Generation of Flexible Domestic Load Profiles to Evaluate Demand Side Management Approaches

Gerwin Hoogsteen, Albert Molderink, Johann L. Hurink, Gerard J.M. Smit

Abstract—Various Demand Side Management (DSM) approaches have been developed in the last couple of years to avoid costly grid upgrades. However, evaluation of these DSM methodologies is usually restricted to a use-case specific example, making comparison between different DSM approaches hard. This paper presents a novel, open source, load profile generator to evaluate and compare DSM approaches. In addition to static load profiles for both active and reactive power, it also provides flexibility information for various classes of controllable domestic devices. Load profiles and flexibility information are generated using a simple behavioural simulation. The output data uses 1 minute intervals and incorporates device measurements. The generated profiles are in sound with the measurement data obtained in a field test on both the household level and aggregated neighbourhood level. The same dynamics, such as unbalanced loading and rapid power consumption fluctuations, are observed in the generated model.

Index Terms—Load modeling, simulation, load management, demand side management

I. INTRODUCTION

Various Demand Side Management (DSM) approaches have been developed in the last decade [1] to cope with larger quantities of intermittent renewable energy resources (e.g., photovoltaics (PV)) in distribution grids on the supply side and with the changing demand for electricity on the demand side (e.g., adoption of Electric Vehicles (EVs)). The optimization objectives within DSM include peak shaving and self-consumption of renewable energy, thereby reducing the stress on grid assets and the need for reinforcements. Without such a DSM system installed, overloading of these assets is likely to happen in the near future [2].

To avoid peak consumption, flexibility offered by devices has to be exploited to reduce these peaks. Therefore, it is essential to know what the available flexibility in a particular grid is. This availability is usually expressed in start times (at which a device becomes available), deadlines and the amount of energy a device demands. This data is generally not easily obtainable. Measurement data merely consists of the aggregated electricity demand of a household and therefore lacks information about the available flexibility. Even if individual devices are measured, it is still unknown what their deadline was and whether a certain minimum (lower) required amount of energy could have been sufficient to satisfy the end-user comfort constraints (e.g., a 50% state of charge for an EV).

This paper presents a novel, open source, artificial load profile generator (ALPG) developed specifically to test and compare the performance of DSM methodologies. The ALPG uses a generated occupancy profile for each household. Static, inflexible, load is derived from this occupancy profile as well as availability times for flexible devices. The resulting load profiles are given in discrete time intervals of one minute, providing both active and reactive power. Using only a limited number of input parameters, the ALPG can deliver customizable scenarios that reflect the characteristics of measurement data reasonably accurately.

The main contributions of this paper are as follows:

- Method to generate realistic, high resolution, artificial load profiles with flexibility specification;
- Comparison between generated profiles and measurements;
- Publication of an open-source generation tool.

The remainder of this paper is as follows: The next section discusses related work on artificial load profile generators. Section III explains how the ALPG works and what configuration options are available. Measurement data obtained in a field test is compared to the generated load profiles in Section IV. Section V concludes the paper and outlines suggestions for future work.

II. RELATED WORK

Predicting and creating energy demand profiles for residential areas is a well studied field within power engineering. Such demand profiles usually consist of the aggregated demand profile for a group of households, where the large number of households helps to obtain realistic profiles. Various load forecasting methods are classified by Alfares et al. in [3]. Jardini et al. [4] show that on this aggregated level, accurate forecasts for residential areas can be made. These forecasted profiles are used to design and size distribution grid assets, or to investigate how to trade on the energy market.

Such demand profiles are used to estimate the flexibility that the supply side has to offer. However, with the transition to renewable energy the demand for flexibility on the demand side is growing. Consumer electricity profiles that specify the available flexibility can be used to investigate in which way DSM methodologies are able to shape demand profiles. An overview of DSM methodologies is given by Siano [1].

Within research on residential smart grids and DSM methodologies, the focus and modelling is usually on a much finer scale: that of individual devices within households. This asks for load demand profiles that explicitly specify consumption patterns of individual devices and households. Obtaining household consumption measurements is not the
hardest part in this. These are available in field tests (e.g., [2]), or in online databases such as provided by Pecan Street inc. [5]. It is even possible to extract usage of individual devices, such as washing machines and EVs, from an aggregated power consumption profile provided by smart meters. The larger devices, suitable for DSM, can be extracted using a load disaggregation algorithm, such as the one proposed in [6].

Pipattanasomporn et al. have investigated the potential for DSM in US households using measurements in [7]. A total of six major devices, suitable for DSM programs, were identified and measurement data is provided to construct residential load profiles. Plugrat et al. [8] describe a novel load profile generator that uses psychological models for user behaviour. The generator combines a solid model with an extensive set of measurement of individual devices. However, it takes quite some time to set up a model due to the vast number of provided options. Lampropoulos et al. [9] notice that prosumer behaviour is an important parameter for load modelling within smart grid research. Kong et al. [10] propose a load profile generator that incorporates usage patterns derived from inhabitant simulations. Other available profile generators include [11] and [12].

The above mentioned approaches require a lot of user input for the purpose of evaluating a DSM approach. Another drawback is that these techniques don’t specify the exact degree of freedom these devices offer. Obtaining this data from a field test is tedious and costly. One would need to equip a large group of households with smart devices and measurement equipment. Another approach is to spread questionnaires to map the flexibility. Such an approach is taken by Stamminger [13], who outlines usage distributions for various devices through Europe.

Claessen et al. [14] used a generic way of specifying flexibility in their model to compare the performance of the TRIANA methodology [15] and the Intelligator methodology. This is a static model for the considered use-case, which cannot be applied directly to other scenarios and regions.

III. APPROACH

As all approaches mentioned above do not fulfil the stated requirements, we propose a new open source ALPG written in Python that enables researchers to evaluate DSM approaches and may be used to generate benchmarking sets, such that different approaches can be compared in a fair manner. An overview of the available classes in the code is given in Subsection III-A. Subsection III-B details the required parameters to generate profiles. A detailed description of the simulation is given in III-C. The resulting output and its applicability for DSM approaches is detailed in III-D.

A. Overview

The proposed profile generator uses a bottom-up approach that initially specifies for each household the available flexible devices, photovoltaic (PV) installations, batteries and the configurations of the inhabitants. The behaviour of individual inhabitants is simulated to obtain an occupancy profile. The occupancy profile is used to create consistent profiles for the devices, meaning that devices that require user interaction can only change their state when a person is at home. For flexible devices the usage patterns are generated using the household configuration and it is guaranteed that the start and end times are synchronized with the occupancy profile. The available flexibility is specified using four generic flexibility classes for devices: Timeshiftables, buffer-timeshiftables, buffers and curtable. The inflexible electricity profiles are categorised in the following categories: stand-by load, consumer electronics, lighting, inductive devices, fridges and other. These individual profiles are scaled in magnitude to reflect the annual electricity consumption. This makes it easy to alter the profile when external factors lead to a change in electricity usage, such as the adoption of LED lighting. Furthermore, power factors are assigned to obtain reactive power figures as well. These inflexible device categories are tabulated in Table I.

The simulation model consists of multiple classes that represent a part of the model. This makes the software flexible to be extended in the future. The following classes are implemented:

Neighbourhood: The Neighbourhood class contains information about the households in a residential area. Furthermore, it is in charge of distributing the available (PH)EVs, batteries and PV setups to these households.

Households: The Households class contains various types of households with corresponding Persons, Devices and Houses, including annual power consumption values. This class is used to collect data of all devices and is in charge of bookkeeping the Devices attached to the household.

Persons: This class defines the behaviour of a person by generating and storing default patterns, such as, wake up times and arrival times from work. It sets parameters for activities such as sporting days, washing days, home working days and the travel distance to work. An uncertainty variable is available to make the person more stochastic.

Devices: The generic flexibility device class is defined within this class to bookkeep the important input values that define the flexibility. It contains behavioural descriptions of other devices such as refrigerators, vacuum cleaners, lighting, etc. For these devices, individual devices were measured to obtain realistic values.

House: In the current version, the House class stores information about the orientation of the roof to determine the PV generation based on the position of the sun and recorded irradiation data for a given location.

An overview of the relation between different classes is given in Fig. 1. Other classes take care of handling and storing the output (see subsection III-D) such that it can be imported

![Fig. 1. Overview of the relation of the different objects to each other.](image-url)
in the DSM software. These classes can be overloaded and imported to write DSM software specific configuration files right away.

B. Input Data

The major problem for generating residential load profiles as observed in Section II is either the lack of configurability or the burden of collecting and analysing enormous amounts of data to create models. The latter one resulting in more accurate and sophisticated profiles being generated. However, such a degree of detail is not strictly required for the evaluation of DSM approaches. Instead, we focus on a minimal set of in general easily obtainable parameters and statistics (such as from surveys or census data) that result in a decent load profile.

Parameters for the model can be configured in the config.py-file, which provides options for the penetration of certain emerging technologies and household types in the neighbourhood. The following parameters must be provided:

- **Simulation parameters:** The number of days to be simulated and the randomization seed.
- **Penetration of emerging technologies:** This includes the percentage of households owning a Full EV (FEV), Plug-in Hybrid EV (PHEV), battery storage and rooftop mounted PV.
- **Power consumption of various devices:** Power ratings and, if applicable, battery capacities for these emerging technologies have to be specified. Most of these values can be obtained through data sheets. Consumption levels of demanding devices, such as induction stoves, have to be provided too.
- **Geographical location of the neighbourhood:** The geographical location to obtain sunrise and sunset times.
- **Household types in the neighbourhood:** the number and types of households are configured by selecting the number of each predefined household type. These predefined households are tabulated in Table II and include, amongst others, families and single households. These values are based on [16] and [20].
- **Predictability of persons:** The predictability of inhabitants can be configured to allow random events by chance. These events represent completely random actions or skipping actions from the standard behaviour pattern. This makes it possible to evaluate the robustness of DSM approaches to prediction errors.

Next to the parameters defined in the configuration file, other classes may be extended to better reflect local behaviour and scenarios. The Persons class can be inherited if personalities need to be changed or new personalities are required. This includes changing the default distributions for wake-up times for example. New devices can be derived from the Devices class, which contains the generic device subclasses and their flexibility parameters.

C. Simulation and Generation

The program runs by executing profilegenerator.py. This starts the initialization of the model and selection of values from predefined distributions, after which the simulation of households starts to obtain profiles and flexibility data.

**Generation:** The complete model is generated in the first stage of the ALPG using values from the input parameters. Various probability distributions are used to choose parameters for households, persons, devices, etc. For each individual household the availability of devices such as dryers and dishwashers is determined by probability, depending on the type of household. Furthermore, the annual power consumption for some of the uncontrollable load categories is chosen using a truncated Gaussian distribution, in which the mean value represents the average consumption per category for the type of household and the number of adults and children.

Within the household generation, the number of children and the age of each inhabitant is chosen using a bounded uniform distribution. All persons are stored in a list, in which the first person is always an adult. For this person, the driving distance to work is chosen using a Gaussian distribution for which the mean value can be configured. This value is zero for those without a job.

An occupancy profile is generated using a simple behavioural model. This model is configured using mean times for general events that change the state of a person to either active (being home), inactive (e.g., sleeping) or away (e.g.,

### Table I

**Uncontrollable device categories**

<table>
<thead>
<tr>
<th>Category</th>
<th>% of annual consumption</th>
<th>Power factor range</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standby</td>
<td>17 - 25</td>
<td>0.75 - 0.85 capacitive</td>
<td></td>
</tr>
<tr>
<td>Electronics</td>
<td>8 - 12</td>
<td>0.99 - 1.00 capacitive</td>
<td>TV, computer, stereo set, etc</td>
</tr>
<tr>
<td>Lighting</td>
<td>3 - 6</td>
<td>0.99 - 1.00 capacitive</td>
<td></td>
</tr>
<tr>
<td>Fridge</td>
<td>N/A</td>
<td>0.55 - 0.65 inductive</td>
<td>Duty cycle based simulation, power consumption in range of 50 to 120 W</td>
</tr>
<tr>
<td>Inductive</td>
<td>N/A</td>
<td>0.70 - 0.90 inductive</td>
<td>Ventilation system</td>
</tr>
<tr>
<td>Other</td>
<td>N/A</td>
<td>1.00</td>
<td>Cooking (microwave oven, oven, induction cooking), vacuum cleaning, ironing</td>
</tr>
</tbody>
</table>

### Table II

**Predefined household configurations**

<table>
<thead>
<tr>
<th>Name</th>
<th>Annual consumption</th>
<th>Persons (Adults)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SingleWorker</td>
<td>1610 - 2410 kWh</td>
<td>1 (1)</td>
</tr>
<tr>
<td>DualWorker</td>
<td>2560 to 4060 kWh</td>
<td>2 (2)</td>
</tr>
<tr>
<td>FamilyDualWorker</td>
<td>3460 - 7060 kWh</td>
<td>3 - 6 (2)</td>
</tr>
<tr>
<td>FamilySingleWorker</td>
<td>3460 - 7060 kWh</td>
<td>3 - 6 (2)</td>
</tr>
<tr>
<td>FamilySingleParent</td>
<td>2600 - 6200 kWh</td>
<td>2 - 5 (1)</td>
</tr>
<tr>
<td>DualRetired</td>
<td>2660 - 4060 kWh</td>
<td>2 (2)</td>
</tr>
<tr>
<td>SingleRetired</td>
<td>1610 - 2410 kWh</td>
<td>1 (1)</td>
</tr>
</tbody>
</table>
to work). The exact times for inhabitants are chosen using a truncated Gaussian distribution. Days with regular events, such as sporting activities, washing days and home working, are also selected.

After the individual households and the corresponding persons are established, the list of households is shuffled and assigned to a physical house in the grid. Next, the other devices with specific penetration levels are distributed over the households to control the penetration of these devices. Induction cooking and PV setups are randomly distributed over the households and the orientation of the roof is randomly chosen using a truncated Gaussian distribution within predefined bounds. The selected amount of battery storage systems is spread over the houses with a PV setup, such that these households can increase the self consumption of their PV production. The EVs and the PHEVs are distributed over the households with the largest commute distance.

**Simulation**: After the complete model is established, the simulation is started. A flow-chart of the simulation process is shown in Fig. 2. To avoid memory issues, each house is simulated individually and written into an output file. Since the behaviour of the inhabitants of the household is leading for interactions with devices, person behaviour is simulated first. During the weekends, the weekend schedule of persons is simulated to obtain their occupancy profile. A working day is simulated on working days, except on a day off, for which a weekend schedule is simulated. The exact activity event times are randomized. Other random events (e.g., shopping) are added to the occupancy profile. In the weekends there is a probability the family may leave, which synchronizes the individual occupancy profiles.

The static stand-by profile is always applied to the load profile. On top of this profile, the individual freezers and refrigerators are simulated. Electronic equipment, such as computers and televisions, are activated using a probability function, which depends on the active amount of inhabitants. Each person that changes its state to active can trigger a device to be turned on. In the event that a device is turned on, a random value between 0.7 and 1.3 is selected as weight for its power consumption. The resulting weighted consumption time series is scaled in magnitude during the post processing step to match the annual power consumption for electronics. As soon as a person leaves the house again, there is a probability that a device turns off again. The simulation ensures that no electronic devices are left on when nobody is at home.

Cooking profiles are simulated at the end of the afternoon or beginning of the evening, based on the occupancy profile. The highest probability is assigned to cook using the stove, but there are chances that the oven or microwave is used instead. Based on the chosen device for cooking, the simulation decides whether the exhaust hood is turned on or not. Coffee makers and (electric) kettles are turned on randomly when at least one person is active. Lighting is turned on approximately 30 minutes before sunset. Power consumption for ironing and vacuum cleaning is added approximately once per week.

Based on the occupancy profile, available devices and preferred washing days, the start times of the washing machine are chosen. A measured profile from washing machines is available, which determines the minimum runtime and therefore the minimum end time. The chosen deadline for these controllable devices is usually later than this minimum end time and is chosen such that a person is at home. The same process is used for the dishwasher. EVs are activated as soon as their owner gets back home from work, unless a random event is scheduled later in the evening. The car has to be charged before this person leaves to work the next day. The required charge is based on the driving distance with an average energy consumption of 1 kWh per 5 km, with a variation to this value to reflect driving conditions and behaviour. The value is slightly increased during the winter months and reduced during the summer, to reflect the effect of different temperatures on the battery performance as shown in [17]. Other random trips, such as family outings, are included in the EV usage pattern. The simulation ensures that no start and end times overlap and that all deadlines can be met by the devices.

**Post processing**: When all usage patterns are determined, the static load profiles are post-processed. First, the lighting, stand-by and electronics profiles are rescaled in magnitude to reflect the annual consumption for these consumption categories. Then white noise is added to the time series to mimick the fluctuations caused in the profile due to voltage and small load variations. A similar rescaling is applied to all categories to obtain a reactive power profile.

**D. Output Data**

The output consists of two major parts: The flexible part and the inflexible part. The latter one consists of a comma separated values (csv) file that specifies the average power consumption in watts for each minute for each household. The rows represent the minute intervals in ascending order.
and each column represents a household in ascending order. Likewise, a csv file with reactive power in var is generated.

The flexible devices are split up in generic flexibility classes. We feel that these classes cover any pure electrical device. This makes it convenient to use a limited amount of device agents and corresponding optimization algorithms to target a broad spectrum of devices. This approach is used in EF-Pi (previously known as FPAI) [18], which provides an interface to connect various DSM methodologies and controllable devices through a generic interface.

The following flexibility classes as used in the TRIANA DSM methodology [15], with a direct link to the EF-Pi equivalent:

- **Timeshiftable**: e.g., washing machines, dryers and dishwashers. Real measurement data is used to obtain a static consumption profile in watts. The output is given in a list of jobs that have both an arrival time and deadline in seconds. The profile should be finished before the deadline and not started before the arrival time of each job. The EF-Pi equivalent for this class is the Timeshifter.

- **Buffer-timeshiftable**: e.g., EVs. The flexibility is specified by jobs with a start time, deadline (both given in seconds) and required energy demand in watt-hours. These devices have a fixed maximum power consumption in watts and buffer capacity in watt-hours. The equivalent EF-Pi device class is the Buffer.

- **Buffer**: e.g., a battery or hot water buffer. The specification gives a maximum power consumption or production level in watts and capacity in watt-hours. The equivalent class in EF-Pi is the Buffer.

- **Curtailable**: e.g., PV. This class specifies a fixed consumption or production profile and the amount of power that can be curtailed. The equivalent class in EF-Pi is the Uncontrollable.

Real measurement data for timeshiftable devices is provided. Other DSM methodologies found in literature usually use comparable input data as constraints. In previous work, a direct comparison between TRIANA and Intelligator was made using the Flex street use-case [14] which uses similar descriptions of flexibility.

The ALPG is able to provide output data using minute based intervals. This is enough to see short term peaks and dynamics, resulting from devices such as kettles and makes it possible to analyse the effect of using longer term schedules, usually 15 minute intervals, on the resulting profile when applying real-time control. This is of particular interest for islanded grids that are (in part) controlled by DSM methodologies.

### IV. Results

To evaluate the profile generator we make a comparison between a generated profile and measurement data obtained in Lochem [2]. Over one year of measurement data from transformers and households was obtained from this test site. From the measurement data, we took a measured feeder with 81 connected households. Census data [19] was used to create a model of the households. Induction cooking, PHEVs and rooftop PV setups were introduced according to survey data [20].

Fig. 3 shows the annual load duration curve for both the generated data and the measurement data. There are still some differences observable in the graph. Especially the static standby power usage from the ALPG is too flat. The average annual electric energy consumption per household is quite similar, with a measured value of 3144 kWh versus 3250 kWh for the generated model. Fig. 4 shows a comparison of the data for a single day. Both graphs show fast fluctuations. The MAE (Mean Absolute Error) for the minute to minute fluctuations is 157W. The overall reactive power consumption differs from the measurement data, but follows an overall trend, which results in a comparable power factor, around 0.94 during the night and is close to 1.00 during the morning and evening.

On the single household level, the generated profiles show less fluctuations than the measured data. This is to be expected from the approach taken, but it shows the same peak and standby loads. At the neighbourhood level, this behaviour ensures that the voltage levels and fluctuations obtained through load-flow calculations match the measured values. The simulated voltage levels for this neighbourhood are generally in the range of 230V to 240V during the afternoon, with dips down to 225V in the evening. These values match the fluctuations observed in the field test. Unbalanced loading of the three phases is observed in the model, with the model topping at 15 kW deviation from the average load and the measurements resulting in a deviation of 16 kW from the average load.

Fig. 5 shows a futuristic setting with the resulting cumulative power profile for a group of 81 households during winter. In this example 20% of the households own a full EV, 20% own a PHEV and 40% of the households have PV. Furthermore all washing machines and dishwashers are controllable.

![Fig. 3. Annual electricity duration curve for a neighbourhood of 81 households in Lochem.](image)

![Fig. 4. Neighbourhood active power consumption for one day, depicting measurement data and artificial data. The two upper lines depict the active power load in kW, the two lower lines show the reactive power consumption in kvar.](image)
The authors would like to thank Dutch distribution system operator Alliander for supplying grid data. This research is conducted within the EASI project (12700) supported by STW and Alliander and within the IPIN program Intelligent net in duurzaam Lochem, supported by the Netherlands Enterprise Agency.

REFERENCES