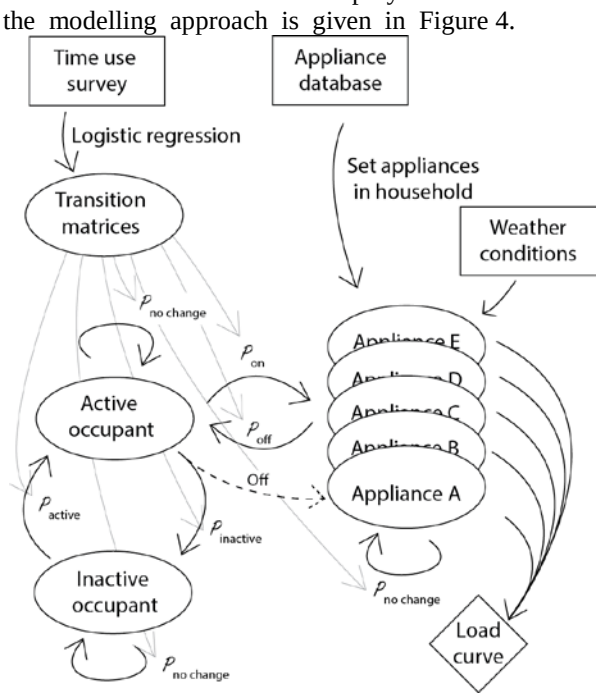

Figure 4: Markov chain model for the bottom-up determination of the household load.

A time use survey by the Netherlands institute for social research in which over 2000 persons have logged their activity for one week on a 15-min basis is used to generate the behaviour of the members of the household. The occupancy of the household is the most important thing which can be determined from the time use data, as it is shown that occupancy is a driving factor for energy usage [6]. The modelling of the occupancy is done based on a Markov Chain Monte Carlo approach [7]. In this approach transition matrices are used to determine whether the state of an occupant in the household changes, whether he/she becomes active or inactive (not present in the household or asleep). The transition matrices are created by using logistic regression on the time use survey data, to gain different transition probability matrices depending on the household size and the age of the household members. By using this approach the occupancy for a certain household can be simulated for each 15-min interval in a day. Next to the occupancy, a similar approach is used to compute transition probability matrices for activities which can be directly linked to the usage of a certain appliance (i.e. watching TV).

With the occupancy of the household known, the appliances in the household need to be modelled. First the appliances are distributed over the households. The households are characterised based on the average level of income of the neighbourhood and the number of household members, as these factors are the main drivers for the number of appliances present [8][9]. For the allocation of appliances to a household, statistics on the penetration levels and the amount of appliances per household type (size & income) are used and normal distributions for these variables are assumed.

The appliances are divided into six types: base-load, night load, heating and cooling loads, lighting, activity linked loads and general loads. New loads like PV and EV can also be introduced as additional appliance types, to contribute in evaluating the household load for future scenarios. Based on the type of appliance, a different simulation approach is chosen.

Base-load (e.g. fridge, modem) can be modelled independently of all external inputs, as these appliances are constantly turned on or are switched on and off with little user interference. The energy use of these appliances is thus modelled as constant or variable switch energy usage.

Night load (e.g. electric boilers, dishwashers) consist of loads which usually only run at night (or in off-peak hours). Depending on the occupancy level on the day, the appliances start to switch on after 9PM with a probability based on an exponential distribution.

Heating and cooling loads (e.g. air-conditioner, electric heating) are dependent on the weather conditions and the occupancy. The main weather conditions which influence the use of these loads are the outside temperature and the wind speed [10]. To evaluate how much heating energy is required, data from the allocation process in the Dutch gas sector [11] is used. The gas use in this approach is dependent on a wind speed adjusted temperature and is converted to a required heat to be delivered to the household. This data is combined with the occupancy in order to generate a lower heating demand when the

household has no active occupants.

Lighting is for the most part only used when the irradiance is low (<60W/m^2). The conversion of the irradiance values to energy use is based on [12] [13]. The number of lights and the level of energy usage from lighting is dependent on the number of active occupants in the dwelling.

Activity-linked loads (e.g. TV, oven) can be seen as loads which have a direct relation to behaviour reported in the time use surveys. The underlying patterns of behaviour are dependent on many variables [14], however as only the resulting energy use is of interest, the appliances are modelled with a Markov Chain, with transition matrices generated through logistic regression of the time use data.

General loads (e.g. water-cooker, hair dryer) are loads which have no clear link to behaviour reported in the time use survey, but can only be turned on when occupants are active. These loads have a chance to switch on depending on the yearly average reported usage time of the appliance.

The total household load can be constructed by adding all the appliance loads together.

ASSESSMENT CRITERIA

To determine the appropriateness of the models for use in the network planning process, both models need to be evaluated on a number of criteria.

One of the most important characteristics of a good household load model for network planning is the peak load, as the strength of the network is usually assessed based on the peak load.

The accuracy of the rest of the load curve is of lesser importance. However for assessing the impact of new technologies on the grid, the determination of the energy losses is also important, and if a more probabilistic approach to the load flow calculations is required, the entire load curve must be known.

For the network planning it is important that the local conditions can be implemented in the load model, as there can be large differences between the electricity usage of certain neighbourhoods.

As the network planning requires an estimate of the load for the coming years, the implementation of scenarios within the load model is also required.

Asset engineers need to evaluate networks with different solutions and under various loading situations. The modelled household load is one of the main inputs parameters for this process. If the computational burden of the household load model becomes large, and many input variables are required, it will hamper the asset engineer and therefore become a less valuable tool for the network planning process.

An overview of the assessment criteria, their measures and relative importance are given in Table 1.

To assess the models, they are compared to smart meter data, transformer data and Velander load estimates [2], which are currently used tools in the network planning process.

Table 1: Overview of the assessment criteria and weights

Criteria	*Measure*	*Weight*
Peak load	Error	0.4
Load curve	RMSE	0.2
Local conditions	Qualitative	0.1
Scenario implementation	Qualitative	0.1
Computational speed	time	0.2

RESULTS

For both models time series of a year of data has been created for the comparison. An example of the load curves of the two models, measured transformer data and smart meter data (of a different neighbourhood) for two days is given in Figure 5. Here it can be seen that the four time series lay close to each other, only the bottom-up model predicts a much higher morning peak on the second day and the peak of the smart meter data and the bottom-up model is significantly higher. The RMSE between the transformer data and the models relative to the RMSE of the smart meter data is used for the comparison and given in Table 2 (lower is better).

As peak load is one of the most important characteristics for the planning of distribution networks the two models are compared to smart meter data and the Velander model (currently used by Dutch DNOs). An overview of how these models compare is given in Figure 6.

The relative error compared to the smart meter data of the two models is used as comparison measure for the peak loading and given in Table 2 (higher indicates a better estimate).

For the implementation of the local conditions the top-down model uses fitting to measured data with similar local conditions, while for the bottom-up model the local conditions can only be implemented by using neighbourhood statistics as provided by Statistics Netherlands.

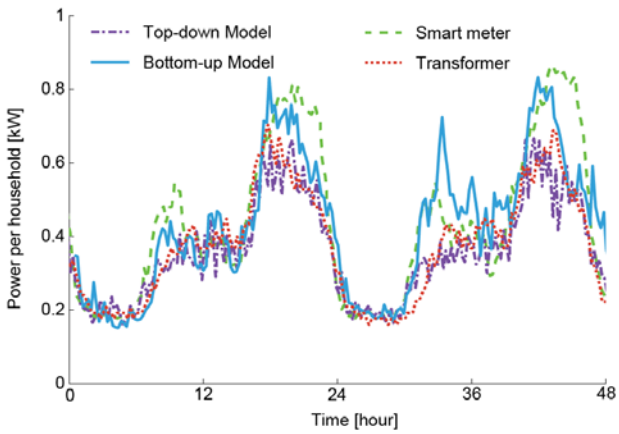

Figure 5: Two sample days from both models, smart meter data and transformer data.

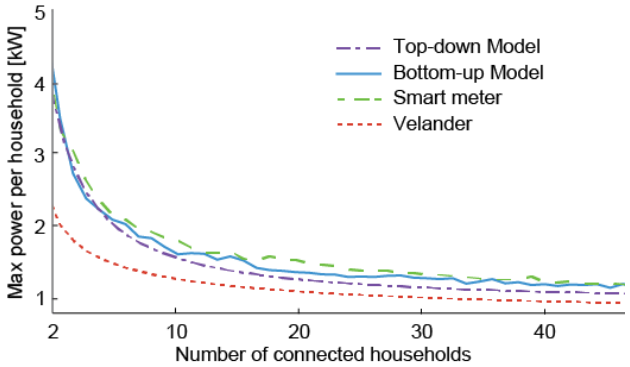

Figure 6: The peak load of the various models versus the number of aggregated households.

The ease of scenario implementation in the models is mostly dependent on the model structure. The top-down model contains black-box models (ANNs) in which the link between the input parameters and the resulting output variable is unknown, and cannot be altered to implement changes which arise from certain scenarios.
For the comparison on the computational speed, the load curve for 100 households over a year has been calculated and timed. The criteria for both models are shown in Table 2.

Table 2: Overview of the performance of the models, the peak load (1 perfect fit) and load curve RMSE (0 perfect fit) are assessed relative to smart meter data

Criteria	*Bottom-up*	*Top-down*	*Smart meter*
Peak load	0.96	0.84	1
Load curve	0.88	0.72	1
Local conditions	+/-	++	++
Scenario implementation	++	+	-
Computational speed	35.1s	0.4s	na

CONCLUSIONS

Two types of household load models have been developed and evaluated based on criteria important for household load planning: the peak loading, load curve, inclusion of local conditions, scenario implementation and computational speed. A purely statistical model based on local measurements can model the loading quickly for most network planning purposes. For the implementation of future scenarios or a more detailed peak load calculation a bottom-up model is more appropriate. Based on what kind of assessment an asset engineer needs to perform the quick top-down method is in most cases sufficient. Only for detailed scenario assessments the bottom-up model would be preferable. For the use in large scale grid planning scenario studies a combination wherein the bottom-up model is an input for the top-down model may be most appropriate.

REFERENCES

[1] L. G. Swan, V. I. Ugursal, 2009, *Modeling of end-use energy consumption in the residential sector: A review of modeling techniques*, Renewable and Sustainable Energy Reviews 13 (2009) 1819–1835.

[2] P.M. van Oirsouw, *Netten voor distributie van elektriciteit.* Arnhem: Phase to Phase B.V., 2011, pp. 85-86.

[3] EDSN - Energy Data Services Nederland, *Verbruiksprofielen, Profielen Elektriciteit 2014.* EDSN; 2014
Available: http://www.edsn.nl/verbruiksprofielen/..

[4] Anssi Seppälä, 1996, *Load research and load estimation in electricity distribution.* Ph.D. dissertation, Helsinki University of Technology, Espoo, Finland.

[5] B.R. Cobb, R. Rumí, A. Salmerón, 2012, *Approximating the Distribution of a Sum of Log-normal Random Variables.* Proc. of the Sixth European Workshop on Probabilistic Graphical Models, Granada, Spain, 2012

[6] S. Firth, K. Lomas, A. Wright, R. Wall, 2008, *Identifying trends in the use of domestic appliances from household electricity consumption measurements*, Energy and Buildings 40 (5) 926–936.

[7] W. R. Gilks, S. Richardson, D. J. Spiegelhalter, 1996 *Markov Chain Monte Carlo in Practice*, Springer US.

[8] Y. G. Yohanis, J. D. Mondol, A. Wright, B. Norton, 2008, *Real-life energy use in the uk: How occupancy and dwelling characteristics affect domestic electricity use*, Energy and Buildings 40 (6) 1053–1059.

[9] A. Druckman, T. Jackson, 2008, *Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model*, Energy Policy 36 (8) 3177–3192.

[10] M. Hart, R. de Dear, 2004, *Weather sensitivity in household appliance energy end-use*, Energy and Buildings 36 (2) 161–174.

[11] NMA, 2012, Allocatievoorwaarden gas, Tech. rep.

[12] I. Richardson, M. Thomson, D. Infield, A. Delahunty, 2009, *Domestic lighting: A high-resolution energy demand model*, Energy and Buildings 41 (7) 781–789.

[13] M. Stokes, M. Rylatt, K. Lomas, 2004, *A simple model of domestic lighting demand*, Energy and Buildings 36 (2) 103–116.

[14] J. Stephenson, B. Barton, G. Carrington, D. Gnoth, R. Lawson, P. Thorsnes, 2010, *Energy cultures: A framework for understanding energy behaviours*, Energy Policy 38 (10) 6120–6129.