Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model

Francesco Lombardi, Sergio Balderrama, Sylvain Quoilin, Emanuela Colombo

DOI: https://doi.org/10.1016/j.energy.2019.04.097

Published in: Energy

Received Date: 27 July 2018

Revised Date: 22 February 2019

Accepted Date: 16 April 2019

Please cite this article as: Lombardi F, Balderrama S, Quoilin S, Colombo E, Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model, Energy (2019), doi: https://doi.org/10.1016/j.energy.2019.04.097.

Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model

Francesco Lombardi[a], Sergio Balderrama[b,c], Sylvain Quoilin[d,b], Emanuela Colombo[a]

[a] Politecnico di Milano, Department of Energy, Milan, Italy

[b] University of Liège, Department of Mechanical and Aerospace Engineering, Liège, Belgium

[c] Universidad Mayor de San Simon, Cochabamba, Bolivia

[d] KU Leuven, Department of Mechanical Engineering, Geel, Belgium

Corresponding author: F. Lombardi, e-mail address: francesco.lombardi@polimi.it

Abstract

Energy access projects in remote off-grid areas would benefit from the adoption of a multi-energy system perspective, addressing all energy needs – not only lighting and power appliances, but also water-heating and cooking – by means of a mix of energy vectors. However, multi-energy analyses in remote areas are hindered by a lack of models allowing for the generation of multi-energy load profiles based on interview-based information characterised by high uncertainty. This study proposes a novel open-source bottom-up stochastic model specifically conceived for the generation of multi-energy loads for systems located in remote areas. The model is tested and validated against data obtained from a real system, showing a very good approximation of measured profiles, with percentage errors consistently below 2% for all the selected indicators, and an improved accuracy compared to existing approaches. In particular, some innovative features – such as the possibility to define and modulate throughout the day appliances' duty cycles – seem to be determinant in marking a difference with previous approaches. This might arguably be even more beneficial for case studies characterised by a larger penetration of appliances that are subject to complex and unpredictable duty cycle behaviour.

Keywords: load profile; energy demand; multi-energy system; off-grid; rural areas

1. Introduction

The transition towards cleaner and more sustainable energy systems requires a comprehensive approach, going beyond the unique aspect of renewable energies penetration into the electricity generation mix and focusing more on the integrated use of multiple energy vectors and storage options that may meet different local needs [1]. In fact, adopting what has been defined by Mancarella et al. [2] a “Multi-Energy System” (MES) configuration ensures an enhanced flexibility [3] and has the potential to unveil synergies that would remain unexploited within a single-vector (typically electricity) perspective [4]. This concept is particularly relevant in the framework of energy access projects in remote rural areas, which are often off the grid: in such contexts, non-electric energy needs (such as space heating, water heating and cooking) are rarely taken into account within energy planning strategies, even though they represent the major share in the total final energy consumption and are traditionally satisfied by non-affordable, non-clean and non-safe sources [5]. One of the reasons behind the lack of multi-energy analyses in those contexts is the intrinsic complexity linked to the accurate assessment of local needs and the consequent generation of appropriate load profiles, as uncertainties about the energy demand depend on different factors and are not easily predictable [6], especially for non-electric types of demand. Indeed, the estimation of load profiles for remote off-grid systems often entails the reliance on interview-based information, which are used – in the absence of measured data – as an input for bottom-up models generating synthetic profiles [7].

In the framework of residential load curves modelling, stochastic approaches are commonly employed to reproduce unpredictable random consumers’ behaviour [8]. However, despite the abundance in the literature of models based on such approaches, those are usually conceived for on-grid residential buildings of industrialised countries, where the availability and detail level of the data is higher. For instance, Widén et al. [9,10] rely on activity diaries collected by Statistics Sweden as a basis to define the activity transition probabilities of a Markov-chain model of households’ activity patterns. Such diaries provide precise data about the daily activities – on 5-min intervals – of a large number of households in different settings. An analogous approach is adopted by Tsagarakis et al. [11] and Collin et al. [12], who draw on detailed time-use surveys for the UK to build a Markov-Chain model of activity patterns, also complemented by databases regarding power profiles per appliance and device ownership statistics.

Similarly, Fischer et al. [13] and McKenna et al. [14] define transition probabilities within their stochastic models based on data provided, respectively, by national time-use surveys for Germany and UK, with similar temporal resolutions and detail. The UK time-use survey datasets are also employed by Richardson et al. [15] to elaborate activity probabilities with traditional statistical techniques, which are subsequently compared against random numbers to determine the occurrence of a given switch-on event. The same data are used as a basis, together with other complementary statistical datasets, by Good et al. [16] to define activity profiles which are then coupled with detailed appliances operation profiles and building models to generate synthetic multi-energy demand profiles. Finally, Marszal-Pomianowska et al. [17] rely on a series of detailed datasets (such as 1-hour yearly profiles of household electricity consumption per appliance group) specific to the Denmark context to build a load demand model based on statistically-derived occupancy profiles and appliances characteristics, highlighting also how the highly context-specific nature of previously-developed similar models prevents their application in different circumstances.

Conversely, when planning an intervention in remote areas, the detailed input parameters that are requested by those approaches are hardly available with an acceptable degree of accuracy. Interview-based information are typically much less detailed than time-use diaries and inherently affected by significant degrees of uncertainty, which prevent a high-resolution characterisation of the activity patterns [18]. Furthermore, models conceived for remote areas need to be tailored for *energy planning* purposes rather than for *load forecasting* [19], i.e. they need to be flexible and adaptable to appliances and loads that might not even be present in the target community when interviews are held. Accordingly, even assuming that more detailed and highly-resolved surveys could be hold – which is typically not a realistic assumption [18] – it would be impossible to capture statistics and probabilities associated with non-yet existing activity patterns (such as switching from traditional biomass cooking to various electric cooking appliances such as boilers, ovens, hobs, etc.), which are nonetheless essential from a policy maker perspective. To this regard, a first attempt to extend the applicability of stochastic load profiles modelling to systems located in remote areas has been recently made by Mandelli et al. [19], by developing a bottom-up stochastic approach specifically conceived for the reliance on interview-based information. In practice, this approach (implemented in the software LoadProGen) embeds the inherent uncertainties of the input data in the modelling process by randomly varying the parameter related to the self-declared activity patterns (such as time frames and duration of appliances use) as well as parameters referred to random consumers' behaviour (such as switch-on times).

Nonetheless, none of the above mentioned approaches is capable of fully responding to the needs of energisation projects in remote rural areas, which require not only to deal with the inherent uncertainty and low detail of input data, but also to extend the range of modelling possibilities to non-electric loads, like water-heating or cooking. Those, in fact, would require a further degree of stochasticity in terms, for example, of random variation of the nominal power of each appliance or of the characteristics of its duty cycles, as recently demonstrated by Lombardi et al. [20]. Moreover, energy planning in remote areas requires a highly flexible and customisable modelling approach, in such a way to ensure its applicability to a wide range of contexts.

This study discusses the design and validation of a bottom-up stochastic model for the generation of high-resolution multi-energy load profiles for energy systems located in remote areas. The model builds on the concept proposed by Mandelli et al. [8] for the use of interview-based input data, but proposes an expanded stochastic approach with an increased degree of stochasticity. It is specifically designed for the characterisation of multi-energy needs. In addition, the scientific community is increasingly recognising the necessity of an open-source philosophy for models that need to cope with high degrees of uncertainties and subjectivities, in order to ensure transparency, testability and knowledge transfer [21]. The model is therefore implemented in a Python environment and is released as open-source software. This also enhances the degree of customisability and adaptability, in line with the previously identified model requirements. The model is freely accessible as "*Remote-Areas Multi-energy systems load Profiles*" (RAMP) from the GitHub repository: https://github.com/SESAM-Polimi/RAMP.

2. Methods

From a conceptual point of view, RAMP is based on three main layers of modelling, namely: i) the *User type*; ii) the *User*; and iii) the *Appliance* layers (Figure 1).

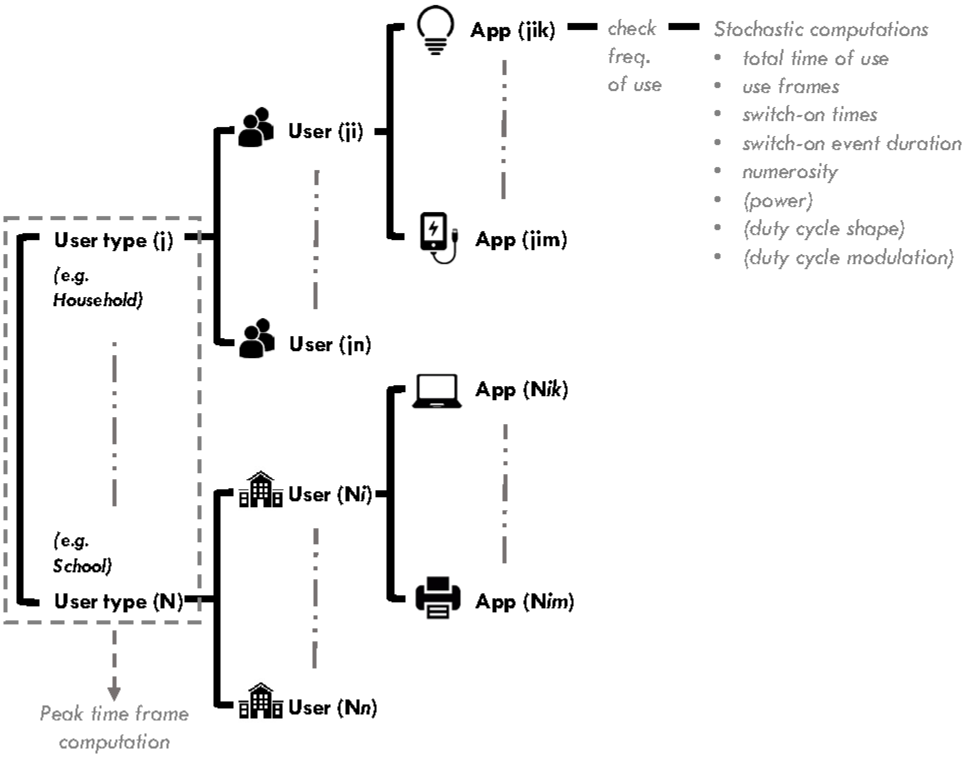

Figure 1 – Graphical sketch of the modelling layers constituting the simulation process

The higher layer consists in the definition of a set of arbitrary *User types* (e.g. Household, Commercial activities, Public offices, Hospitals, etc.), whose level of discretisation depends on the modeller's needs; for instance, when more precise information is available, a "Households" *User type* may be further subdivided by income classes or building type. Each *User type* is subsequently characterised in terms of the number of individual *Users* associated to that category (second layer) and in terms of *Appliances* owned by each of those *User*s (third layer). As shown in Figure 1, the three-layer structure allows to independently model the behaviour of each jik-th *Appliance*, so that each individual ji-th *User* within a given i-th *User type* will have a unique an independent load profile compared to the other *Users* of the same *type*. The aggregation of all independent *User* profiles ultimately results in a total load profile, which is uniquely generated at each model run. Multiple model runs generate different total load profiles, reproducing the inherent randomness and unpredictability of users' behaviour and allowing to obtain a series of different daily profiles.

All the inputs required to run the model are summarised in Table 1, and consist of information that can be obtained from common field surveys, in analogy with and expanding those defined by Mandelli et al. [19].

User type* and *Users	
$\boldsymbol{User\ type_j}$	Name of the *User type* (e.g. “households”, “commercial activities”, etc.)
$\boldsymbol{n}$	Number of $\boldsymbol{Users_{ij}}$ (for $\boldsymbol{i = 1{:}n}$) within $\boldsymbol{User\ type_j}$
Appliances	
$\boldsymbol{Appliance_{jik}}$	Name of the *k-th Appliance* associated with the *j-th User type* and the *i-th User*
$\boldsymbol{m_{jik}}$	Numerosity of $\boldsymbol{Appliance_{jik}}$ (e.g. numerosity of “indoor light bulbs”)
$\boldsymbol{P_{jik}}\ [\boldsymbol{W}]$	Power absorbed by a single item of $\boldsymbol{Appliance_{jik}}$ (i.e. assuming numerosity = 1)
$\boldsymbol{tot\_use_{jik}}\ [\boldsymbol{min}]$	Total time of use of the $\boldsymbol{Appliance_{jik}}$ in a day
$\boldsymbol{t\_min_{jik}}\ [\boldsymbol{min}]$	Minimum time that the $\boldsymbol{Appliance_{jik}}$ is kept on after a switch-on event
$\boldsymbol{\delta_{t\_min,jik}}$ [%]	Percentage random variability applied to $\boldsymbol{t\_min_{jik}}$
$\boldsymbol{use\_frames_{jik}}$	Time frames in which a random switch-on of $\boldsymbol{Appliance_{jik}}$ can occur
$\boldsymbol{\delta_{frames,jik}}$ [%]	Percentage random variability applied to $\boldsymbol{use\_frames_{jik}}$
Appliances’ optional attributes	
$\boldsymbol{cycle_{jik}}$	Duty cycle of $\boldsymbol{Appliance_{jik}}$ (up to 3 per appliance)
$\boldsymbol{\delta_{cycle,jik}}$ [%]	Percentage random variability applied to the duration of the segments composing $\boldsymbol{cycle_{jik}}$
$\boldsymbol{cycle\_mod_{jik}}$	Association between time frames and different duty cycles
$\boldsymbol{frequency_{jik}}$ [%]	Weekly frequency of use of $\boldsymbol{Appliance_{jik}}$ (<100% for “occasional-use” appliances)
$\boldsymbol{fixed\_num_{jik}}$	Constraint for all the $\boldsymbol{m_{jik}}$ appliances to always switch-on simultaneously
$\boldsymbol{\delta_{P,thermal}}$ [%]	Percentage random variability applied to $\boldsymbol{P_{jik}}$, conceived for thermal appliances

Table 1 – Summary of the input data required by the model.

2.1. Core stochastic algorithm

From a mathematical point of view, the stochastic algorithm that constitutes the core of the model (without including the *Appliances*’ optional attributes mentioned in Table 1) is articulated in the following steps:

1. identify the expected peak time frame;

2. for each $User\ type_j$, for each $Users_{ij}$ and for each $Appliance_{jik}$, check if the appliance is used based on the weekly frequency of use ($frequency_{jik}$). If not, ignore the appliance; otherwise, compute:

   a. the randomised total time of use $tot_{use\,jik}$;

   b. the randomised vector of time frames in which the appliance can be on $use_{frames\,jik}$;

   Subsequently:

   c. compute a random switch-on time (with random switch-on even duration ≥ $t_{min\,jik}$) within the available use frames;

   d. compute the randomised power required by the appliance for the switch-on event under consideration P_{jik}.

   e. compute the actual power absorbed by $Appliance_{jik}$ during the switch-on event considering a random numerosity in the range $[0, m_{ijk}]$.

   Repeat steps 2.c – 2.e until the sum of the durations of all the switch-on events equals the randomised $tot_{use\,jik}$;

3. Aggregate all profiles in a total load profile.

The identification of a peak time frame allows differentiating between off- and on-peak switch-on events, which are associated with different probability distributions for the computation of the random numerosity. To this end, a theoretical peak time frame is identified, as proposed in [19], as the time frame associated with the maximum load in a virtual total load profile resulting from the fictitious assumption that each *Appliance* is always switched-on with maximum power and numerosity during all of its potential time frames of use. Within such theoretical peak time frame, a unique *peak time* (t_{peak}) is hence randomly sampled with uniform distribution. Finally, an actual *expected peak time frame* is defined as per Equation 1.

$$peak\ time\ frame = \left[t_{peak} - k, t_{peak} + k\,\right] \tag{1}$$

Where k is the product of a random sampling with normal distribution around t_{peak} and standard deviation equal to $t_{peak} \cdot \delta_{peak}$. By default, δ_{peak} is set to 15% of t_{peak}, but it represents a potential calibration parameter that allows to modulate the extension of the peak time frame and may serve to simulate, for instance, a different social behaviour during holidays or weekends.

Peak-load periods correspond to periods in which a large share of *Users* is interested by intensive activity patterns and when, consequently, they are more likely to switch-on multiple *Appliances* of the same kind (e.g. “Households” might be likely to switch-on multiple indoor lights simultaneously in the evening, when they also cook, watch TV, etc.). To this regard, the model acts on the modulation of the “coincident numerosity” factor, defined by Equation 2.

$$f_{coinc\_num} = \frac{m_{ON,ijk}}{m_{ijk}} \tag{2}$$

Where $m_{ON,ijk}$ represents the numerosity of appliances that are simultaneously switched on during a switch-on event related to $Appliance_{ijk}$. Such factor can assume values ranging from $\frac{1}{m_{ijk}}$ to 1. During off-peak periods, $m_{ON,ijk}$ is randomly chosen based on Equation 3, i.e. by relying on a uniform distribution. During peak-load periods, conversely, $m_{ON,ijk}$ is randomly chosen based on a Gaussian distribution (Equation 4).

$$off\ peak\text{:}\ m_{ON,ijk} = max\{1, unif(0, m_{ijk})\} \tag{3}$$

$$on\ peak\text{:}\ m_{ON,ijk} = max\{1, norm(\mu_{\%} \cdot m_{ijk}, \sigma_{\%} \cdot m_{ijk})\} \tag{4}$$

Where the parameters $\mu_{\%}$ and $\sigma_{\%}$ are set by default so as to have, respectively, a mean value of the on-peak distribution that is the average of the range $[0, m_{ijk}]$ and a standard deviation that reaches the extremes of the range. Indeed, $\mu_{\%}$ and $\sigma_{\%}$ represent two further calibration parameters that can be manipulated by the modeller to reproduce behavioural patterns that are typical of its context of application, as well as to force the model towards the generation of “extreme” profiles, which may be required by robust optimisation tools [22].

2.2. Optional stochastic attributes

As shown in Table 1, RAMP offers the possibility to define several optional *Appliances*’ attributes, which allow to further enhance the customisability and the stochasticity of the model.

2.2.1. Modular duty-cycles and cooking cycles

Key optional attributes are those allowing to model pre-defined duty cycles and to modulate (if needed) the behaviour of such cycles throughout the day. For instance, a pre-defined duty cycle may be set to reproduce the behaviour of a fridge; however, considering that actual fridge's cycles are not fixed but rather dependent on the temperature and on user's activity patterns [23,24], different duty cycles (e.g. standard, intensive, etc.) can be modelled and associated with different time frames to follow the variation of such parameters during the day (Figure 2). Alternatively, duty cycles segments can be allowed to randomly vary within a user-defined range, to reproduce the behaviour of highly random and subjective tasks, such as cooking (Figure 3).

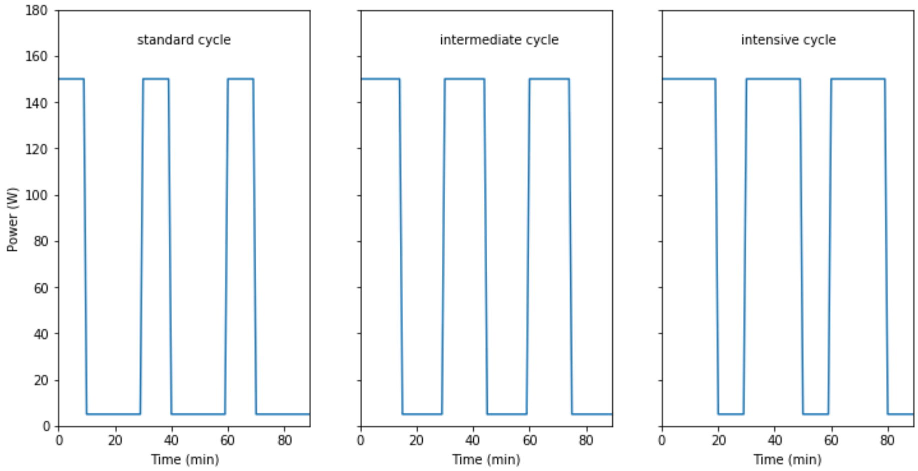

Figure 2 – Example of duty cycle modulation throughout the day for a fridge.

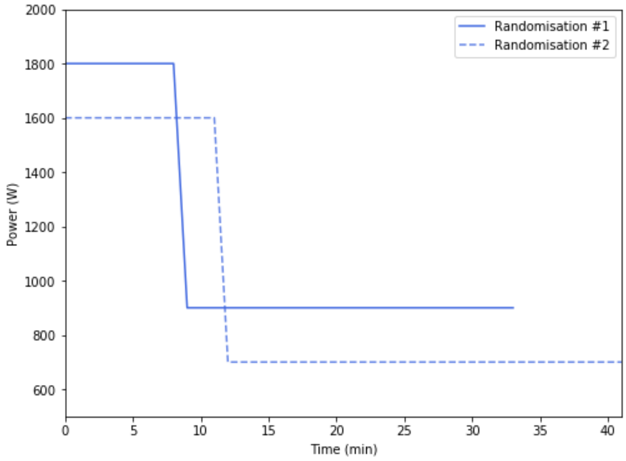

Figure 3 – Example of two different randomisations of a cooking cycle (in this case representing a boiling task followed by a simmering period).

2.2.2. Frequency of use

It is also possible to mark *Appliances* as "occasionally-used": in this case, the latter will be included in the set of *Appliances* that the i-th *User* will switch on during the day only conditionally to a random probability check (Equation 5), independently evaluated for each *User*.

$$if\colon\ frequency_{ijk} > unif(0{,}1) \rightarrow\ \exists\ Appliance_{ijk} \tag{5}$$

As a result, on a given day (i.e. a single model run) some of the *Users* of a given *type* may use them, while others may not; this functionality is conceived to reproduce the real patterns of use of appliances such as irons or mixers, and strengthens the unique random characterisation of each individual *User*. Figure 4 shows an example of different daily load profiles for a single household – owning iron and using it with a frequency of 3 days a week – over a 7 days period; the appliance is used only occasionally, and its relative weight on the weekly average profile is thus opportunely represented.

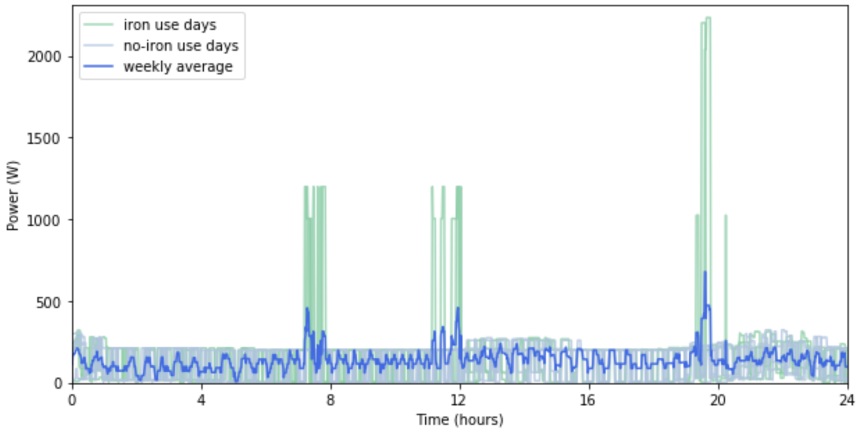

Figure 4 – Example of 7 different stochastic daily profiles for a single household using iron with an average frequency of 3 days a week, modelled by the "occasional-use" attribute. Some of the stochastic daily profiles (in light green) include iron use, while others (in grey) do not.

2.2.3. Thermal appliances and random power regulation

A special functionality is included in the model to better simulate the behaviour of thermal appliances. Those, in fact, are typically characterised by a high degree of variability in terms of absorbed power, which is a function of subjective and random preferences, for example in terms of hot tap water temperature. Such variability is embedded in the model by allowing to set a percentage random variability for thermal appliances' power ($\delta_{P,thermal}$), which RAMP exploits to uniquely characterise each switch-on event, as shown in the example in Figure 5. The possibility to randomly variate *Appliances*' power is nonetheless useful for modelling any other kind of appliance that allows for power regulation (e.g. electric heating stoves, ovens, etc.) as already shown in Figure 3 for the cooking cycle example.

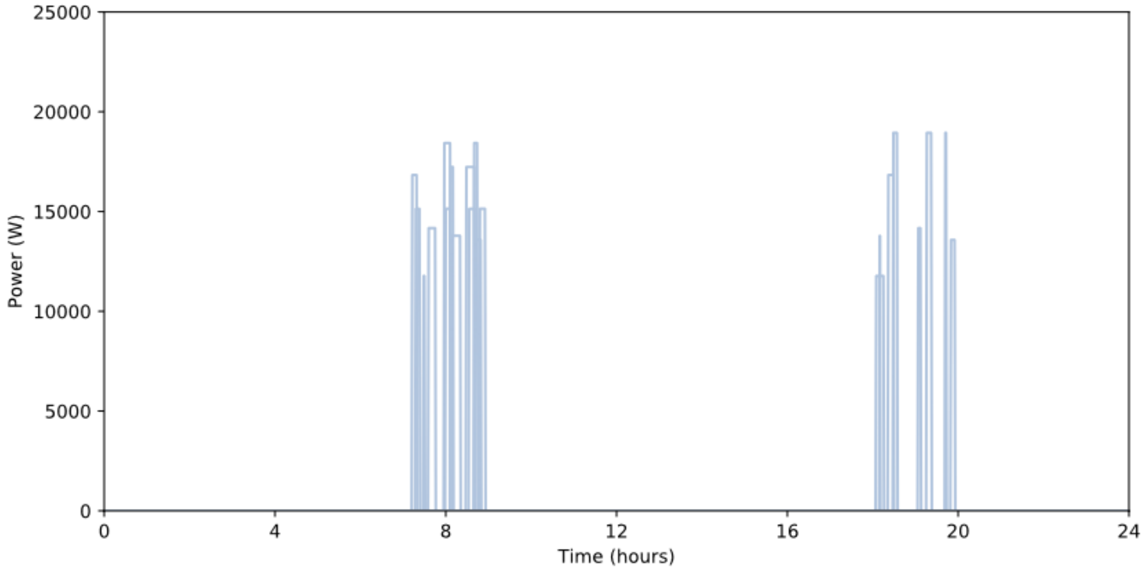

Figure 5 – Example of multiple stochastic runs (10) for a thermal appliance, in this case reproducing a "shower" task: the model variates not only switch-on times and shower duration, but also the absorbed power (i.e. hot water temperature).

3. Empirical data and validation criteria

To validate the model, we test it against empirical data measured by the isolated hybrid micro-grid system of the village "El Espino" (-19.188, -63.560), in Bolivia, installed in September 2015 and composed of 60 kW of PV panels, 464 kWh of battery storage and a 58 kW Gen-Set. The system serves a community of 128 households, a hospital and a school, as well as the public lighting service. A comprehensive description of the system and of the data is available in Balderrama et al. [25]. Aggregate electric load data are available as an indirect measure, i.e. as the sum of direct measurements retrieved from PV arrays, Gen-Set and batteries by means of smart meters. No measurement systems is instead available in the area for the quantification of non-electrical energy usage by individual users (e.g. traditional biomass or LPG for cooking). Accordingly, the latter is not considered for the present analysis.

3.1. Validation dataset

A survey campaign was conducted in El Espino in November 2016 to collect the interview-based information needed as an input for the bottom-up RAMP model. As a result of the interviews, it is possible to identify 8 *User types*, summarised in Table 2, whilst Table 3 shows the main appliances owned, on average, by *Users* comprised in the identified *types*.

User type **name**	**Acronym**	***Users*** **number**
High-Income households	HI	11
Higher Middle-Income households	HMI	38
Lower Middle-Income households	LMI	34
Low Income households	LI	45
Hospital	HO	1
School	SC	1
Public lighting service	PL	1
Church	CH	3

Table 2 – Summary of the identified *User types*, with their respective number of *Users*.

	HI	**HMI**	**LMI**	**LI**	**HO**	**SC**	**PL**	**CH**
Indoor bulbs	6	5	3	2	12	8	0	10
Outdoor bulbs	2	2	2	1	1	6	0	7
TV	2	1	1	1	0	1	0	0
DVD	1	1	1	1	0	1	0	0
Antenna	1	1	1	1	0	0	0	0
Phone charger	5	4	4	2	8	5	0	0
Freezer	2	1	0	0	0	1	0	0
Fridge	0	0	0	0	3	0	0	0
Stereo system	0	0	0	0	0	1	0	0
Mixer	1	1	1	0	1	0	0	0
PC	0	0	0	0	1	18	0	0
Printer	0	0	0	0	0	1	0	0
Radio	0	1	0	0	0	0	0	0
Large public light	0	0	0	0	0	0	25	0
Small public light	0	0	0	0	0	0	12	0
Speaker	0	0	0	0	0	0	0	1

Table 3 – Summary of the appliances owned, on average, by each *User* belonging to the different *User types*.

Appliances' power, time frames of use and average total daily time of use, are also gathered from the interviews. For the simulation, months from May 2016 to April 2017 – i.e. the six months before and after the interview period – are considered. As summarised in Table 4, the input files for each month take into account: (i) monthly-averaged variation of dawn and dusk timings with respect to the interview period; (ii) holidays, weekends and periods of seasonal work outside the village; and (iii) monthly-averaged daily temperature trends affecting the behaviour of fridges and freezers. In fact, dawn timing are used to variate accordingly the switch-off timing of public lighting, while dusk timing correlates with both public lighting switch-on and with the start of the time frame related with households evening activities. Modular duty cycles are employed to better reproduce the actual behaviour of fridges and freezer with respect to the identified seasonal temperature trends (further details are given in Appendix A). School vacations are identified to exclude the school load in the corresponding periods, whilst other holidays and Christmas vacations are considered as weekends (see sub-section 3.2). In January and February, a few households (precisely 6 LMI and 9 LI) move outside the village to work as farmers, determining a corresponding reduction in the number of modelled users. It is also worth specifying that mixers are flagged as "occasionally-used" and public lights are of course set to always switch-on altogether ($fixed\_num_{ijk}$ = True), following the specified dawn and dusk timings. All the detailed input files used to model each month in RAMP are reported in the Supplementary material.

Month	**Dawn**	**Dusk**	**Season type**	**Special conditions**
May	06:01:00	17:49:00	Cold	
June	06:20:00	17:46:00	Cold	
July	06:22:00	17:54:00	Cold	School vacations (1^{st} – 17^{th})
August	06:10:00	18:03:00	Warm	
September	05:52:00	18:07:00	Warm	
October	05:22:00	18:14:00	Hot	
November	05:10:00	18:27:00	Hot	
December	05:14:00	18:43:00	Hot	Start school vacations: 8^{th} Vacation for all: 8^{th} – 31^{st}
January	05:31:00	18:52:00	Hot	School vacations Seasonal work
February	05:45:00	18:45:00	Hot	End school vacations: 8^{th} Seasonal work
March	05:54:00	18:25:00	Warm	
April	06:02:00	18:02:00	Cold	

Table 4 – Summary of changes to the input files in each month in relationship with seasonal factors.

Given the high degree of uncertainty inherent to synthetic loads generated with a bottom-up approach based on data collected via interviews, the validation of the model against experimental data is also complemented – only for the representative month of November 2016 – with a further comparison against the only previously-published model conceived for this kind of application, i.e. LoadProGen. Given the formal analogy in the definition of the inputs for the two models, these are kept identical when possible, with the exception of the additional *Appliances' optional attributes*, introduced in the present study, that are not applicable to LoadProGen. The inputs for the latter model are also reported in the Supplementary material.

3.2. Validation parameters

For each month, a number of stochastic profiles equivalent to the number of days in the month is generated, discriminating between weekdays and weekends or holidays. These are differentiated by acting on the parameter $\mu_{\%}$, which is incremented by 50% for weekends and holidays in order to skew the on-peak distribution of f_{coinc_num} towards higher values. Detailed information for each month about the number of weekdays, weekends and holidays are reported in the Supplementary material.

A set of indicators are defined to evaluate how much the generated profiles adhere to the experimental ones. The accuracy of the shape of the average daily aggregate load profile is evaluated by means of the *Normalised Root-Mean-Squared Error* (NRMSE) (Equation 6).

$$NRMSE = \frac{\sqrt{\frac{\sum_{x}^{N_t}\left(P_{model}(x) - P_{measured}(x)\right)^2}{N_t}}}{P_{measured,average}} \ [\%] \tag{6}$$

Where $P_{model}(x)$ $[W]$ is the value of the modelled load (via either RAMP or LoadProGen) at each time-step x, $P_{measured}(x)$ $[W]$ is the one related to the measured load and N_t is the total number of time-step (e.g. 1440 for a 1-minute resolution). $P_{measured,average}$ $[W]$ is the average value of the measured daily average load profile.

Other parameters that are widely used to validate load profile models against measured data and that are also critical with regards to the sizing of the associated energy systems are the *Load Factor*, the *Coincidence Factor*, the value of the *Peak Load* and the value of the *Aggregate demand* [9,18,26]. The mathematical definitions of *Load Factor* and *Coincidence Factor* are reported in Equations 7 and 8, respectively.

$$LF = \frac{P_{L,avg}}{P_{L,peak}} \tag{7}$$

$$CF = \frac{P_{L,peak}}{P_{L,max}} \tag{8}$$

Where $P_{L,avg}$ $[W]$ and $P_{L,peak}$ $[W]$ are, respectively, the daily average and the maximum (or peak) values of the average profile resulting from those measured or modelled; and $P_{L,max}$ $[W]$ is the theoretical maximum load, i.e. the total installed load.

4. Results and discussion

The comparison between measured data and synthetic profiles generated based on the existing approach (LoadProGen) and on RAMP are represented in Figure 6 and Figure 7, respectively. Figure 6 shows how the existing approach, though well reproducing the measured degree of day-to-day variability, leads to a significant overestimation of the *Peak Load* and of the load related to the first hours of the day, which is in turns compensated by an underestimation of the load during mid-day hours. This behaviour can be partly attributed to the reliance on an empirical correlation for the computation of coincidence and load factors, and partly to the impossibility to modulate the load of fridges and freezers throughout the day; in fact, this last effect plays a major role since cooling appliances are among the few high-power appliances of the village. Conversely, Figure 7 shows that RAMP is capable of reproducing both the average daily profile and the day-to-day fluctuations with good approximation, as a consequence of its higher degree of stochasticity and of its additional features – such as duty cycles modulation.

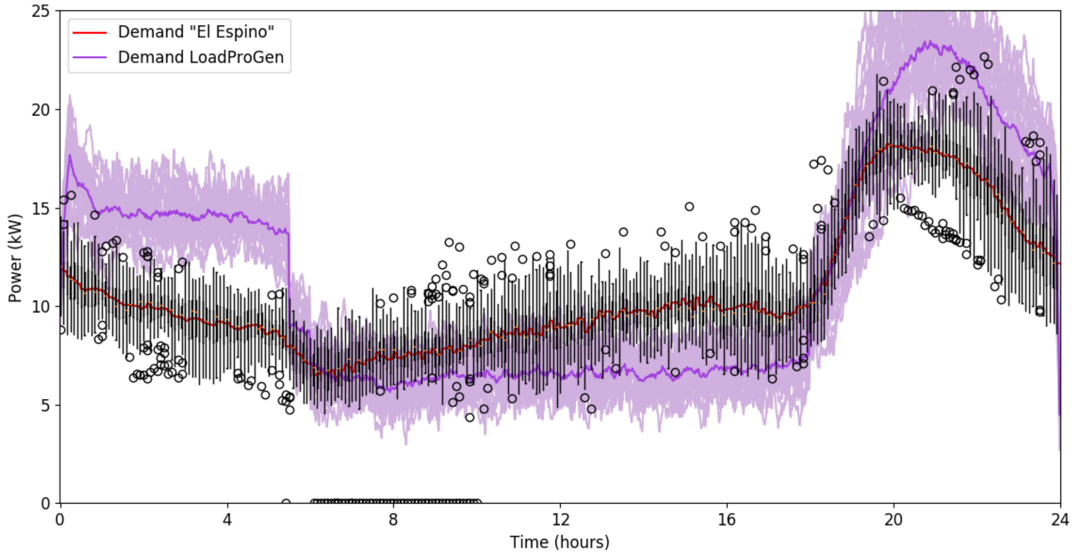

Figure 6 – Comparison between measured data (boxplots) for the weekdays of November 2016 and stochastic profiles generated by LoadProGen.

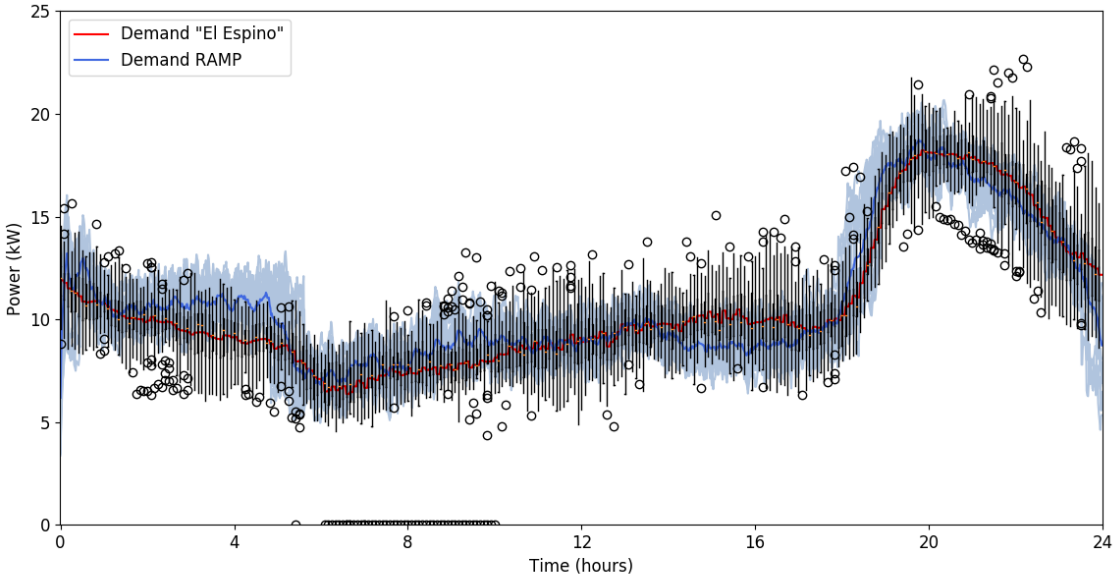

Figure 7 – Comparison between measured data (boxplots) for the weekdays of November 2016 and stochastic profiles generated by RAMP.

Table 5 provides a quantitative overview of the comparison between measured data and the two models, based on the previously defined indicators. The LoadProGen approach confirms a marked discrepancy with measured data in terms of *Peak Load, Load Factor* and *Coincidence Factor* as well as a higher NRMSE as compared to RAMP. The latter, instead, provides a good approximation for all the indicators, with percentage errors compared to measured data consistently small – the highest being 2.9% for the *Peak Load* – and with a NRMSE value below 10%.

	Measured data	**LoadProGen**	**RAMP**	**Err% LoadProGen**	**Err% RAMP**
Peak Load [W]	18222	23466	18751	28.8%	2.9%
LF [-]	0.585	0.491	0.578	16.1%	1.2%
CF [-]	0.114	0.147	0.117	28.9%	2.6%
Aggregate Demand [kWh]	256	269	260	5.1%	1.6%
NRMSE [%]	-	30.6%	9.6%	-	-

Table 5 – Quantitative comparison between measured data, LoadProGen and RAMP, for the reference month of November 2016.

	Err% *LF*	**Err% *CF***	**NRMSE [%]**
May-16	0,5%	8,3%	14,3%
Jun-16	0,4%	9,1%	13,5%
Jul-16	2,4%	1,9%	11,1%
Aug-16	9,0%	0,4%	12,9%
Sep-16	8,4%	2,2%	10,2%
Oct-16	0,7%	0,9%	9,7%
Nov-16	2,4%	2,7%	9,1%
Dec-16	4,8%	1,8%	9,8%
Jan-17	2,2%	2,7%	9,0%
Feb-17	1,5%	5,1%	8,1%
Mar-17	7,6%	4,3%	8,5%
Apr-17	2,1%	5,8%	11,8%
Average	**3,5%**	**3,8%**	**10,7%**

Table 6 – Quantitative summary for all months of the agreement between RAMP results and measured data.

The above values are computed for the reference month of November 2016. However, it is also relevant to check the effect of seasonal and/or weather effects on the model validation. Figure 8 shows the consistency of RAMP-generated profiles for each month of the year. The modelled average monthly profiles follow with a good agreement the changes in the real load associated with seasonal patterns. Changes in the peak time are well approximated by the shift of evening activities and public lighting in correlation with sunset timings, whilst the drop in the load during the coldest months (from April to July) is well captured by the lower intensity of fridges and freezers cycling behaviour. As further highlighted by Figure 9, which reports the corresponding Load Duration Curves, the modelled profiles match with a satisfying degree of accuracy the peak load in all months. Some discrepancies remain for medium loads but are deemed acceptable, especially considering the lack of precise information about the users' behaviour in all months. Quantitative indicators are reported in Table 6.

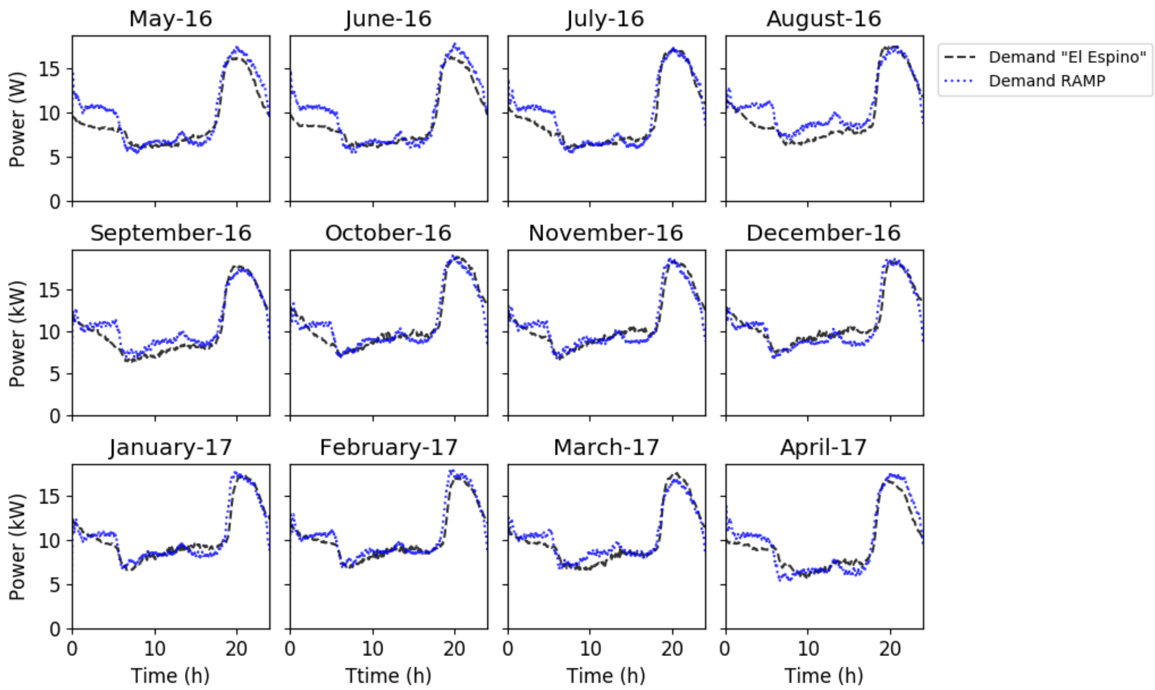

Figure 8 – Monthly-averaged daily load profiles throughout the year. Comparison between measured data and RAMP.

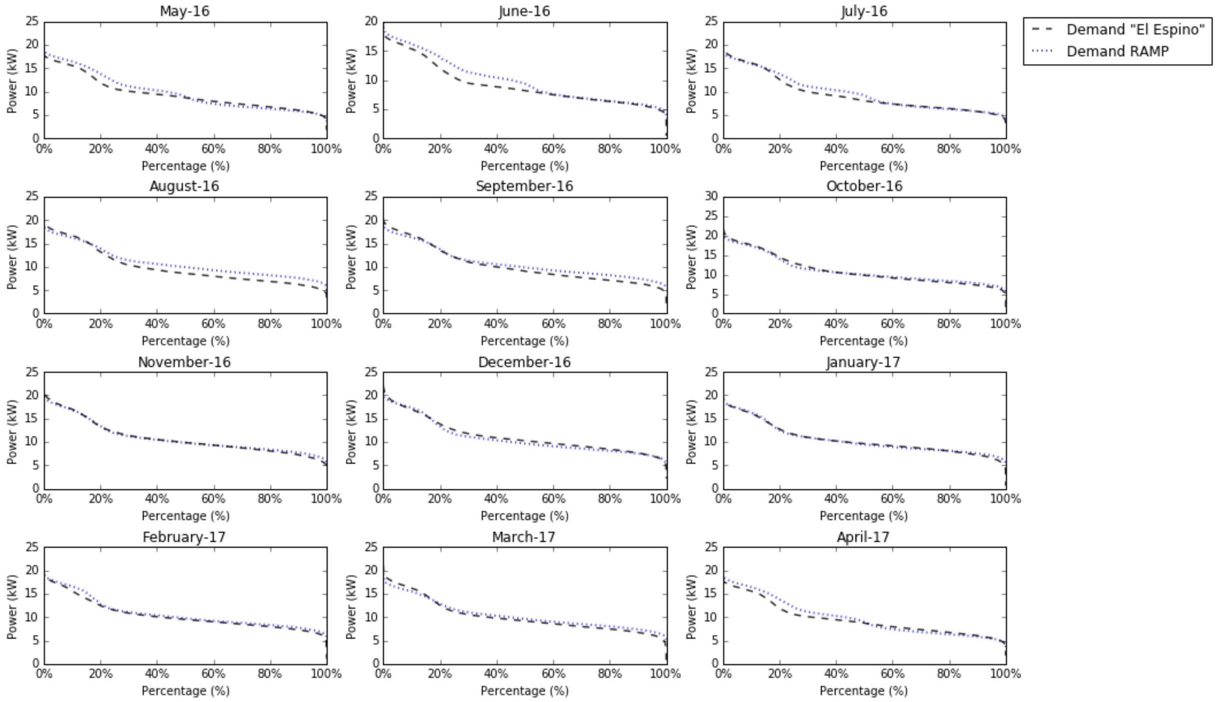

Figure 9 – Monthly-averaged daily Load Duration Curves throughout the year. Comparison between measured data and RAMP.

To evaluate the robustness of the model with respect to some of the calibration parameters assumed, such as δ_{peak} and $\mu_{\%}$, a sensitivity analysis on the latter is performed. For the reference month of November, a set of 15 synthetic profiles is generated and averaged for each calibration parameter value. As shown in Figure 10, $\mu_{\%}$ is varied between -25% and +50% of tis default value – i.e. the mean value of the on-peak distribution, as per Equation 4 – without having significant effects on the peak time frame behaviour. As expected, increasing $\mu_{\%}$ produces a slightly higher load during the peak time frame, as a result of the more likely coincident switch-on of multiple appliances of the same kind. The higher increase of 50%, which is assumed for week-ends and holidays as discussed in sub-section 3.2, produces a variation comparable to an increase of 25%, allowing to realistically represent a weekend-like behaviour without producing significant changes in the load shape. Similarly, the variation of δ_{peak}, aimed at shortening or enlarging the peak time frame duration in relationship with days in which the users may stay awake for shorter or longer periods (e.g. special events), is only marginally affecting the results, demonstrating a satisfying robustness of the model with respect to the assumed values.

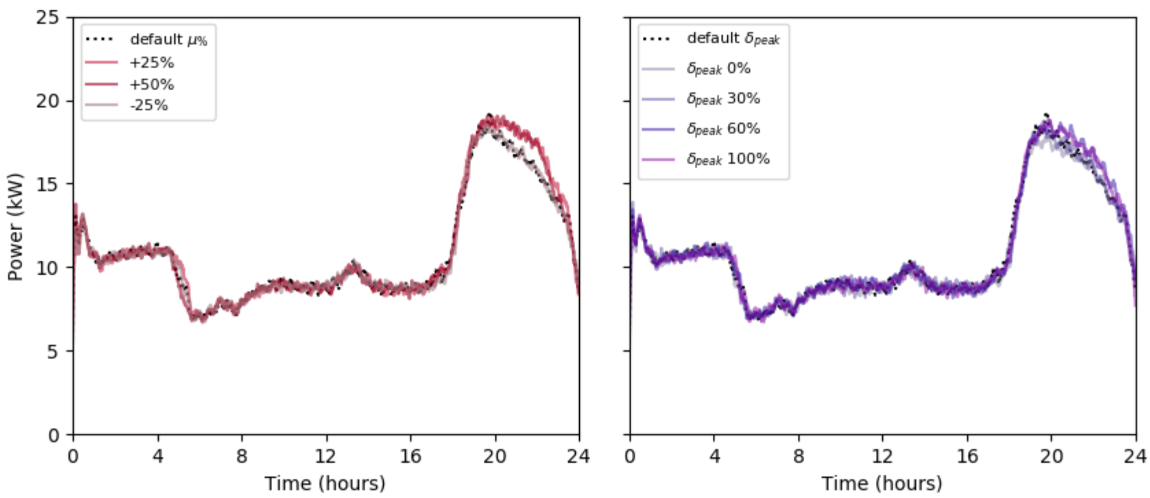

Figure 10 – Sensitivity analysis on the values assumed for the calibration parameters $\mu_{\%}$ and δ_{peak}.

Conclusions

The analysis demonstrates that the designed RAMP model provides a good agreement with measured data for all the selected indicators (with an average *NRMSE* around 10%). Moreover, it exhibits a higher performance than previous approaches, in terms of both profile shape and quantitative indicators. The increased adherence to reality of the newly designed RAMP is explained, among others, by the increased degree of stochasticity and the possibility to define modular duty-cycles for selected appliances. The latter functionality has revealed to be very useful to reproduce correctly the seasonal variations of the load, by adjusting the cycling depths of some appliances as a function of the ambient temperature. This is especially relevant if the load is dominated by cooling appliances. A good consistency has also been obtained for the parameters that can be manipulated to force the model towards reproducing social behaviours that are typical of weekends, holidays or special events. The sensitivity analysis has also shown that there are no excessive changes in the results depending on the modeller's assumptions. Other innovative functionalities introduced within RAMP – such as cooking cycles, thermal appliances power regulation, frequency of use, etc. – that are only partially relevant for the context considered for the validation, would likely mark an even larger differentiation from the existing approach in cases in which they are predominant.

The model validation focused only on energy uses satisfied by electrical appliances due to the complete lack of measurement systems for other types of energy vectors in the area. Nonetheless, a key innovation introduced by RAMP is that the model features are not conceived for the mere generation of electric load profiles (in fact, the model is independent from any empirical correlation) but rather for the generation of “multi-energy” loads that may help designing energy systems capable of satisfying all final energy uses. What is more, the model is tailored for energy planning purposes. As such, it is adaptable to the simulation of additional loads than those assessed via interviews. To this regard, it is also worth underlining how RAMP is released as open-source software, allowing for a better user- and context-adaptability as well as for the simulation of “extreme scenarios” (e.g. one with a high probability of coincident behaviour) that may be relevant within the framework of robust optimisation. The code openness also facilitates further developments, which may include standby power consumption and simplified and context-adaptable building models for space heating loads simulation.

Acknowledgements

The authors gratefully acknowledge Mario Carmelo Paz Duran, general manager of the “Cooperativa rural de electrificacion” that operates the El Espino micro-grid system, for sharing data about measured load profiles and outdoor temperature profiles. They also acknowledge Stefano Sabatini and Andrea Tarantino for performing the field campaign for the collection of survey-based information.

Appendix A – Fridges cycle modulation

As discussed in sub-section 3.1, fridges and freezers cycles are modulated within RAMP based on the main parameters influencing their behaviour, i.e. room temperature and users’ activity level (as a proxy for door openings) [23,24]. Given the lack of data about indoor temperatures for the simulated building types, and considering that those are not well insulated and lack any air conditioning or space heating system, outdoor temperature is considered as approximately equal to the indoor one. The assumption is also supported by the fact that indoor and outdoor temperatures correlate better at warmer outdoor temperatures, as those considered [27].

As shown in Figure 11, the outdoor temperature profile for the month of November is, on average, often above 25 °C, and approaches 30°C during mid-day hours. Accordingly, a “standard” cycle (compressor working for 1/3 of the time) is modelled only for those time frames in which the temperature is in its lowest range (24-25 °C) and users’ activity levels are low. An “intensive” cycle (compressor on for 2/3 of the time) is instead modelled for the time frame in which temperature is higher than 27°C and/or users’ activity is high, whilst an “intermediate” cycle (compressor on for 1/2 of the time) is modelled for the time frame in which temperature starts to increase but users’ activity level are medium-high to medium-low.

As regards LoadProGen, an “intermediate” behaviour has been set as the unique input to the model, given the impossibility to reproduce such cycle modulation.

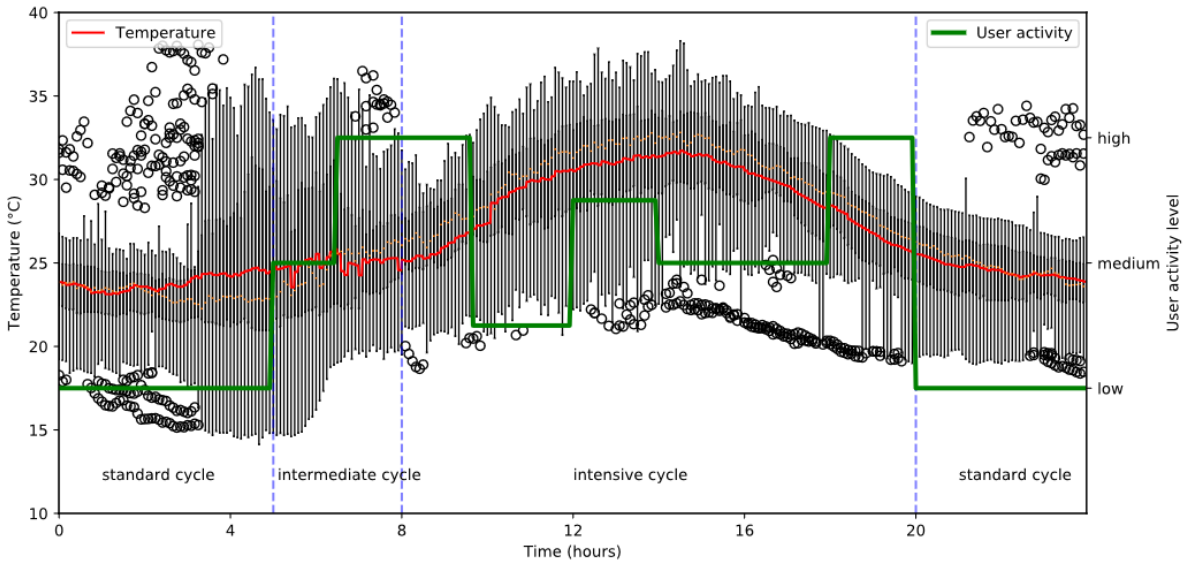

Figure 11 – Outdoor temperature average profile and boxplot for El Espino in November 2016, compared with the corresponding typical residential user activity level (i.e. qualitative level of interaction with fridges and freezers).

Such kind of considerations are repeated for the other months so as to capture potential variations of fridges cycling behaviour throughout the year. Figure 12 shows the temperature trends in El Espino in the period considered for the validation. Months from October to February all present similar temperature profiles, and are classified as “hot months”, for which the cycling behaviour is the same as November. The periods preceding and following such hot months (namely August, September and March) can be instead treated as “warm months”, for which the cycling behaviour experiences some minor variations.

Finally, months from April to July are “cold months” for which the cycling behaviour is significantly less intense, as reported in Table 7. The detailed cycling behaviour assumed for all months is reported in the Supplementary material.

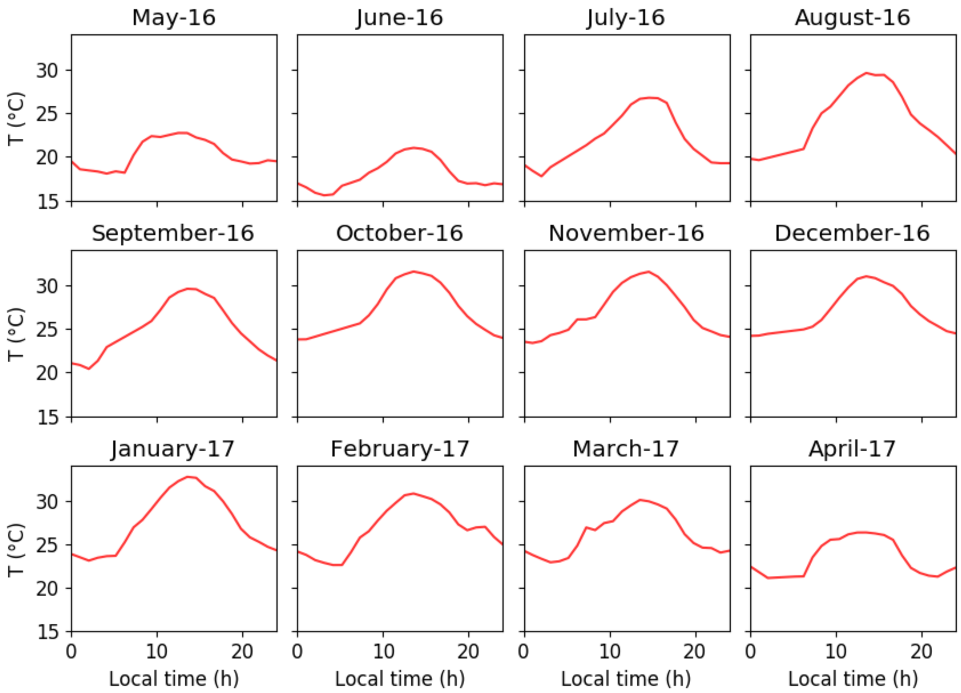

Figure 12 – Monthly-averaged daily temperature profiles in El Espino.

Season type	**Standard cycle**	**Intermediate cycle**	**Intensive cycle**
Hot	00:00:00 – 04:59:00 20:01:00 – 23:59:00	05:00:00 – 07:59:00	08:00:00 – 20:00:00
Warm	00:00:00 – 04:59:00 18:01:00 – 23:59:00	05:00:00 – 09:39:00	09:40:00 – 18:00:00
Cold	00:00:00 – 04:59:00 20:01:00 – 23:59:00	08:00:00 – 20:00:00	-

Table 7 – Summary of the fridges duty-cycle modulation estimated based on the seasonal temperature trends.

Nomenclature

δ_{peak}	Factor multiplying t_{peak} to get st.dev. of peak time frame random sampling [%]
$\mu_{\%}$	Factor multiplying m_{ijk}o get average of $m_{ON,ijk}$e random sampling [%]
CF	Coincidence Factor [-]
f_{coinc_num}	Coincident numerosity factor [-]
LF	Load Factor [-]
m_{ijk}	Numerosity of appliances of a certain type [-]
$m_{ON,ijk}$	Numerosity of appliances of a certain type switched-on in a switch-on event [-]
$NRMSE$	Normalised Root-Mean-Squared Error [%]
t_{peak}	Peak time [min]

References

[1] Mathiesen B V., Lund H, Connolly D, Wenzel H, Ostergaard PA, Möller B, et al. Smart Energy Systems for coherent 100% renewable energy and transport solutions. Appl Energy 2015;145:139–54. doi:10.1016/j.apenergy.2015.01.075.

[2] Mancarella P. MES (multi-energy systems): An overview of concepts and evaluation models. Energy 2014;65:1–17. doi:10.1016/j.energy.2013.10.041.

[3] Good N, Mancarella P. Flexibility in multi-energy communities with electrical and thermal storage: A stochastic, robust approach for multi-service demand response. IEEE Trans Smart Grid 2017. doi:10.1109/TSG.2017.2745559.

[4] Lund H, Alberg P, Connolly D, Vad B. Smart energy and smart energy systems. Energy 2017;137:556–65. doi:10.1016/j.energy.2017.05.123.

[5] International Energy Agency. Energy Access Outlook 2017 - From poverty to prosperity. 2017.

[6] Riva F, Ahlborg H, Hartvigsson E, Pachauri S, Colombo E. Electricity access and rural development: Review of complex socio-economic dynamics and causal diagrams for more appropriate energy modelling. Energy Sustain Dev 2018;43:203–23. doi:10.1016/j.esd.2018.02.003.

[7] Riva F, Gardumi F, Tognollo A, Colombo E. Soft-linking energy demand and optimisation models for local long-term electricity planning: An application to rural India. Energy 2019;166:32–46. doi:10.1016/j.energy.2018.10.067.

[8] Grandjean A, Adnot J, Binet G. A review and an analysis of the residential electric load curve models. Renew Sustain Energy Rev 2012;16:6539–65. doi:10.1016/J.RSER.2012.08.013.

[9] Widén J, Wäckelgård E. A high-resolution stochastic model of domestic activity patterns and electricity demand. Appl Energy 2010;87:1880–92. doi:10.1016/j.apenergy.2009.11.006.

[10] Widén J, Lundh M, Vassileva I, Dahlquist E, Ellegård K, Wäckelgård E. Constructing load profiles for household electricity and hot water from time-use data-Modelling approach and validation. Energy Build 2009;41:753–68. doi:10.1016/j.enbuild.2009.02.013.

[11] Tsagarakis G, Collin AJ, Kiprakis AE. Modelling the electrical loads of UK residential energy users. Proc Univ Power Eng Conf 2012:1–6. doi:10.1109/UPEC.2012.6398593.

[12] Collin AJ, Tsagarakis G, Kiprakis AE, McLaughlin S. Development of low-voltage load models for the residential load sector. IEEE Trans Power Syst 2014;29:2180–8. doi:10.1109/TPWRS.2014.2301949.

[13] Fischer D, Härtl A, Wille-haussmann B. Model for electric load profiles with high time resolution for German households. Energy Build 2015;92:170–9. doi:10.1016/j.enbuild.2015.01.058.

[14] McKenna E, Thomson M. High-resolution stochastic integrated thermal-electrical domestic demand model. Appl Energy 2016;165:445–61. doi:10.1016/j.apenergy.2015.12.089.

[15] Richardson I, Thomson M, Infield D, Clifford C. Domestic electricity use: A high-resolution energy demand model. Energy Build 2010;42:1878–87. doi:10.1016/j.enbuild.2010.05.023.

[16] Good N, Zhang L, Navarro-Espinosa A, Mancarella P. High resolution modelling of multi-energy domestic demand profiles. Appl Energy 2015;137:193–210. doi:10.1016/j.apenergy.2014.10.028.

[17] Marszal-Pomianowska A, Heiselberg P, Kalyanova Larsen O. Household electricity demand profiles - A high-resolution load model to facilitate modelling of energy flexible buildings. Energy 2016;103:487–501. doi:10.1016/j.energy.2016.02.159.

[18] Hartvigsson E, Ahlgren EO. Comparison of load profiles in a mini-grid: Assessment of performance metrics using measured and interview-based data. Energy Sustain Dev 2018;43:186–95. doi:10.1016/j.esd.2018.01.009.

[19] Mandelli S, Merlo M, Colombo E. Novel procedure to formulate load profiles for off-grid rural areas. Energy Sustain Dev 2016;31:130–42. doi:10.1016/j.esd.2016.01.005.

[20] Lombardi F, Riva F, Sacchi M, Colombo E. Enabling combined access to electricity and clean cooking with PV-microgrids: new evidences from a high-resolution model of cooking loads. Energy Sustain Dev 2019;49:78–88. doi:10.1016/j.esd.2019.01.005.

[21] Pfenninger S, Decarolis J, Hirth L, Quoilin S, Staffell I. The importance of open data and software : Is energy research lagging behind ? Energy Policy 2017;101:211–5.

doi:10.1016/j.enpol.2016.11.046.

[22] Balderrama S, Canedo W, Fernandez M, Lemort V, Quoilin S. Techno-economic optimization of isolate micro- grids including PV and Li-Ion Batteries in the Bolivian context. 29th Int. Conf. Effic. Cost, Optim. Environ. Impact Energy Syst., 2016, p. 1–12.

[23] Saidur R, Masjuki HH, Choudhury IA. Role of ambient temperature, door opening, thermostat setting position and their combined effect on refrigerator-freezer energy consumption. Energy Convers Manag 2002;43:845–54. doi:10.1016/S0196-8904(01)00069-3.

[24] Geppert J, Stamminger R. Analysis of effecting factors on domestic refrigerators' energy consumption in use. Energy Convers Manag 2013;76:794–800. doi:10.1016/j.enconman.2013.08.027.

[25] Balderrama S, Haderspock F, Canedo W, Orellana R, Quoilin S. Techno-economic evaluation of rural electrification in Bolivia: lessons learned from the "El Espino" micro-grid. ECOS 2018, 2018.

[26] Jambagi A, Kramer M, Cheng V. Residential electricity demand modelling: Activity based modelling for a model with high time and spatial resolution. 2015 3rd Int Renew Sustain Energy Conf 2015:1–6. doi:10.1109/IRSEC.2015.7455047.

[27] Nguyen JL, Schwartz J, Dockery DW. The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. Indoor Air 2014. doi:10.1111/ina.12052.

Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model

Francesco Lombardi[a], Sergio Balderrama[b,c], Sylvain Quoilin[d,b], Emanuela Colombo[a]

[a] Politecnico di Milano, Department of Energy, Milan, Italy

[b] University of Liège, Department of Mechanical and Aerospace Engineering, Liège, Belgium

[c] Universidad Mayor de San Simon, Cochabamba, Bolivia

[d] KU Leuven, Department of Mechanical Engineering, Geel, Belgium

Corresponding author: F. Lombardi, e-mail address: francesco.lombardi@polimi.it

Supplementary material

Considering the large amount of information contained, all the input files used for modelling load profiles with RAMP and LoadProGen have been uploaded on the GitHub repository conceived for hosting also RAMP's source code. It can be accessed at the following link: https://github.com/SESAM-Polimi/RAMP.