

MATLAB Applications to Generate Synthetic Electricity Load Profiles of Office Buildings and Detached Houses

Daniel A. Brodén
School of Electrical Engineering
Royal Institute of Technology
Stockholm, Sweden
danbro@kth.se

Kaveh Paridari
School of Electrical Engineering
Royal Institute of Technology
Stockholm, Sweden
paridari@kth.se

Lars Nordström
School of Electrical Engineering
Royal Institute of Technology
Stockholm, Sweden
larsno@kth.se

Abstract—In this paper we present two MATLAB applications that generate synthetic electricity load profiles for office buildings and detached houses down to 1-minute resolution. The applications have been developed using App Designer, a MATLAB environment for application development. The applications are based on consumer load models for office buildings and detached houses published in previous research work. The aim of this paper is to present an overview of the application functionalities, code design, assumptions and limitations, and examples of their potential use in power system education and research. To the author’s knowledge these are the first applications which allow generating synthetic load profiles for office buildings and houses in practical and intuitive manner where building attributes can be easily configured.

Keywords—MATLAB, App designer, Applications, Consumer load models, Synthetic load data

I. INTRODUCTION

The electric power system is undergoing a paradigm shift where intermittent energy resources such as wind and solar power are increasingly being integrated into existing grids. This integration is disrupting the traditional centralized power production, transmission and distribution approach. In consequence, increases of imbalances in the power system are expected which has raised emphasis on controlling so called flexible loads. In short, these are electric loads that can easily be rescheduled or controlled with agreed, little or no impact to the particular consumer. Heating systems in residential or commercial buildings are examples of this. Many research papers in recent years focus on proposing optimization and control methods for flexible loads to cope with future challenges, often through simulations of futuristic scenarios. Papers [1], [2], [3] are few of many examples of this. To be able to perform similar type of studies, researchers are often faced with the lack of realistic load data. The high rate of smart meter roll outs in many European countries has allowed measurements to take place at the load segments. However, the access to this data for research remains poor and is often under the property of electric utilities. When provided, the data is typically in half-hourly or hourly resolution, depending on the structure of the electricity market of the issuing country. For example, in Sweden, the Nordic electricity market Nord Pool operates on hourly resolution, where utility customers have the right to request hourly billing, which in that case requires the utilities

to collect hourly data for those customers. Consequently, if provided, the data is on the hourly time scale at best.

An alternative solution is to use consumer load models to generate synthetic load data. For power system studies there are often situations where it is sufficient to have realistic variability of the aggregation of loads over time rather than precise measurements. One of the advantages of using models to generate load data is its configuration abilities. One could generate load data in lower resolution than the typical hourly measurements which could be useful for power system stability studies requiring second to minute data. One could also specify load attributes such as the area size or the number of occupants of residential buildings to capture the load profile for specific characteristics of consumers.

In this paper we present two MATLAB applications that generate synthetic electricity load data of office buildings and detached houses. These are typical loads considered as flexible and are the subject of many research studies in the topic of demand-response and demand-side management. The applications generate and display synthetic load profiles in hourly, half-hourly, quarterly or minute resolution. It uses consumer load models which were proposed and validated in [4] and [5]. The models are first order systems which have been proposed using a combined physical and behavioral approach. It should be noted that the validation showed good statistical resemblance with aggregated hourly measurements of buildings in Sweden which are further explained in [4], [5]. The applications generate data for up to 100 buildings, over a period of 1 to 365 days, for specific time resolutions, and select model parameter values (such as heating system capacities, number of occupants, building area size, and others). The resulting load profiles can be analyzed in aggregated or disaggregated form along with other exogenous variables such as outdoor temperature and solar radiation. The applications were developed using App Designer in MATLAB R2016b and are available for download on GitHub [6], [7] or through the MATLAB user community website [8].

II. GRAPHICAL USER INTERFACE

The graphical user interface for both applications are very similar. Screen captures from the applications are presented in Figure 1 and 2. In both cases the applications consist of

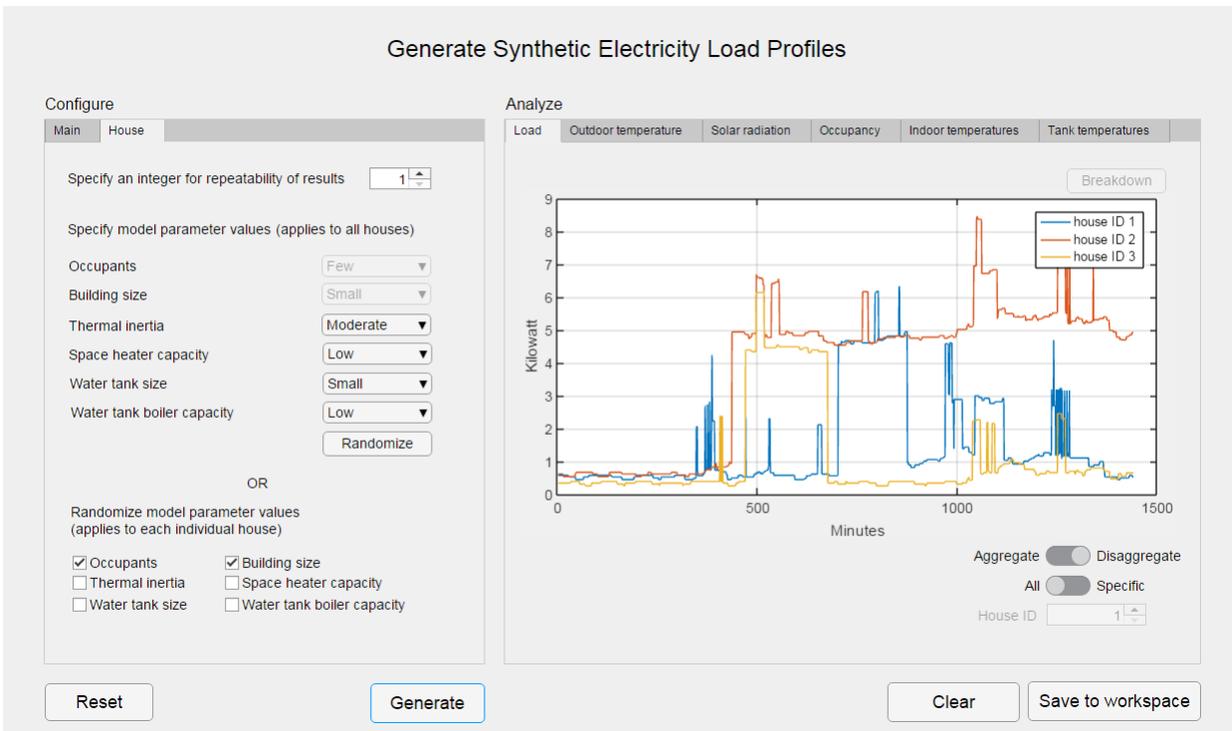


Fig. 1. Application screen capture: the generated load for three detached houses on minute resolution on Thursday, January 1st, 2015, in Stockholm, Sweden

a panel for configuring simulations parameters and one to analyze the generated outputs. These are denoted "Configure" and "Analyze" in the application windows.

A. Configuration panel

There are two tabs under the configuration panel: one where main configurations such as the number of buildings to simulate (1-100), the time period (1-365 days), and the time resolution (minute, quarterly, half-hourly or hourly) are specified. The other tab, where model configurations are specified, this includes values for select model parameters (1-3 options per model parameter). Additional configurations are displayed under the model configurations tab when data is being generated for more than one building. The user can tick check boxes for each model parameter to randomize its value for all individual buildings. If unchecked, the value originally selected will apply for all buildings being simulated. The model parameter values are categorical, for example, the number of occupants in the house is specified as 'Few', 'Average', 'Many' and are mapped to specific quantitative intervals in which they are randomized (more on this in section IV). Thus, to ensure repeatability of results, the user needs to specify an integer which sets the sequence of the random number generator. Pushing the reset button will return all user entries in the configuration panel to its original values. In Figure 1 the model parameter values for the number of occupants and building size is randomized for each of the three houses.

B. Analyze panel

There are multiple tabs under the analyze panel which display the generated outputs. Pushing the generate button

will generate the aggregated load profile along with other indigenous and exogenous variables to one of the tabs of the panel. In Figure 1 and 2 the load profiles displayed have been disaggregated as can be observed by the state of the corresponding switcher. Thus the load profile is displayed for each individual building. For the house loads one can observe varying profiles throughout the day with load spikes from around 400 minutes (corresponding to the early morning). For the office loads one can observe varying profiles throughout the day with a reoccurring pattern during typical work-hours (i.e., around 08:00-17:00). The loads can be further analyzed by pushing the breakdown button which will break down the load into the individual load components assumed by the consumer load models. Pushing it a second time will sum the individual load components back to its original state. Switchers and spinners can be further manipulated to analyze the data for particular buildings or building floors. This helps to further explain the shape of the different load profiles. Pushing the clear button will clear all the outputs from the analyze panel. Pushing the save to workspace button will save the outputs from the analyze tabs as a data structure to the MATLAB workspace for further use or analysis.

III. CODE DESIGN

App Designer imposes a code structure for its application development which has been depicted in Figure 3. An application generally consist of callback functions, utility functions, and properties.

A. Callback functions

The callback functions execute in response to a user interaction with one of the application components, such as

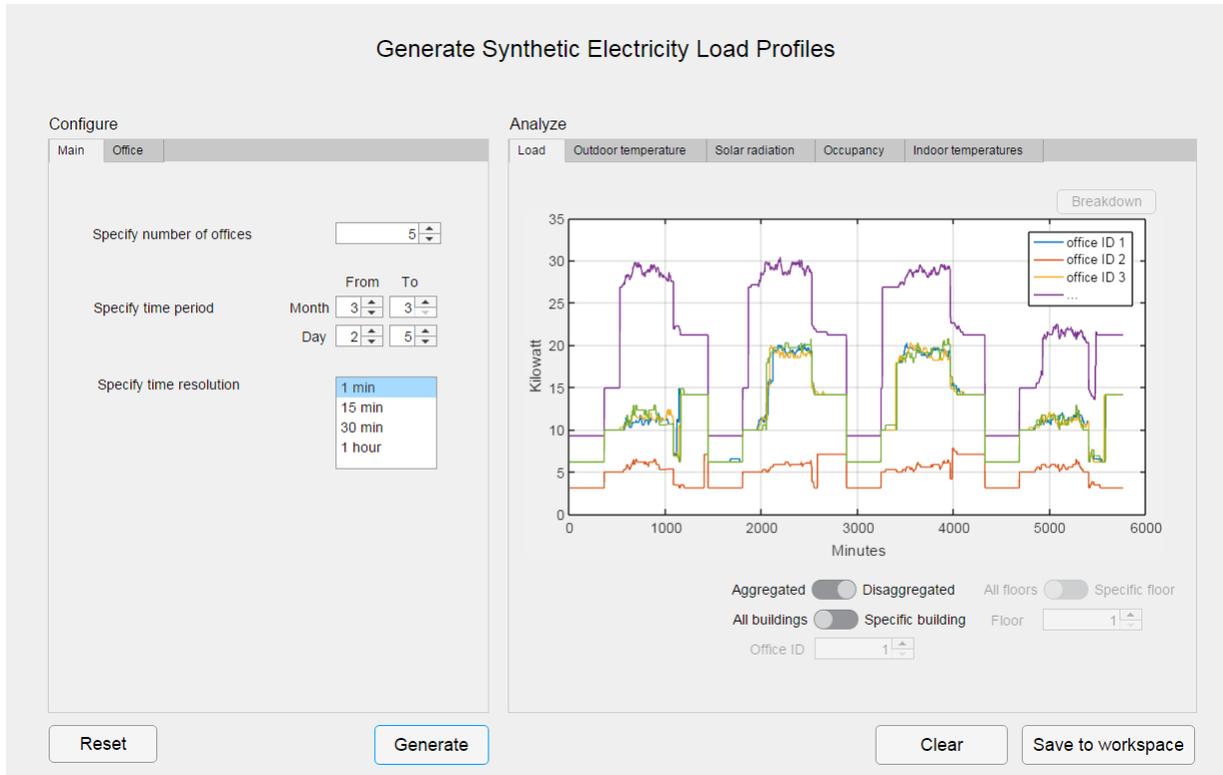


Fig. 2. Application screen captures: the generated load for five office buildings in minute resolution between Monday-Thursday, March 2nd-5th, 2015, in Stockholm, Sweden

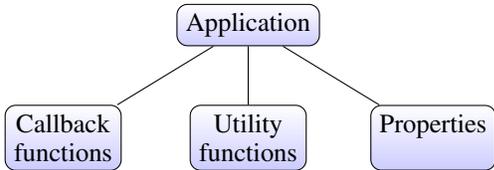


Fig. 3. The application code structure in App Designer (MATLAB R2016b)

pushing the generate button or changing the value of the spinner for the number of buildings to display data. Most switchers and spinners described in section II have value changed or changing callback functions associated to them.

B. Utility functions

The utility functions perform a task that is reused across the application. The office building application consist of two utility functions, one for generating the office model parameters, and the other to generate the Heating Ventilation and Air Conditioning (HVAC) power consumption and indoor temperatures for each building floor. The latter utility function correspond to load components assumed by the consumer load model in [4]. The house application on the other hand consist of three utility functions, one for generating the house model parameters, one for computing the space heating power consumption and indoor temperatures, and one for computing the domestic hot water power consumption and tank temperatures. The two latter utility functions correspond to load components assumed by the consumer load model in [5]. The results from the utility functions are later compiled in a callback function.

C. Properties

Finally, the branch referred to as properties, stores data and other variables introduced across the application. Both load applications use a default data set consisting of weather, occupancy, and behavioural-related power consumption variables (more on this in section IV).

IV. ASSUMPTIONS & LIMITATIONS

A series of assumptions are made to generate the load profiles for both applications. The model-related assumptions are meticulously described in [4], [5] and are outside the scope of this paper. However, assumptions on how the select model parameter values map to quantitative intervals and values, as well as their correspondence to variables from the original papers are presented in the following two subsections. Furthermore, both applications assume that the heating and cooling systems are thermostatically controlled where the heating systems are either consuming at a constant maximum or minimum capacity. For the thermostat control we assume a constant reference temperature of 20°C for space heating and cooling, 80°C for tank heating, with a dead-band in both cases of $\pm 1^{\circ}\text{C}$. For the house application we assume that measures are taken to maintain the indoor temperature when it is either above 27°C or below 15°C . Examples of measures include opening the windows to cool down the house if the outdoor temperature is less than 27°C or to activate additional space heating when the indoor temperature is below 15°C . The same assumption is made for the office load. For the water tank in houses we assume that occupants are showering with cold water to maintain the tank temperature when it drops

below 60°C (which corresponds to a recommended minimum temperature to avoid the occurrence of legionella bacteria [9]). Furthermore, we assume that office buildings have their own space heating systems where cooling is available throughout the day only when the daily mean outdoor temperature is above 15°C.

Some of the limitations of the applications include the maximum number of buildings (100) and the time period to generate data (365 days). In some configuration cases the time to display the generated data in the analyze panel can become cumbersome. The display time for 100 buildings and 365 days in minute resolution for the house load application resulted in approximately 20 minutes when generating data on an Intel Core i7-3687U CPU @ 2.10 GHz with 8 GB RAM. Another limitation is that the only distinction made in the occupancy patterns and activities in the buildings are between weekdays and weekends. This means that the effect on the load profile of typical national holidays such as New Years (January 1st), Christmas day (December 25th) or others are not considered.

A. Office load application

The model parameters, their values, and variable correspondence to the original papers are presented in Table I.

TABLE I. MODEL PARAMETER VALUES AND VARIABLE CORRESPONDENCE TO [4]

Model param.	Categ. value	Quant. values	Var. corresp.
Building floors	Few	1-3	N_{off}
	Average	4-6	
	Many	7-9	
Floor size	Small	440 sqm	A_{floor}
	Average	640 sqm	
	Large	840 sqm	
Occupants/floor	Few	5	$N_{occ, persons}$
	Average	15	
	Many	25	
Thermal inertia	Low	1200e3 min	C_{in}
		80e3 qtr	
		40e3 hhr	
	Moderate	20e3 hrs	
		3600e3 min	
		240e3 qtr	
	High	120e3 hhr	
		60e3 hrs	
		7200e3 min	
Heating/cooling capacity	Low	6 kW	Q_{hp}^{max}
	Moderate	10 kW	
	High	14 kW	

Most of the quantitative values are based on those originating from the validation data sets in [4]. Most of the values have been subtracted or added by a constant to map the remaining categorical values. The thermal inertia parameter is adjusted according to the time resolution specified, either in minutes, quarters, half hours or hours, abbreviated by min, qtr, hhr, and hrs in the tables.

B. House load application

The model parameters, values, and variable correspondence to the original papers are presented in Table II.

A similar mapping approach to the office application is used from quantitative values introduced in [5].

TABLE II. MODEL PARAMETER VALUES AND VARIABLE CORRESPONDENCE TO [5]

Model param.	Categ. value	Quant. values	Var. corresp.
Occupants/house	Few	1-3	β_{family}
	Average	4-6	
	Many	7-9	
Building size	Small	50 sqm	$A_{floor} \& A_{roof}$
	Average	100 sqm	
	Large	200 sqm	
Thermal inertia	Low	12000 min	τ
		800 qtr	
		400 hhr	
	Moderate	200 hrs	
		6000 min	
		400 qtr	
	High	200 hhr	
		100 hrs	
		3000 min	
Space heater capacity	Low	4 kW	Q_{heat}^{max}
	Moderate	6 kW	
	High	8 kW	
Water tank size	Small	150 L	V_{tank}
	Average	300 L	
	Large	600 L	
Water tank boiler capacity	Low	1.5 kW	P_{boil}
	Average	3 kW	
	Large	6 kW	

C. Default application data

A default data set is used for both of the applications to generate the load profiles. The data set consists of weather, occupancy and behavioural-related power consumption and heat gain variables (which are derived from the occupancy data). The weather data consists of outdoor temperature and solar radiation measurements from the Stockholm region in Sweden for the year 2015. The measurements are originally provided in hourly resolution and are linearly interpolated to support the different time resolution options provided by the applications. The occupancy data is simulated using a stochastic model based on non-homogeneous Markov chains proposed in [4], [5], [10]. To generate the occupancy and behavioural-related data, ten simulations are performed in minute resolution for 365 days and for each of nine different options of number of building floors (office load application) or number of occupants (house load application), resulting in a total of 90 simulations for each application. The data is stored as a three-dimensional array, where the first dimension corresponds to the number of minutes in 365 days, i.e., 525600. The second dimension is related to either the maximum number of building floors (office load application) or the maximum number of occupants (house load application), i.e., 9. The third dimension corresponds to ten stochastically generated occupancy profiles. Thus, when pushing the generate button the mapping for each building is performed as follows: the appropriate range is selected based on the time period specified, next, the appropriate column based on either the number of building floors (office load application) or the number of occupants (house load application), and finally, 1 out of 10 possible profiles is randomly selected based on the user-specified integer which sets the sequence of the random number generator. The behavioral-related power consumption variables such as appliance loads, or heat gains from appliance usage and occupancy, or energy drained from the water tank, are derived similarly using the approaches proposed in [4], [5], [10].

V. POTENTIAL USE

Two areas (not limited to) where these applications could be useful are described in the following subsections.

A. In power system education

Once installed, the applications are easily accessible within the MATLAB environment and can be used for educational purposes when the end-user segment is the subject of interest. The analyze panel allows the user to identify causal effects between the load profile and indigenous and exogenous variables. The control logic of the heating systems can also be analyzed by breaking down the load profiles into individual components and comparing them against the weather conditions of the specified period. The following two examples describe the analysis of generated load profiles for each of the applications:

1) *Example 1 — office load profile:* The load profile of an office building is generated in minute resolution for January 1st 2015 using default model parameter values. Figure 4 and 5 presents the load profile (broken down into components) and the corresponding indoor temperature of the building floors. One can observe variability in the appliance load during typical work-hours (7:00-18:00, i.e., 420-1080 minutes) which confirms expectations. The appliance load follows a constant base load for the remaining hours of the day which corresponds to the non-work hours. The indoor temperature for the two floors of the building drop by 1°C at around 13:00 and 18:00 (i.e., around 800 and 1100 minutes). At this point the HVAC load is activated and the temperatures begins to rise back to the set-point. This is observed by the steady rise in temperatures and the two steps in the HVAC load profile. One can deduce the effect of the outdoor temperature and occupancy in the office from the indoor temperature by observing the drop rate between minutes 0-500 (around 00:00-8:00) and 500-1000 (around 08:00-17:00). The drop rate during typical work-hours is lower due to rising outdoor temperature during midday compared to night-time and rise in occupancy due to employees coming as opposed to leaving work.

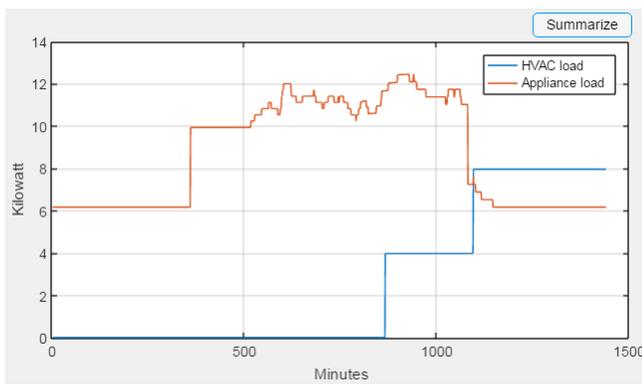


Fig. 4. Application screen capture: shows the load profile of a two-floor office building in minute resolution on Thursday, January 1st, 2015, in Stockholm, Sweden. The load has been broken down into HVAC and appliance load

2) *Example 2 — house load profile:* The load profile of a detached house is generated in minute resolution on January 1st 2015 using default model parameter values. Figure 6 and 7 presents the load profile (broken down into components) and the corresponding tank temperature in the house. One

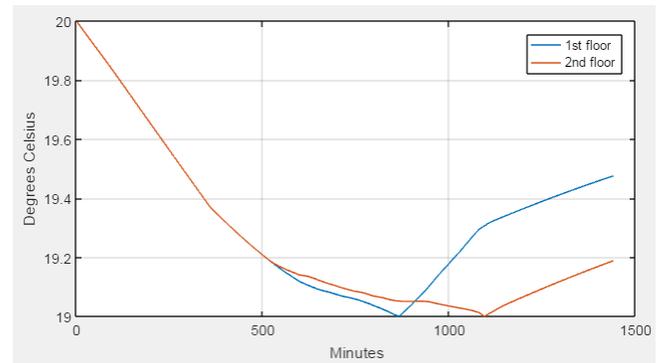


Fig. 5. Application screen capture: shows the indoor temperature in minute resolution for each floor of an office building on Thursday, January 1st 2015, in Stockholm, Sweden

should note that the indoor temperature has been omitted. One can observe peaks in the appliance loads at around 400-600 minutes (i.e., around 7:00-10:00) which typically corresponds to wake-up hours (before leaving for work). One can also observe a drop in tank temperature around this time, corresponding to a probable shower from the house occupants. The tank temperature quickly increases due to the assumption of thermostatic control. One can observe that the hot water load also peaks around this time to heat the water.

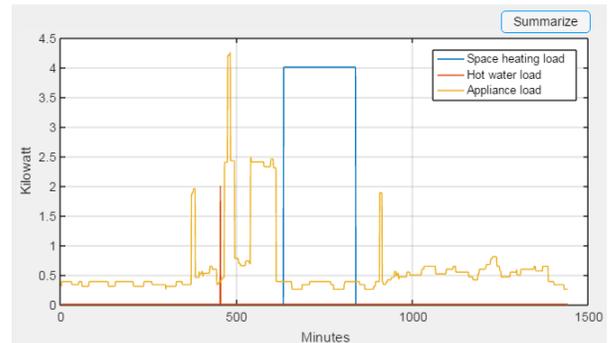


Fig. 6. Application screen capture: shows the load profile in minute resolution of a detached house on Thursday, January 1st 2015, in Stockholm, Sweden. The load has been broken down into space heating, hot water, and appliance load

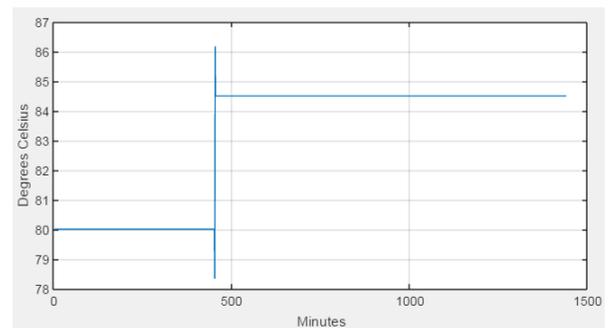


Fig. 7. Application screen capture: shows the tank temperature in minute resolution of a detached house on Thursday, January 1st 2015, in Stockholm, Sweden

B. In power system research

Applying classification and clustering algorithms on consumer load data with the purpose of identifying patterns or load shapes have become increasingly popular in research. Papers [11], [12], [13] are few of many examples of this. In Figure 8 we present an example with results from applying the k-means algorithm with 3 clusters, 100 times, on 100 load profiles generated from the house load application. The data is generated in hourly resolution for January 1st 2015 with randomized building size parameter value for each individual house. The purpose of the clustering is to identify the load shape with respect to the building size attribute of the houses.

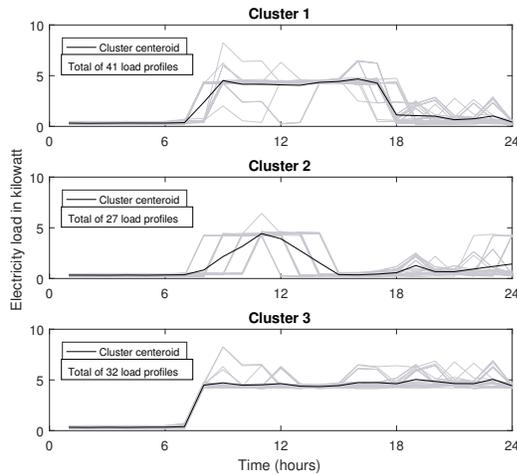


Fig. 8. Results from applying k-means with 3 clusters, 100 times, on 100 load profiles in hourly resolution using data generated by the house load application on Thursday, January 1st 2015, in Stockholm, Sweden

TABLE III. SHARE OF LOAD PROFILES PER CLUSTER WITH RESPECT TO BUILDING SIZE ATTRIBUTE

	Small	Average	Large
Cluster 1	10%	90%	0%
Cluster 2	100%	0%	0%
Cluster 3	0%	0%	100%

Table III presents the share of load profiles per cluster with respect to the building size attribute. For the 41 load profiles in cluster 1, 90% share the attribute of being 'Average'-sized houses and 10% being 'Small'-sized. In cluster 2 and 3, all 27 and 32 profiles respectively correspond to 'Small'-sized and 'Large'-sized houses. The cluster centroid in Figure 8 gives an indication of the load shape for each cluster. One can observe that 'Small'-sized houses have a load shape characterized by two peaks: one very distinctive at 11:00 and a less notable one at 19:00. 'Average'-sized houses have a load shape characterized by a step between 7:00 and 18:00. 'Large'-sized houses have a load shape characterized by a step from 7:00. From the clustering results we can conclude that the building size attribute of a house has a significant effect on the load shape for this particular time frame. Further analysis in the house load application shows that houses with larger building sizes tend to require more heating load which explains the step shape in the load profiles as the building size increases.

This example demonstrates one of many ways to apply classification and clustering algorithms using the data gen-

erated by the house load application. Similar studies can be performed using the office load application.

VI. CONCLUSIVE SUMMARY

We presented two MATLAB applications that generates synthetic electricity load profiles for office buildings and detached houses down to 1-minute resolution. The graphical user interface was described as well as the code design implemented in App Designer. The assumptions and limitations of the applications were described with references to the original consumer load models. The potential use of the applications were described with examples in the areas of power system education and research. Furthermore, the applications are available for download at [6], [7].

ACKNOWLEDGMENT

This work was funded through SweGRIDS, by the Swedish Energy Agency, and Vattenfall [14]. The authors would also like to acknowledge former colleague Dr. Claes Sandels who originally proposed the consumer load models used in the applications.

REFERENCES

- [1] Y. Zhou, Y. Yi, G. Cui, P. Jin, X. Guo and C. Yang, "Demand response control strategy of groups of central air-conditionings for power grid energy saving," 2016 IEEE International Conference on Power and Renewable Energy (ICPRE), Shanghai, 2016, pp. 323-327
- [2] B. Bhattarai; M. Levesque; B. Bak-Jensen; J. Pillai; M. Maier; D. Tipper; K. Myers, "Design and Co-simulation of Hierarchical Architecture for Demand Response Control and Coordination," in IEEE Transactions on Industrial Informatics , vol.PP, no.99, pp.1-1
- [3] I. Diaz de Cerio Mendaza, I. G. Szczesny, J. R. Pillai and B. Bak-Jensen, "Demand Response Control in Low Voltage Grids for Technical and Commercial Aggregation Services," in IEEE Transactions on Smart Grid, vol. 7, no. 6, pp. 2771-2780, Nov. 2016
- [4] C. Sandels, D. Brodén, J. Widén, L. Nordström, E. Andersson, Modeling office building consumer load with a combined physical and behavioral approach: Simulation and validation, Applied Energy, Volume 162, 15 January 2016, Pages 472-485
- [5] C. Sandels, J. Widén, L. Nordström, Forecasting household consumer electricity load profiles with a combined physical and behavioral approach, Applied Energy, Volume 131, 15 October 2014, Pages 267-278
- [6] Github repository for office load application, accessed 2017-06-05, <https://github.com/danbro/officeloadapp>
- [7] Github repository for house load application, accessed 2017-06-05, <https://github.com/danbro/houseloadapp>
- [8] MathWorks user community website, accessed 2017-06-19, <https://se.mathworks.com/matlabcentral/fileexchange/?term=type%3AApp>
- [9] Homepage of Health and Safety Executive (UK), accessed 2017-06-07, <http://www.hse.gov.uk/healthservices/legionella.htm>
- [10] J. Widén, E. Wäckelgård, A high-resolution stochastic model of domestic activity patterns and electricity demand, Applied Energy, Volume 87, Issue 6, June 2010, Pages 1880-1892
- [11] Joshua D. Rhodes, Wesley J. Cole, Charles R. Upshaw, Thomas F. Edgar, Michael E. Webber, Clustering analysis of residential electricity demand profiles, Applied Energy, Volume 135, 15 December 2014, Pages 461-471
- [12] Christoph F., Nicolay D., Tobias C., Clemens V.D., Lilia F-N., Cluster Analysis of Smart Metering Data, Business & Information Systems Engineering, Volume 4, 2012, Pages 31-39
- [13] Ali Al-Wakeel, Jianzhong Wu, K-means Based Cluster Analysis of Residential Smart Meter Measurements, Energy Procedia, Volume 88, 2016, Pages 754-760
- [14] Homepage of SweGRIDS, accessed 2017-05-22, <https://www.kth.se/en/ees/omskolan/organisation/centra/swegrids>