Realistic Multi-Scale Modelling of Household Electricity Behaviours

(Article begins on next page)

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier 10.1109/ACCESS.2017.DOI

Realistic Multi-Scale Modelling of Household Electricity Behaviours

LORENZO BOTTACCIOLI[1,3] (Member, IEEE), SANTA DI CATALDO[1], ANDREA ACQUAVIVA[2,3] (Member, IEEE) and EDOARDO PATTI[1,3] (Member, IEEE).
[1]Dept. of Control and Computer Engineering, Politecnico di Torino, Italy; {lorenzo.bottaccioli, santa.dicataldo, edoardo.patti}@polito.it.
[2]Interuniversity Dept. of Regional and Urban Studies and Planning, Politecnico di Torino, Italy; andrea.acquaviva@polito.it.
[3]Energy Center Lab, Politecnico di Torino, Italy.
Corresponding author: Lorenzo Bottaccioli (e-mail: lorenzo.bottaccioli@polito.it).

ABSTRACT To improve the management and reliability of power distribution networks, there is a strong demand for models simulating energy loads in a realistic way. In this paper, we present a novel multi-scale model to generate realistic residential load profiles at different spatial-temporal resolutions. By taking advantage of information from Census and national surveys, we generate statistically consistent populations of heterogeneous families with their respective appliances. Exploiting a Bottom-up approach based on Monte Carlo Non Homogeneous Semi-Markov, we provide household end-user behaviours and realistic households load profiles on a daily as well as on a weekly basis, for either weekdays and weekends. The proposed approach overcomes limitations of state-of-art solutions that do not consider neither the time-dependency of the probability of performing specific activities in a house, nor their duration, or are limited in the type of probability distributions they can model. On top of that, it provides outcomes that are not limited on a per-day basis. The range of available space and time resolutions span from single household to district and from second to year, respectively, featuring multi-level aggregation of the simulation outcomes. To demonstrate the accuracy of our model, we present experimental results obtained simulating realistic populations in a period covering a whole calendar year and analyse our model's outcome at different scales. Then, we compare such results with three different data-sets that provide real load consumption at household, national and European levels, respectively.

INDEX TERMS Household Load Profile, Non Homogeneous Semi-Markov Model, Monte Carlo, Time Use survey, Use of Energy, Load Modelling, Behavioural Modelling.

I. INTRODUCTION

The latest international conference on climate changes (COP21) highlighted a real need to reduce greenhouse gas emission [1], that can be achieved i) by promoting installations of Renewable Energy Systems (RES) [2] and ii) by fostering a smart use of energy in cities [3]. For these reasons, we are moving forward the novel concept of smart grid, that will change current power distribution systems. Future smart grids will introduce advanced ICT solutions in the power systems, that will be combined together with RES and low carbon technologies.

The smart grid view promotes novel services and improved solutions for network management and system reliability. In this scenario, novel internet-connected smart meters [4] and Advanced Metering Infrastructures will play a crucial role, allowing a fine-grained monitoring of the entire distribution network and enabling services like: i) Non Intrusive Load Monitoring (NILM) [5], [6], ii) State Estimation (SE) [7] and iii) Demand Side Management (DSM) and Demand Response (DR) [8], [9].

NILM aims at desegregating load consumption of individual domestic appliances starting from a single point of measurement sampled by the smart meter, for example at 1 Hz frequency [5].

SE estimates the most probable state of a grid starting from redundant measurements sampled at different levels of the distribution network [10]. In scenarios where RES are widespread deployed across the grid, distributed SE algorithms are needed to provide Distribution System Operators with crucial information to evaluate operating conditions at any level of the grid (i.e. Building, Low Voltage (LV) and Medium Voltage (MV) levels) [7] .

DSM and DR allow to obtain a temporary virtual power plant by modifying the energy consumption patterns of end-users to address grid operation requirements [8] and to facilitate the integration of RES and energy storages [11]. Main objectives of both DSM and DR are peak shaving and self-consumption of RES [12]. Hence, they aim at changing the energy demand of end-users (either customers or prosumers). As any changes in power demand strictly depend on those appliances that are really shifted, DSM and DR policies need detailed information at appliance level to achieve their purpose [13].

To design and, especially, test the aforementioned services, we need massive information at different spatial levels (respectively appliance, household, building, district or city) and time resolutions (from few seconds up to one year). As the internet-connected smart meters and Advanced Metering Infrastructures are not widespread deployed in our cities yet, retrieving such data, or even realistic open data-sets, is extremely difficult. To overcome this issue, there is a strong demand for multi-scale models simulating energy loads, especially in the residential sector.

Power demand in households is strongly influenced by user behaviours [14]. Hence, we need an accurate knowledge of activities and behaviours of end-users (or inhabitants) to estimate the energy use in houses [15]. "Presence in the home is important for consumption" [16] and domestic load patterns are strongly affected by users occupancy [17]. On top of that, models of activities and behaviours of end-users need to take into account different family compositions and lifestyles of numerous households [18]. For example, two families of the same size and comparable daily power demand might still show different load trends.

Literature reports two different approaches to estimate residential loads over the time [19], [20]: *Top-down* and *Bottom-up*. The *Top-down* approach estimates the total load profiles of residential sector based on aggregations or statistical information (e.g. measurements at MV/LV substation level or national energy statistics, respectively). Then, energy consumption patterns are assigned to households according to their characteristics. The *Bottom-up* approach builds load profiles of statistically representative households exploiting information on i) activities and behaviours of end-users, ii) single appliance load consumption and iii) set of appliances for each house. The aggregation of individual household load profiles in a specific area, for example in a district, determines the energy consumption trends over the time in a specific portion of the power grid [21].

On these premises, Bottom-up models are the most suitable for serving the purpose of services like NILM, SE, DSM and DR. On the other hand, obtaining a performing Bottom-up model comes with two major challenges, that are i) modelling with a good approximation activities and behaviours of end-users, and ii) keeping the model structure simple while guaranteeing realistic consumption patterns [15], [19].

In order to solve these issues, in this paper we propose a novel Bottom-up *multi-scale model* to simulate energy consumption trends with different spatial-temporal resolutions. Our solution exploits results of national survey on Time Use and other statistical information on the population to provide i) a realistic model of the activities and behaviours of the end-users and ii) an accurate estimation of the distribution of heterogeneous families with appliances. To achieve these purposes, we exploit a simple Monte Carlo Non Homogeneous Semi-Markov model, that takes into account both the probability of performing an action at a certain time of the day and the duration of the action itself. The final outcomes are realistic residential load profiles, for either weekdays and weekends, with multi-level aggregation.

By doing so, our model fulfils five major requirements pointed out in a recent survey by Grandjean et al. [19]:

- *parametrization*, in order to simulate different scenarios (e.g. definition of households composition with different users and appliances);
- *presence of appliances information* impacting load trend;
- *adaptiveness*, allowing the addition of new simulated appliances;
- *generation of multi-level aggregate results* (e.g. household, district and city);
- *inclusion of domestic end-uses* affecting load trends (more specifically, domestic hot water and specific electricity appliances).

The rest of this paper is organised as follows. Section II reviews relevant literature solutions. Section III introduces the needed data-sources that are input of our solution. Section IV presents our model to simulate load consumption in households based on a stochastic approach. Section V presents the experimental results. Finally, Section VI discusses our concluding remarks.

II. RELATED WORK AND CONTRIBUTIONS

The novel concepts of smart grid promote novel services for an intelligent management of distribution networks. This new paradigm aims at revolutionising both power grids and energy marketplaces. However, to develop and test such new services, we need massive and pervasive information about the status of the grid even at household and appliance level. To overcome the lack of real information, we need realistic models to simulate the residential energy consumption patters. As highlighted in [14], [16], consumption is strongly affected by behaviours of end-users. Thus, we need first to study and model the activities and behaviours of the end-users at home.

The available solutions for Bottom-up approaches can be broadly categorised into two main groups: i) *non-Markov* and ii) *Markov* models.

i) non-Markov models. To the best of our knowledge, ARGOS, proposed by Capasso et al. [22], is the first model that reconstructs the residential power demand. ARGOS models the energy behaviours in dwellings exploiting a statistical approach that takes into account information on demography, life-style and socio-economical status. The final results are load trends with a 15 $minutes$ time resolution. Carpaneto and Chicco [23] proposed a probabilistic characterisation

of aggregated load consumption of houses served by same feeders or substations. This characterisation starts from a statistical study performed in a single house to obtain the probability distribution of aggregated load trends. They also analysed how the number of customers impacts on load patterns. In [24], Gruber at al. presented a consumer demand model that determines different households configurations by using a methodology based on normal distributions. This model computes households load profiles considering each single appliance at home discerning between workdays and holidays. Kong et al. [18] proposed a rule based model to simulate domestic load profiles according to the given: i) family composition, ii) daily schedule and iii) appliance preferences. This model is flexible in adding new appliance load profiles to each virtual house. Lan et al. [25] developed a model that consider only some household appliances: i) heating and cooling systems, ii) electric water heaters, iii) dryers and iv) lighting systems. Load profiles of heating and cooling systems are computed by TRNSYS software tool [26]. Load consumption of electric water heaters and dryers is given by probability distribution. Lighting systems profiles is computed by using the fundamental load profile of an house. In their model, Lan et al. combined this three approaches to compute the final household energy profile. However, this model neglects several appliances that have a strong impact on household consumption patterns (e.g. oven, electric cooker and fridge). Hoogsteen et al. [12] developed ALPG, an artificial load profile generator. This model exploits information on household occupancy profiles (one per house) and availability times for flexible devices to generate, as final result, household load profiles with 1 $minute$ time resolution. Gonzalez et al. [27] presented a user-friendly MATLAB-Simulink toolbox that simulates optimal on/off strategies of residential appliances to study residential energy profiles on a 24h horizon. However, it neglects occupancy profiles as well as end-user actions. In [28], authors proposed a model to simulate residential end-user electricity consumption by considering different parameters: i) types and number of electrical appliances in a house, ii) usage pattern of each appliance, iii) number end-users in a building and iv) activities performed by each end-user.
In our view, the main limitation of all this family of solutions is that, while trying to provide a probabilistic estimation of user behaviours, they completely neglect relevant statistical information that can be derived, for example, from national surveys on *Time Use* (TUS).

ii) Markov models. Other Bottom-up solutions implement a Markov chain to model the domestic activities of end-users. In [20], [29], the authors proposed two models to estimate energy consumption starting from American TUS. Thus, both models simulate the user behaviours patterns that are translated into load consumption over the day. Widen et al. [30] developed a stochastic framework to generate high-resolution load profiles. They implemented non-homogeneous Markov chains to simulate activities of end-users that are tuned to TUS. The final output is a power demand for individual or aggregated households with 1 $minute$ time resolution that embeds in its core two models to simulate domestic lighting systems [30] and electric water heaters [31]. In [32], Richardson et al. proposed an high-resolution energy demand framework that exploits two models to estimate domestic building occupancy [33] and simulate domestic lighting systems [34]. Even in this case, the framework implements a Markov chain to generate realistic behaviours of end-users at home. CHAP [35] is a stochastic multi-energy model that extends the models proposed by Richardson et al. [32], [33] and customises them for residential buildings in the UK. In [13], [14], authors presented two Bottom-up models that employ Monte Carlo Markov chain to develop residential demand profiles combined with electrical characteristics of appliances. Both models provide information on short- and long-term variations load patterns. Finally, Sancho et al. [36] presented a model based on both discrete-time Markov processes and survival analysis. Discrete-time Markov processes model the transitions' probability occurring between energy states. Whilst, survival analysis model both the switching-on/off of household appliances together with their sojourn time in different energy states.
The main limitation of classic Markov models is that, while modelling user behaviour, they do not consider neither the relation between the probability of taking actions and the specific time of the day (e.g. the probability of cooking at 12 a.m. is very different from the probability of performing the same activity at 12 p.m.), nor the duration of the activity. To address this issue, we propose a solution exploiting a Monte Carlo Non Homogeneous Semi-Markov Model, that is inherently capable of embedding such information.

With respect to presented literature solutions, we propose a *multi-scale model* addressing all the specifications identified by Grandjean et al. [19], and generating realistic household load profiles with different spatial-temporal resolutions. Simulation-space spans from a single household up to districts and cities. Simulation-time ranges from 10 minutes to years, with a granularity spanning from 1 second to 1 hour. The inputs of the model are i) *Time Use* surveys [37] that include information of twelve different classes of users (e.g. Part-time Working Male, Full Working Female and Kid); ii) surveys on *Use of Energy* [38] that provide distribution of appliances according to family-size and weakly statistics on usage of household appliances in families; iii) *Census data* [39] on families and population; iv) *Load profiles of real appliances* sampled at 1 Hz [40]; v) Real *weather data* [41]. Based on this information, our solution computes households activities and behaviours of each single end-user (according to its belonging class) exploiting a Monte Carlo Non Homogeneous Semi-Markov Model [42], [43]. Then, it builds heterogeneous families to create a virtual and realistic population and it assigns to each family a realistic set of appliances and their respective weekly usage. This is, to the best of our knowledge, a complete novelty, in that most of the available solutions provide only daily and not weekly usage profiles. For each end-user's activity, our model associates an appliance load consumption to create the final residential load trends. Finally, these load trends are aggregated by households, districts or cities according to the needed spatial-resolution.

Summarising, the specific contributions of our *multi-scale* model are:

- accurate modelling of activities and behaviours of end-users;
- realistic estimation of distribution of families and their respective appliances;
- realistic weekly residential load profiles for either week-days and weekends;
- possibility of simulation with very large spatial-temporal resolution range and granularity;
- multi-level aggregation of the simulation outcomes.

III. DATA SOURCES

In this section, we present the data-sources that are needed as input to our *multi-scale model* (presented in next Section IV) to create a virtual and realistic population of heterogeneous families. More specifically, we exploit the following sources: i) National survey on *Time Use*, ii) National survey on *Use of Energy*, iii) *Census data*, iv) *Load profiles of real appliances* and v) *Real weather data*. The rest of this section describes each input data-source in more details.

A. NATIONAL SURVEY ON TIME USE

A survey on *Time Use* (TUS) consists of a time diary covering 24h, where household members (or end-users) report all their daily activity sequences (e.g. sleeping, cooking and working) at 10 minutes intervals. TUS is a very useful instrument because it provides statistical information on activities and behaviours of both adults and children. These activities can be even grouped by type of the day (i.e. weekdays and week-ends) [19], [44]. Thus, such data are extremely important to analyse crucial aspects of daily life of people, such as: i) activities and needs, ii) ratio between work-time and free-time, iii) use of communication medias and iv) use of spaces and services. Our *multi-scale model* uses TUS data to build a user-activity model that simulates activities and behaviours of individual household members and, consequently, their respective electricity consumption at home (see Section IV).

The Italian National Institute of Statistics (ISTAT) performs every 5 years a national survey on TUS that is available for free download. In this work, we used the results of the most recent Italian TUS, that was performed in 2013-2014 [37]. These results consist in diaries filled in by around $60'000$ persons, grouped into about $27'000$ families, reporting: i) daily diary with performed actions and places, and ii) anagraphical information such as age, gender and profession.

Before giving this information as input to our *multi-scale model*, we grouped end-users into following twelve categories according to their age, gender and profession: i) *Full Working Female*, ii) *Full Working Male*, iii) *Part-time Working Female*, iv) *Part-time Working Male*, v) *No Working Female*, vi) *No Working Male*, vii) *Retired Female*, viii) *Retired Male*, ix) *Housewife*, x) *Student*, xi) *Kid* and xii) *Infant*. Then for each category, we grouped their related activities according to weekdays and weekends. To reduce the complexity of the model, we marked as *"Not Present"* all those activities that do not require an appliance or are not performed at home. Finally, we associated the remaining activities with one or more appliances (see Table 1) that have been modelled following a stochastic methodology (see Section IV). We also modelled washing machine, dishwasher and dryers following a probabilistic approach.

TABLE 1: Mapping between user activities and appliances

Activity	Appliance
Cooking, Eating, Ironing, Being at Bathroom, Showering, Reading, Studying, Watching Television	Lights
Studying, Using Personal Computer	Personal Computer
Cooking	Oven, Electric Cookers
Watching Television	Television
Ironing	Iron
Cleaning	Vacuum Cleaner
Listening to Radio	Radio
Dishwashing, Showering	Electric Water Heater

Regarding the usage of washing machine and dishwasher, we analysed in-dept the results of TUS and we computed their usage probability distributions for both weekdays and weekends. These information is used by our *multi-scale model* to simulate their energy consumption over the whole week. Analysing these data, we found a strong correlation between gender and usage of such appliances. As shown in Figure 1, washing machine and dishwasher are mainly used by Women. Deepening this analysis, Figure 2 highlights some remarkable differences in the usage over the week of washing machine and dishwasher between *Female workers* and *Housewives* and between *Male workers* and *Not Working Males*, respectively.

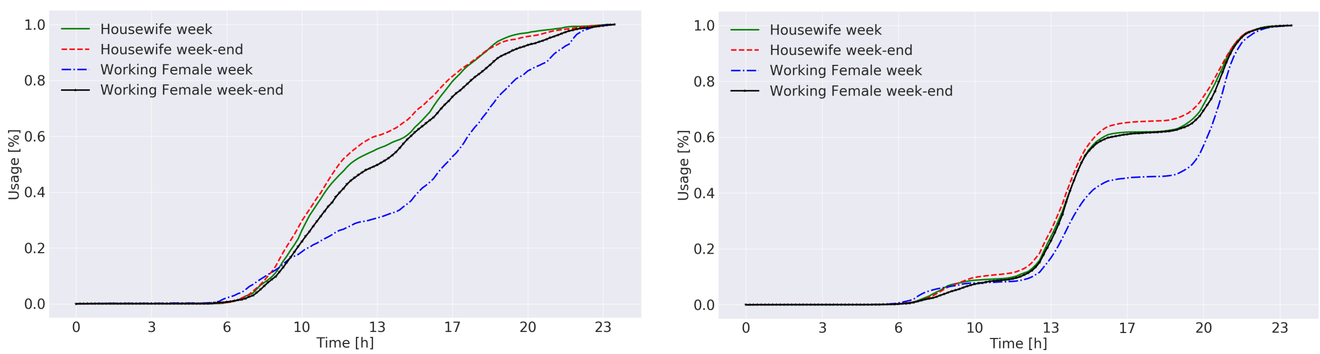
B. NATIONAL SURVEY ON USE OF ENERGY

In 2013, ISTAT performed a survey on more than $20'000$ Italian families to have an overview of their energy consumption [38]. Results of this survey provide also a statistical distribution of appliances with respect to family size (see Figure 3). Washing machine and oven are present in almost all Italian families regardless of their size (see Figure 3a and Figure 3b, respectively). As shown in Figure 3c, the probability that a family have a dishwasher rises from almost 20% (families with one member) to 60% (families with six members). Presence of an electric water heater slightly decreases from 20% to almost 10% (see Figure 3d). Dryer and electric cooker are not widespread in Italian families. Indeed, a dryer is present only in families with more than two members, with a probability not exceeding 10% (see Figure 3e). Electric cooker is present in about 5% of families, regardless of their size (see Figure 3f). Finally, Figure 3g and Figure 3h show the distribution of televisions (TVs) and personal computers (PCs), both increasing with family size.

(a) Usage profile of washing machine

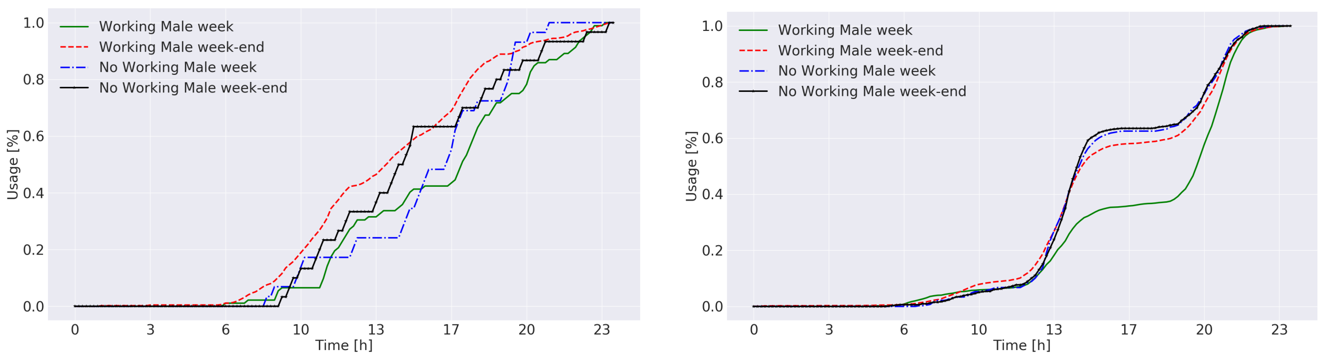
(b) Usage profile of dishwasher

FIGURE 1: Usage profiles of washing machine and dishwasher



(a) Probability of using the washing machine by *Female worker* and *Housewife*

(b) Probability of using the dishwasher by *Female worker* and *Housewife*



(c) Probability of using the washing machine by *Male worker* and *Not Working Male*

(d) Probability of using the dishwasher by *Male worker* and *Not Working Male*

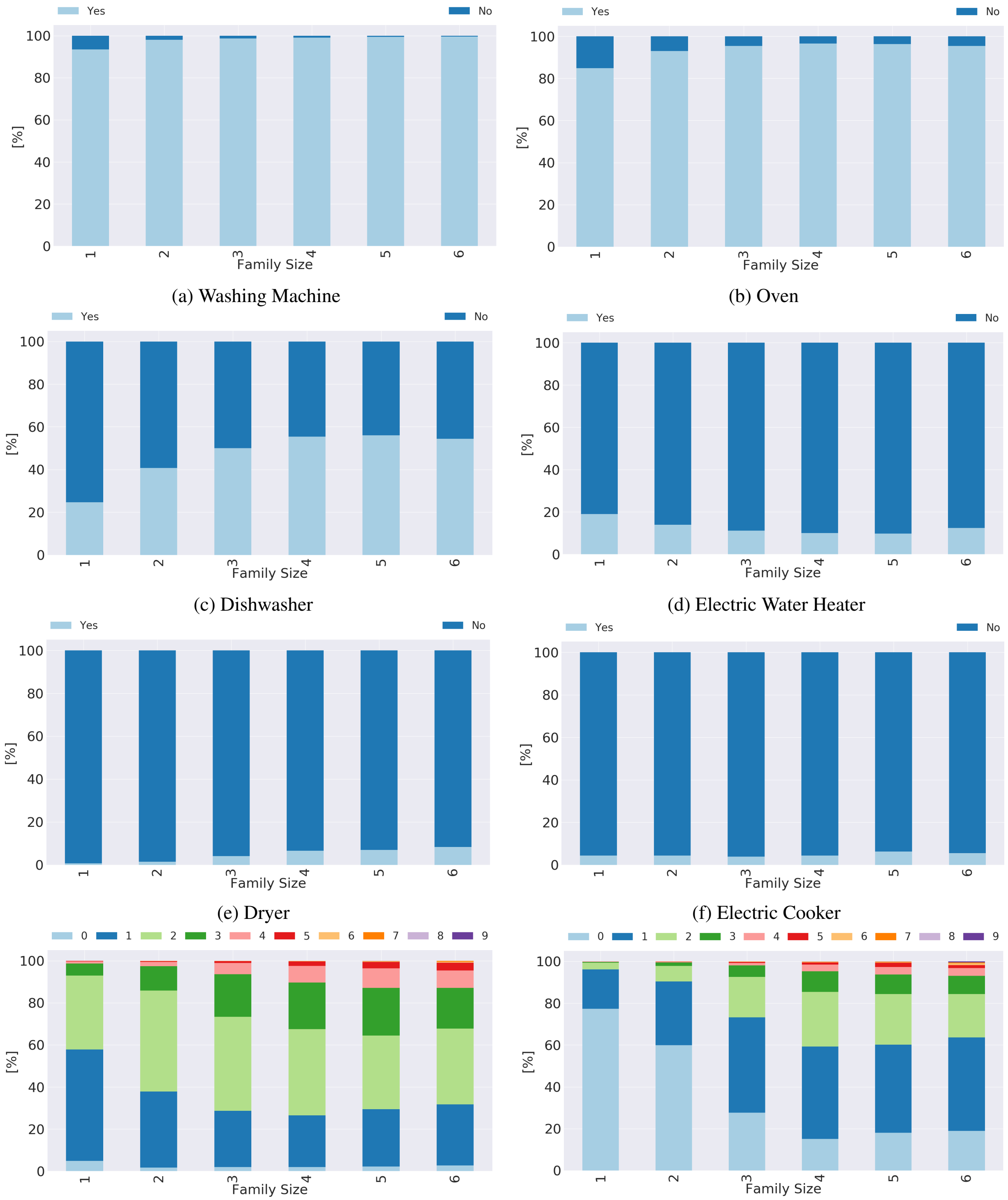
FIGURE 2: Probability of using washing machine and dishwasher

Furthermore, results of this survey provide the percentage of use of washing machine, dishwasher and oven with respect to family size, grouped by weekdays and weekends. As shown in Figure 4, the number of weekly activations rises with family size. This survey reports also additional information on some characteristics (e.g. load-size and production year) of washing machines, fridges and electric water heaters.

Our *multi-scale model* exploits both statistics on distribution and usage of appliances to build a virtual and realistic population to simulate. It associates a consistent set of appliances (among those mentioned above), together with their respective percentage of use, to each virtual family.

C. LOAD PROFILES OF REAL APPLIANCES

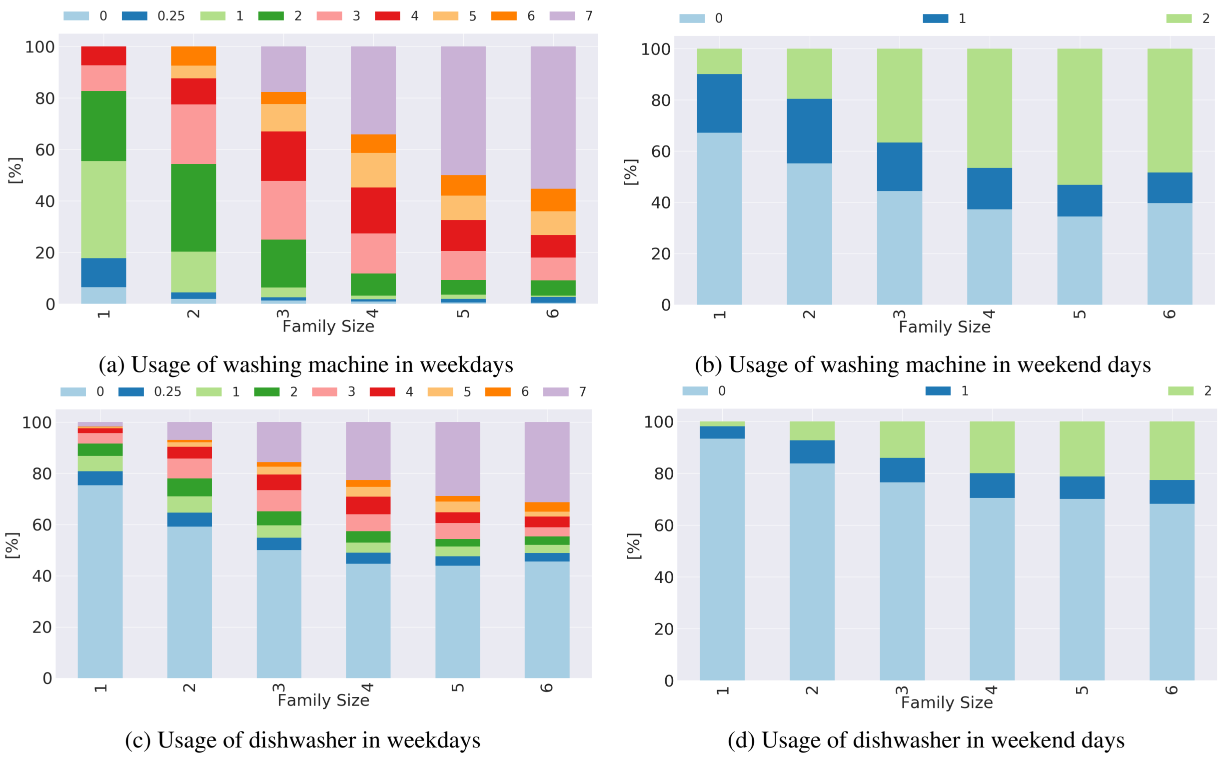
Our *multi-scale model* exploits *load profiles of real appliances*. We collected these load profiles by sampling different appliances with $1\ Hz$ resolution. We extended this data-set integrating also the *Tracebase* data-set [40] (available at http://www.tracebase.org under the Open Database License). The combination of both data-sets covers all the appliances



(a) Washing Machine
(b) Oven
(c) Dishwasher
(d) Electric Water Heater
(e) Dryer
(f) Electric Cooker
(g) Television
(h) Personal Computer

FIGURE 3: Distribution of appliances with respect to family size

reported in the national survey on *Use of Energy* (i.e. washing machine, dishwasher, dryer, electric water heater, TV, vacuum cleaner, PC, radio, oven, iron, electric cooker and fridge). The use of sampled load trends makes the whole *multi-scale model* flexible in including easily further appliances with different characteristics (e.g. load-size, model, brand and production year). For example, two similar virtual families can have a similar set of appliances with different characteristics, and hence different load profiles. In this scenario, the aggregated household load consumption of both



(a) Usage of washing machine in weekdays

(b) Usage of washing machine in weekend days

(c) Usage of dishwasher in weekdays

(d) Usage of dishwasher in weekend days

FIGURE 4: Percentage of use of appliances with respect to family size

families are different.

We chose a resolution of $1\ Hz$ to allow further tests of services that need either fine-grained information or aggregated data, such as NILM [5], [6] or Demand Response [8], [9], respectively. However, this input data-set can be easily replaced with other load profiles of appliances, even with different sampling period (either over-sampled or down-sampled).

D. CENSUS DATA

Census data provides statistics on families and population. In this work, we used desegregated information of the 15^{th} census in Italy [39], which is publicly available in free download. This data-set describes $248'500$ families in Italy, giving information on i) their composition and ii) age, gender and profession of each household member. During our simulations, this information has been used to generate a virtual and realistic population consisting of heterogeneous families statistically consistent.

E. REAL WEATHER DATA

Real *weather data* are needed to compute energy consumption of domestic lighting systems according to natural light (i.e. solar radiation). In this work, we retrieved such information from *Weather Underground* [41].

IV. METHODOLOGY

In this section, we present our proposed *multi-scale model* to simulate household load profiles with different spatial-temporal resolutions exploiting a Monte Carlo Non Homogeneous Semi-Markov chain [42], [43]. As shown in Figure 5, it consists of four computational modules grouped into following three macro-areas (or steps):

1) *Single Person Behaviour.* Simulation of activities and behaviours of individual end-users.
2) *Household Behaviour.* Composition of families and simulation of household activities.
3) *Aggregated and Single Load Profiles.* Multi-scale spatial-temporal simulation of aggregated and desegregated load profiles.

The rest of this section provides an in-depth description of each macro-area.

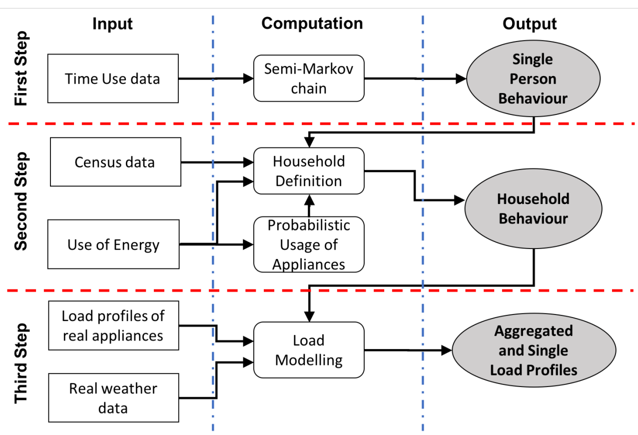


FIGURE 5: Schema of the proposed *multi-scale model*

A. SINGLE PERSON BEHAVIOUR

The first step in our *multi-scale model* consists of one main computation module called *Semi-Markov chain* in Figure 5. This module simulates, and gives as output, activities and behaviours of each end-user typology (see Section III-A) exploiting a Monte Carlo Non Homogeneous Semi-Markov chain and data obtained from the ISTAT Time Use [37] (see Section III-A).

Markov chains are a simple stochastic method to describe a sequence of possible events, where the probability of transitioning from state i to state j (i.e. from a certain activity to another, in our specific case) only depends on the current state and not on former occupied ones. Nonetheless, in a classic homogeneous Markov process the sojourn times of the states (i.e. the duration of the individual activities of the end users, in our case) should be exponentially distributed [45], which does not fit our specific case. Moreover, the time-dependency of the transition probabilities is not taken into account. To solve these issues, we decided to exploit a non-homogeneous semi-Markov process, with a two-fold advantage over previous literature methods based on classic Markov. i) Semi-Markov processes are distribution-free, and can adequately model processes with a non-exponential distribution of the sojourn times. ii) More specifically, *non-homogeneous* semi-Markov can adequately model time-dependent transition probabilities. This makes it perfectly suitable for modelling household electricity behaviours, where the probability of performing a specific action is strongly affected by the time of the day this action is performed. For example, the probability of a household member going to sleep at 11:30 a.m. is obviously very different from the probability of going to sleep at 11:30 p.m.

For our model, we adopted the formulation described in [42], [43]. We assume that a person can occupy one of 13 possible states, correspondent to activities reported in TUS, as in Table 2.

TABLE 2: TUS action codes and corresponding states

TUS Action Codes	State
All events with location different from home	Not Present
011	Sleeping
212, 219	Studying
821	Watching TV
617,831	Listening to Radio
PC, 733	Using PC
714, 811, 812, 813, 814 ,819	Reading
311, 421	Cooking
021	Eating
312	Dishwashing
321, 422	Cleaning
332	Ironing
031	Being in Bathroom

When switching from a time step to the next, the probability of going from state i to state j is represented by a *transition probability matrix*, as shown in Figure 6. In our application, probability fluctuations over the 24 hours are reproduced by computing 144 time-dependent transition probability matrices (i.e. one every 10 minutes). The i, j elements of such matrices are computed as follows:

$$P_{i,j}(t) = \frac{n_{i,j}(t)}{n_i(t)} \tag{1}$$

where $P_{ij}(t)$ is the probability to switch from state i to state j at time t. $n_{ij}(t)$ is the number of transitions from state i to state j at time t. $n_i(t)$ is the total number of switches from state i to any state j at time t. This is done per each category of end user, respectively for working days, Saturdays and Sundays.

The distribution of sojourn times of a certain state j at a specific time t is represented by a so-called *sojourn probability vector* $S_j(t)$, as represented in Figure 6. More specifically, $S_j(t)$ is a 144-dimensional vector where each k_{th} element represents the probability of state i having a sojourn time equal to $k \times T$, with $T = 10\ min$.

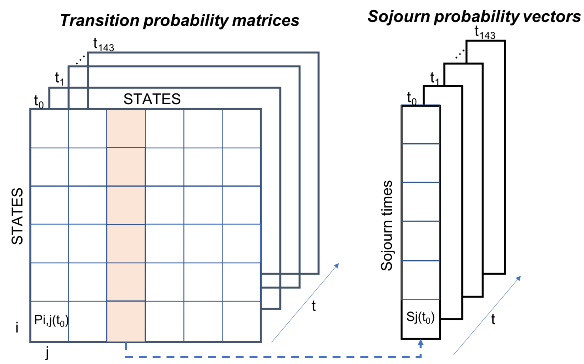


FIGURE 6: Transition probability matrices and sojourn time vectors

The cumulative probability distributions of transitioning from a specific state i to any other state at the time t are represented by *cumulative transition probability vectors*, obtained as follows:

$$Ct_i(t)(z) = \sum_{0}^{j} P_i(t)(i,j) \tag{2}$$

Likewise, the cumulative distributions of sojourn times of a new state j at time t are represented by *cumulative sojourn probability vectors*, computed as follows:

$$Cs_j(t)(w) = \sum_{0}^{k} S_j(t)(k) \tag{3}$$

Even in this case, probability fluctuations over the 24 hours are represented by computing 144 time-dependent vectors, grouped by category of end user and type of day.

Once the complete set of transition probabilities and sojourn probabilities are computed, the Monte Carlo simulation is performed by setting an initial state i and a corresponding sojourn time S_i. Then, at each time step:

1) a uniformly distributed [0,1] random number is compared with the cumulative transition probabilities of

state i at that time, in order to decide which transition is going to take place (see Figure 7);

2) a [0,1] random number is compared with the cumulative sojourn probabilities of the new state, in order to decide its duration τ_{new} (see Figure 7);
3) the system will be updated next only after a time step of τ_{new} has passed. By doing so, we obtain that, unlike classic Markov models, transitions from one state to another can happen at different time steps, as shown in Figure 8.

Simulation ends after a number of iterations defined as configuration set-up.

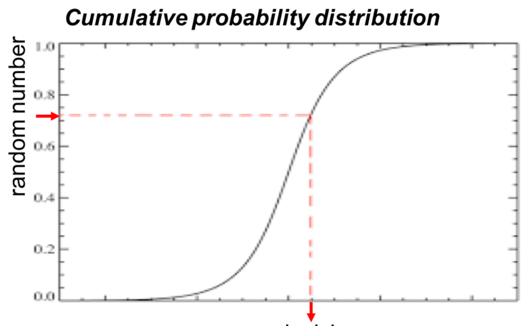


FIGURE 7: Decision process for new transition states/sojourn times

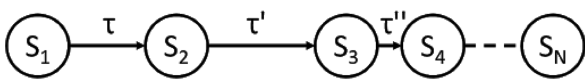


FIGURE 8: Evolution of state transitions over time

B. HOUSEHOLD BEHAVIOUR

Household Behaviour is the the second step of our *multi-scale model*. Its function is to assemble individual persons into heterogeneous families and to compute their household behaviours. To do so, each family is associated with a set of appliances, depending on the family size.

As shown in Figure 5, *Household Behaviour* consists of two computational modules: i) *Household Definition* and ii) *Probabilistic Usage of Appliances*, respectively.

The *Household Definition* module creates ad simulates a virtual population of heterogeneous households with appliances that are statistically coherent with the real population.

First, a set of virtual families are generated starting from the information of *Census data* provided by ISTAT (see Section III-D). Such information includes family compositions as well as age, gender and profession of each family member.

Once the virtual families are built, the *Household Definition* module simulates household activities and behaviours by summing up the *Single Person Behaviour*s of each family member, based on the twelve different categories of individuals defined in Section III-A.

Then, the *Household Definition* module assigns to the virtual families a set of appliances with their corresponding usage, statistically determined based on the results of the national survey on *Use of Energy* by ISTAT (see Section III-B). To do so, we make the following assumptions: i) all houses consist of a room per family member plus a living room, a kitchen and a bathroom; ii) all rooms are illuminated by one artificial light; iii) all houses have one fridge, one iron, one vacuum-cleaner and one radio.

Finally, the *Household Definition* module translates family members activities into appliances usage, based on the mapping reported in Table 1. It is worth specifying that some extra assumptions are needed for few specific activities, that are not reported in the Time Use survey. More specifically, we assume that "showering" has a 3 to 5 min duration, with a needed water flow rate of 0.2 $[l/s]$. The manual activity "dish washing" happens only when the virtual house is not equipped with a dishwasher and has a needed water flow rate of 0.15 $[l/s]$. This information is exploited to model the behaviour and energy consumption of the electric water heater during this activity (more details are given in next Section IV-C). The electric water heaters range between three different sizes: i) *Small* (30 to 50 $[l]$ capacity and 1000 W nominal power); ii) *Medium* (51 to 70 $[l]$ capacity and 1200 W nominal power); iii) *Large* (71 to 120 $[l]$ and 1500 W nominal power).

As reported in the survey on *Use of Energy*, some appliances (e.g. washing machine, dryer, dishwasher and oven) are regularly used by families throughout the week. Typically, such appliances (hereafter we will refer to them as *probabilistic appliances*) have well-defined weekly usage patterns that need to be taken into account in our *multi-scale model*. The *Probabilistic Usage of Appliances* module estimates the usage of the probabilistic appliances based on i) the composition and characteristics of the family and ii) the day of the week (weekday, Saturday or Sunday, respectively). At the beginning of each simulated day, this module estimates the activation of the probabilistic appliance. This is done as follows:

1) a uniformly distributed $[0, 1]$ random number r is extracted;
2) if $r < \frac{a}{nw}$, the appliance will be turned-on,

where a is the occurrence of the appliance's weekly usages as from *Use of Energy* survey and nw is the total number of days that are left to the end of the week.

The starting time is obtained as follows:

1) for the washing machine and dishwasher, we adopt a decision process similar to the one depicted in Figure 7: we extract a uniformly distributed $[0, 1]$ random number and compare it with the corresponding cumulative probability distributions presented in Section III-A;
2) for the dryer, we impose that the appliance can be turned on only after the end of the washing machine operation cycles;

3) for the oven, we decide the time based on the outcome of corresponding "cooking" activity of the Monte Carlo Non Homogeneous Semi-Markov chain, described in the previous section.

Additional features of our model are the following:

1) whenever all the household members are out of home, the module postpones the activation of the probabilistic appliances until someone is back;
2) to avoid more than one probabilistic appliance running at the same time, the module waits that an appliance ends before turning-on the next one;
3) the fridge is constantly active according to its operational cycles.

These features make sure that the probabilistic appliances' usage does not exceed statistical information reported by the *Use of Energy* survey (see Section III-B and Figure 4).

As depicted by Figure 5, the output of the *Probabilistic Usage of Appliances* module is given as input to *Household Definition* module to better define the set of appliances associated to each virtual family and simulate their usage.

For the sake of making our model usable by any Demand Response and/or Demand Side Management strategies, the appliances in the *Household Definition* module were categorised as follows:

1) *Non-shiftable appliances* are those devices that cannot be shifted from a time-slot to another in the same day. This category includes lighting system, TV, radio, PC, iron, vacuum-cleaner, oven, electric cooker and fridge.
2) *Shiftable appliances* are those devices that can be shifted from a time-slot to another in the same day. This category includes washing machine, dryer and dishwashers. It is important to notice that these appliances can be only shifted and not interrupted during their operating cycles.
3) *Buffering appliances* are those devices that can be turned-on or turned-off during their operating cycles. Thus, they can be considered as a buffer to regulate and modify the household power demand. So far, this category includes the electric water heater with the only constraint that it must provide hot water when requested by household members.

So far, the *Household Definition* module does not include the following appliances: i) microwave and kettle, because statistics on daily and weekly usage are missing; ii) internet modem, because its energy consumption is negligible; iii) air conditioning system, because detailed information on heat transfer of the dwelling would be needed. Nonetheless, this module is ready to support the inclusion of such appliances.

C. AGGREGATED AND SINGLE LOAD PROFILES

The *Aggregated and Single Load Profiles* is the third and last step of our model, as depicted in Figure 5. The module aims at translating household activities mapped to appliances usage (which is the outcome of the *Household Behaviour* step) into realistic household load profiles, that is the final output of our *multi-scale model*.

Results can be generated at different spatial-temporal resolutions:

1) simulation-space includes trends of load profiles over the whole week, either desegregated per single appliance or aggregated from single household up to districts and cities;
2) simulation-time ranges from $10\ min$ up to years, with a granularity spanning from $1\ s$ to $1\ hr$.

As represented in Figure 5, the simulation at this stage consists of one computation module, *Load Modelling*, which takes *Load profiles of real appliances* and *Real weather data* (see Section III) as input, together with the usage of appliances per family resulting from *Household Behaviour*. More specifically, the *Load Modelling* module associates to each appliance usage over the time the corresponding load profile from real appliances. Then, it aggregates these results based on the decided spatial-temporal resolutions, which are given as a configuration set-up.

The usage of lighting systems, and hence their energy consumption, depends on the activities performed by household members as well as by solar radiation during the day (this information is retrieved by *Real weather data*). Indeed, if an ongoing activity needs illumination and the solar radiation is lower than a threshold equal to $60\ kWh/m^2$, the *Load Modelling* module will turn the lighting on, which is a $100\ W$ lamp by default.

To model energy consumption of the electric water heater, *Load Modelling* module implements the model presented in [46], that includes four different operational status: i) *full*, ii) *water temperature lower than the given set-point*, iii) *discharging or recharging* and iv) *completely empty*.

When the electric water heater is *full*, the energy consumption is affected by the heat dispersion on its surface, as in the Equation 4:

$$\frac{dT_w}{dt} = \frac{Q_{elec} - \dot{m}C_p(T_w - T_{inlet}) + UA_{wh}(T_{amb} - T_w)}{C_w}, \quad (4)$$

where Q_{elec} is the heating capacity of the resistor in the water heater, $\dot{m}$ is the mass flow rate of the hot water, C_p is the thermal capacitance (in BTU/(lb*$^\circ F$)), T_w is the water temperature, T_{inlet} is the water temperature, UA_{wh} is the thermal conductance of the tank shell, T_{amb} is the air temperature in the room (in our model, $20^\circ\ C$) and C_w is thermal capacitance (in BTU/$^\circ F$). If the *water temperature T_w is lower than the given set-point*, the electric water heater will be turned-on until T_w reaches the set-point, according to Equation 4.

If the operational status is in *discharging or recharging*, the electric water heater will be turned-on following Equation 5:

$$\frac{dh}{dt} = a - b * h \quad (5)$$

with:

$$a = \frac{Q_{elec} + UA_{wh}(T_{amb} - T_{lower})}{C_w * (T_{upper} - T_{lower})} * H - \frac{\dot{m} * C_p}{C_w} * H \quad (6)$$

$$b = \frac{UA_{wh}}{C_w} \quad (7)$$

where $\dot{m}$ is equal to 0.2 $[l/s]$ and 0.15 $[l/s]$ for the actions "showering" and "dish washing", respectively (as anticipated in Section IV-B), H is the maximum height of the water tank, T_{lower} is equal to T_{inlet}, T_{upper} is equal to T_w and h is the height of hot water slug, as a state variable.

Finally, if the operational status is *completely empty*, the electric water heater is turned-on following Equation 4, estimating also the time needed to completely refill the tank with hot water.

When the electric water heater is in any of the operational status (ii), (iii) and (iv), it consumes a nominal power defined by the *Household Definition* module based on size (either 1000 $Watt$, 1200 $Watt$ or 1500 $Watt$).

V. EXPERIMENTAL RESULTS

In this section, we present the experimental results obtained simulating a virtual population of $1'000$ heterogeneous families, including a total number of around $2'000$ end-users, with a simulation period covering a whole calendar year. This virtual population was built based on the latest Italian Census data (see Section III-D). Thus, virtual families are statistically coherent with the distributions of family-size and end-users categories reported in Figure 9a and Figure 9b, respectively.

In the following, we first show the outcome of our simulations at different scales and time resolutions, in order to provide a complete overview of the functionality of our model. Then, to assess the goodness of the model, we validate our simulation results using real data from different sources as a reference.

A. SIMULATION OUTCOMES

In Figure 10, we show the outcome of the first step of our multi-scale model, *Single Person Behaviour*, for each of the 12 end-users categories in a randomly simulated working day. In each plot, we show in the x-axes the simulation time and in y-axes the different activities, as decided by our model. Blue bars indicate whether the end-user is respectively present at home, sleeping or active. In the latter case, red bars show the performed activity and the corresponding duration.

By analysing the plots, we can obtain some interesting insights on the behavioural patterns of different end-user categories. For example, both *Full Working Female* and *Male* are never present during the standard working times. When they are at home, they are most of the time sleeping. Even *Part-time Working Female* and *Male* have similar behaviours, in that they are never at home during morning time, which is presumably their working period. Both *Retired Female* and *Housewife* are mostly at home during the whole day and behave quite similarly. *No Working Female*, *No Working Male* and *Retired Male* have spotted presence during the day and perform similar activities when they are at home. *Kid* and *Student* are at home at similar times, but the *Student* is more active that the *Kid*, especially during the morning. Finally, the activities of the *Infant* are mostly not comparable with the ones of *Kid* and *Student*.

As described in Section IV, the outcome of *Single Person Behaviour* are exploited to simulate the overall activities and behaviours of a virtual population. Then, these behavioural patterns are combined with *Load profiles of real appliances* to provide, as final outcome of our *multi-scale model*, aggregated or desegregated households load profiles.

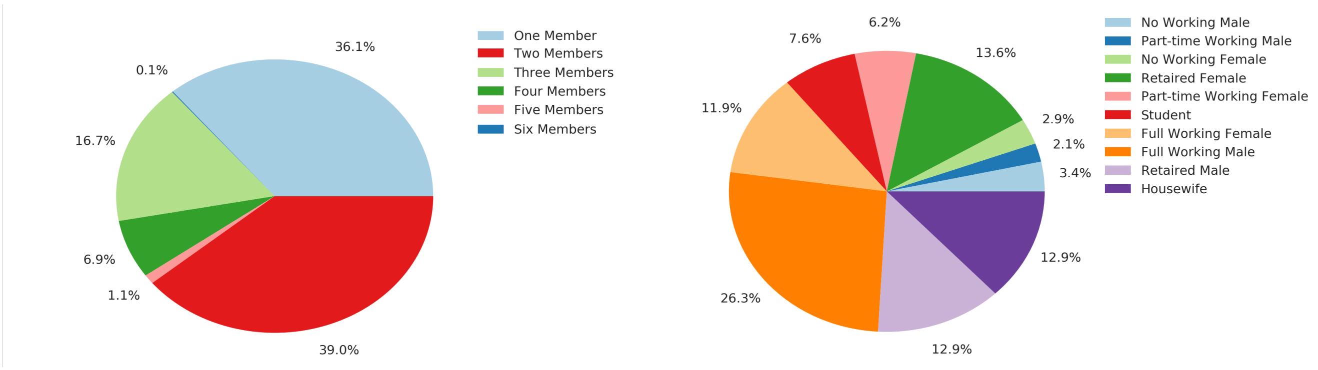
In Figure 11 and Figure 12, we show the household load profiles and the corresponding end-users behaviours (grouped into Presence or Sleeping) provided by our model. We show data corresponding to four different virtual families with respectively one (Figure 11a), two (Figure 11b), three (Figure 12a) and four (Figure 12b) members, for a simulation time of one week, chosen randomly.

As expected, by looking at Figure 11 and Figure 12, we can observe that energy consumption increases with family-size. Washing machines, dishwashers and dryers are periodically turned-on over the week, which is consistent with the real data reported by the survey on *Use of Energy*.

Figure 13 reports the aggregated load profile of $1'000$ families. We can observe that, as expected, energy consumption of fridges is constant over the whole week (about 0.06 MW). Peaks are registered in the evening (between 0.30 MW and 0.35 MW), with a spike of around 0.40 MW on Sunday, mostly due to lighting systems, dishwashers and electric water heaters.

To provide a better view of the multi-scale functionality of our system, in Figure 14 we show the outcome of our model, and more specifically the simulated household load profiles, at different space and time resolutions. In particular Figure 14a and Figure 14b show the yearly energy consumption of a district area, taken as example, respectively at the building or at the census section level. For a better understanding, it is worth noting that: i) the considered district was characterised by very non-homogeneous distributions of family sized in different buildings and census sections; ii) the non-residential areas (white-coloured, in the figures) were not included in the analysis.

In Figure 14c and Figure 14d we show respectively the daily load profiles of four representative buildings and four census sections, taken from the same district area shown above, as provided by our model. From the figures we can draw two observations: i) both the load profile groups are quite similar in trend, with peaks in the early morning and in the evening. This is an expected outcome, as these are generally the times when people prepare to go or get back from work. ii) as expected, the absolute value of the load profiles generally increases with the size of the building or census section.



(a) Distribution of family-size

(b) Distribution of end-users categories

FIGURE 9: Composition of virtual families

B. VALIDATION

To assess our model, we compared our simulated activity patterns with real data obtained from TUS survey, as introduced in Section III-A. On top of that, we validated the load curves provided by our model by exploiting three additional data-sets from the following sources:

1) RSE, an Italian research centre on electrical systems, that published reference curves of mean aggregated household load profiles of Italian users [47];
2) REMODECE [48], European project monitoring different houses across Europe by sampling both appliances and household energy consumption. For such houses, REMODECE reports both mean aggregated household load profiles and mean appliance load profiles, grouped by weekdays and weekend days.
3) IREN, an Italian Distribution System Operator [49]. IREN measured the load profile of one Italian house in Turin for a one year period. Such data were provided under a non-disclosure agreement.

Hereafter, we will refer to the aforementioned data-sets with the name of the corresponding source.

To assess the results of our model compared to reference data, we exploited the following indexes, that are widely used in descriptive statistics:

1) *Bias* is the mean value of the differences between simulated and observed values;
2) *Correlation coefficient* is a 0 to 1 number representing the strength of the correlation between the simulated and observed values;
3) *Index of Agreement* is the standardised measure of the agreement of the model with the observed values. It ranges from -1 to 1, where 0 means no agreement, -1 a perfect negative agreement and 1 a perfect positive agreement, respectively;
4) *Mean Absolute Error (MAE)* is defined as the average of the absolute difference between the model prediction and the reference data;
5) *Root Mean Square Difference (RMSD)* is the standard deviation of differences between predicted and observed values.

Figure 15 shows the time profiles of the end-users' activities. Each plot shows in the x-axes the time of the day and in the y-axes the percentage of end-users that are, respectively, i) present at home (*Presence profile*), ii) sleeping (*Sleeping profile* or iii) doing some activity at home (*Active profile*) at that specific time. Values are averaged over one year period, that is our whole simulation time. The green line refers to the outcome of our multi-scale model and the blue one to the real data obtained from the TUS surveys, plotted for comparison.

As shown by Figure 15, during weekdays the simulated profiles follow TUS data with a very good accuracy. During weekend days, the model slightly underestimates the *Presence profile* and *Active profile* in terms of absolute values, but still maintains a good overall agreement with the TUS trends.

The observed relation between simulated and real TUS data is confirmed by the results of the quantitative assessment. In Table 3, we show a whole set of indexes to measure the agreement between the simulated time profiles of the three categories of activities (for either weekdays and weekends) and the TUS data. As it can be seen from the Table, we obtained low values of Bias, MAE and RMSD, as well as very close to 1 values of the Correlation Coefficient and Index of Agreement for most the configurations. The *Active profile* in the weekends is the only one showing relatively lower values of Bias, MAE and RMSD (respectively equal to 0.01, 0.11 and 0.13). The higher variability and lower predictability of end-user behaviours during the weekends can be held responsible for the lower agreement between model and real data. Nonetheless, even in this case we registered reasonably good values of the Correlation Coefficient and of the Index of Agreement (both equal to 0.96). This confirms our considerations about the overall match between the two trends.

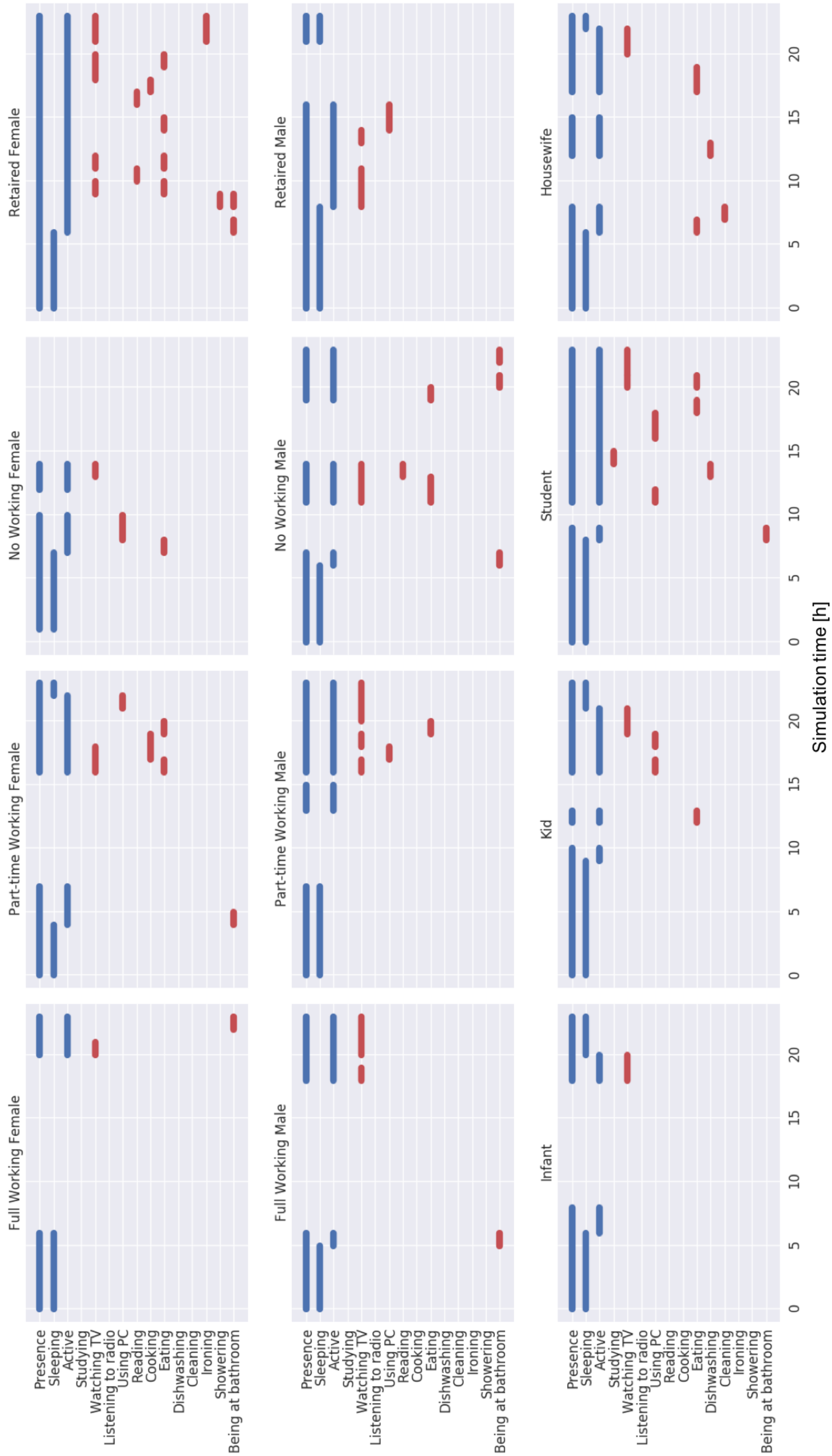
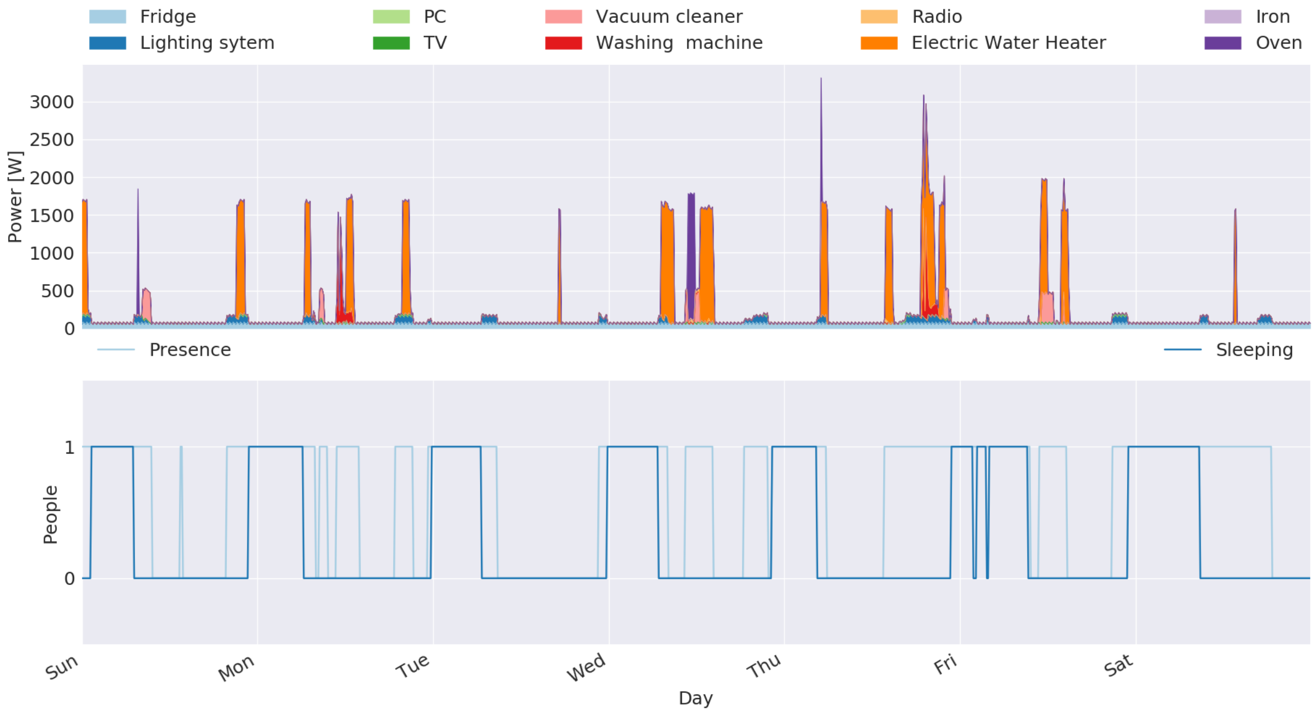
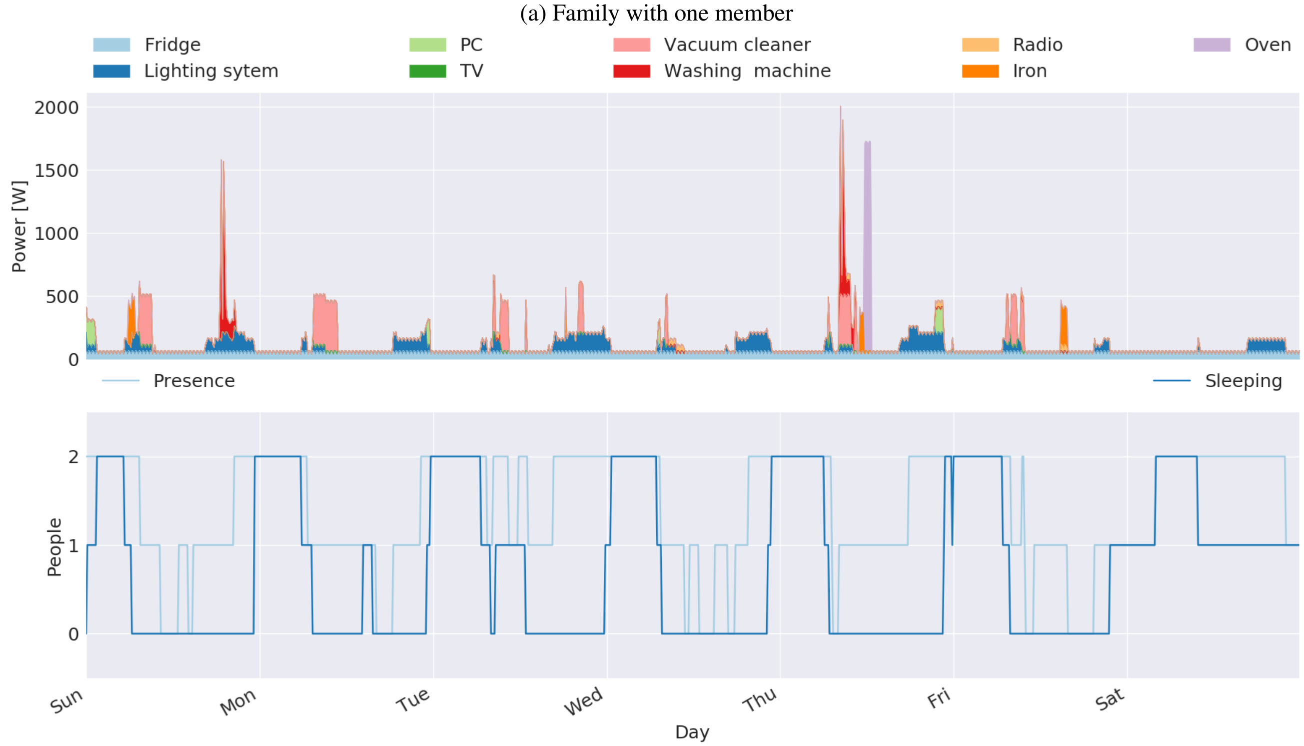


FIGURE 10: Activities of end-users in one simulated working day

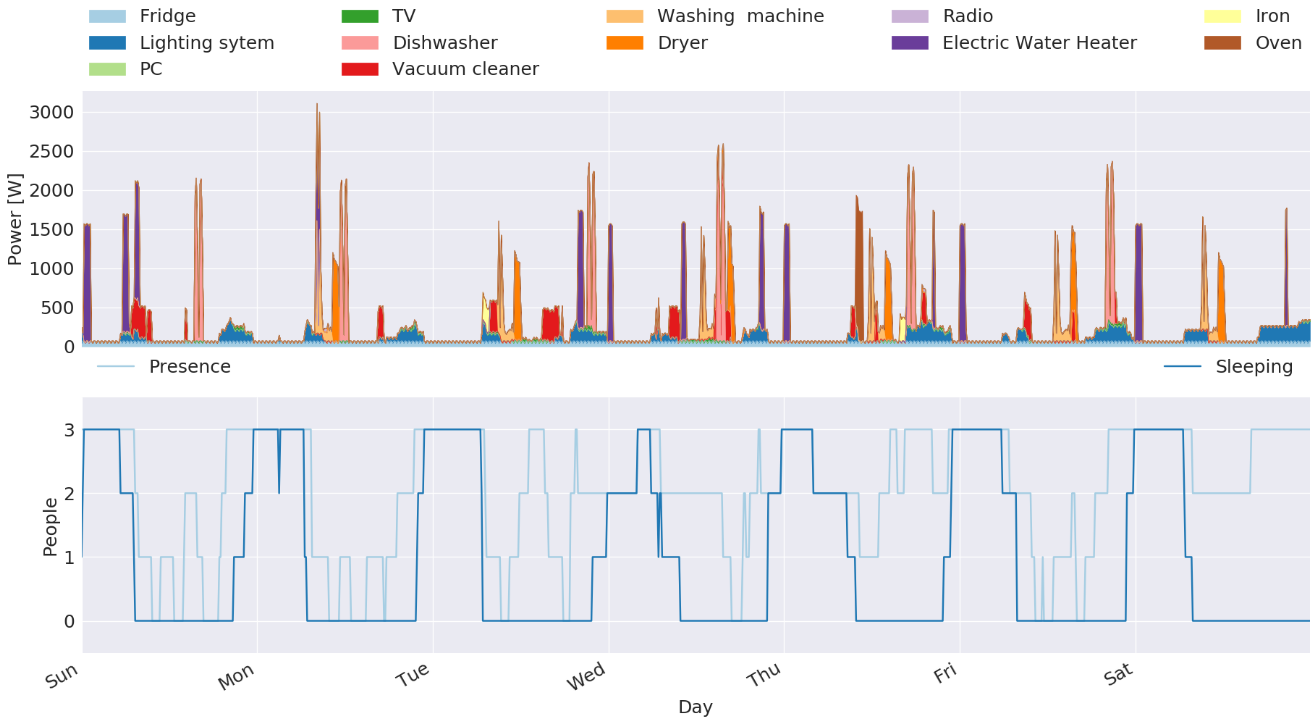


(a) Family with one member

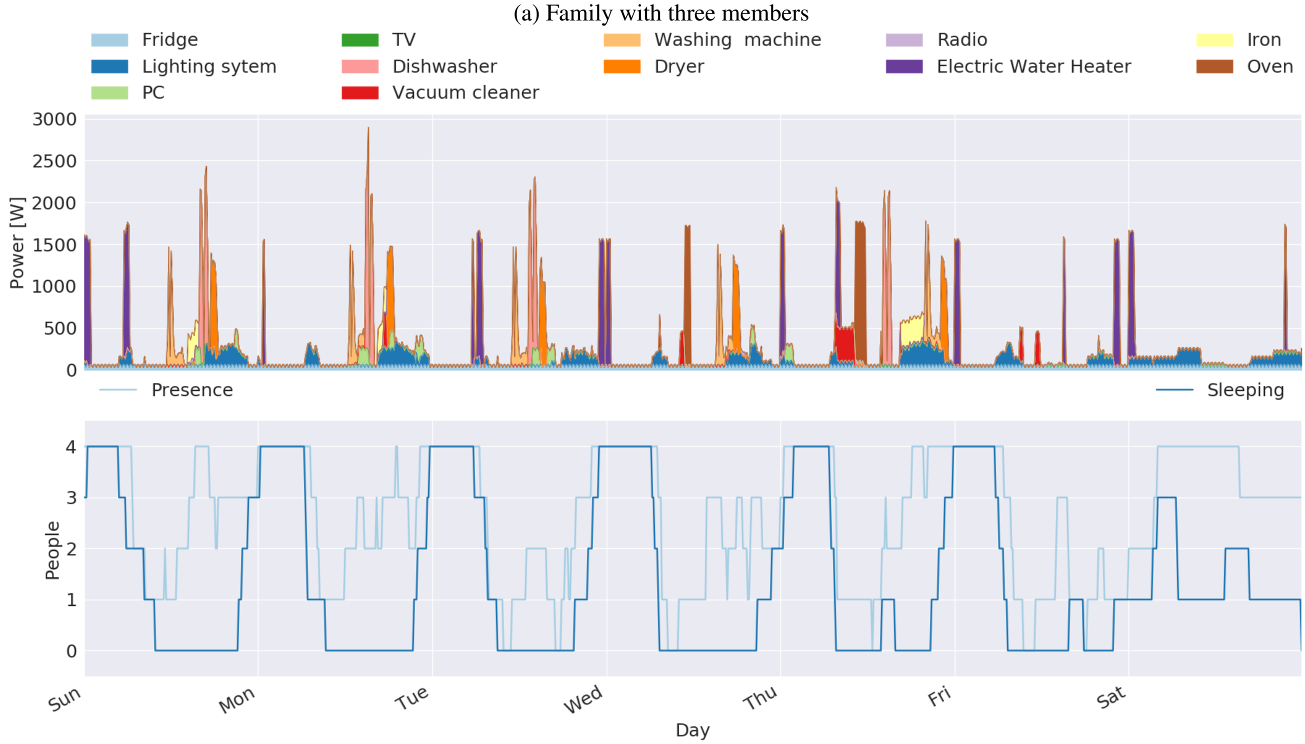


(b) Family with two members

FIGURE 11: Household load profiles of one simulated week (part 1).



(a) Family with three members



(b) Family with four members

FIGURE 12: Household load profiles of one simulated week (part 2).

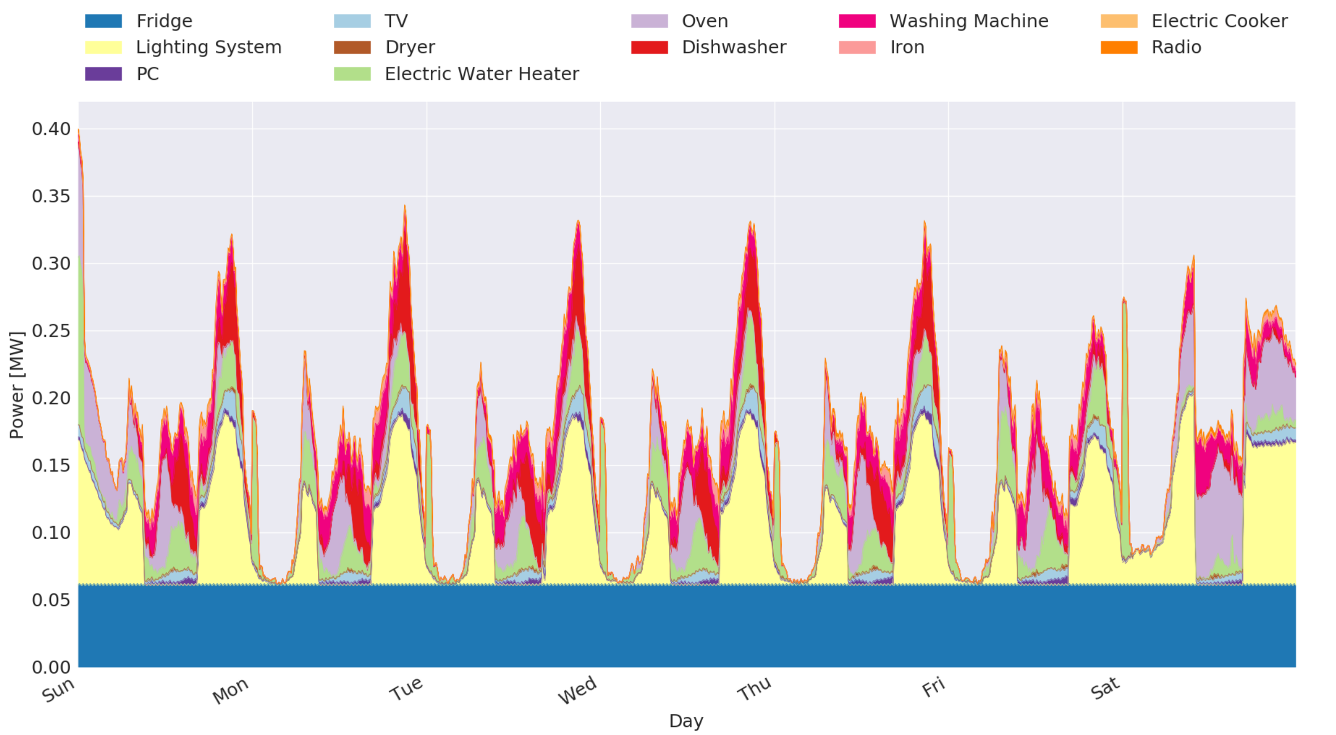
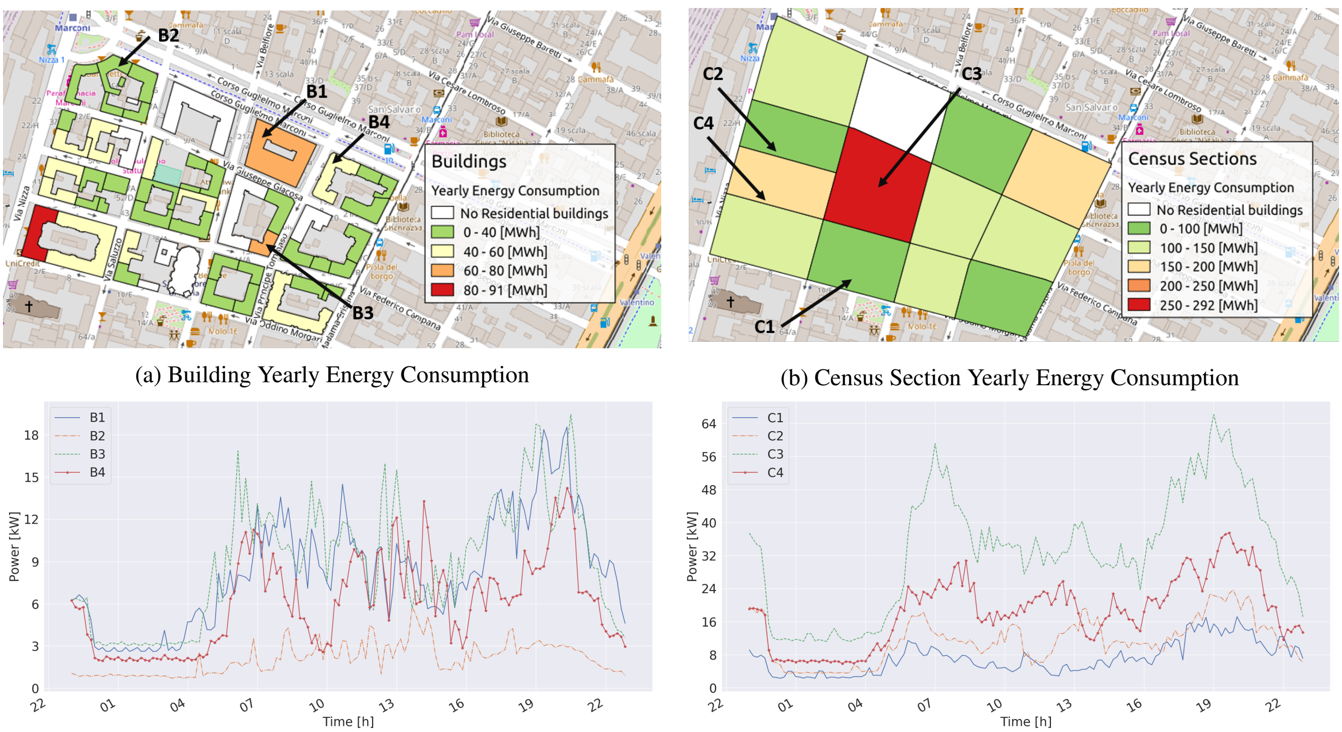


FIGURE 13: Aggregated load profile of $1'000$ simulated families



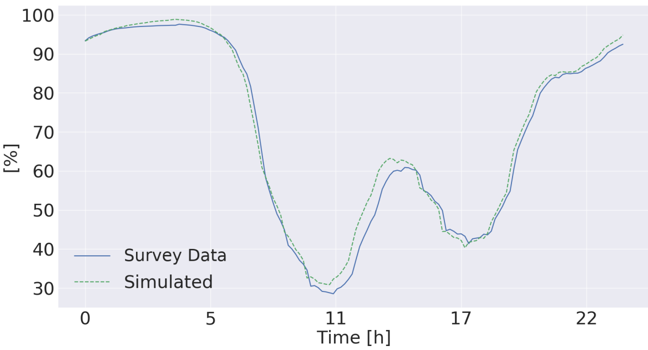
(a) Building Yearly Energy Consumption

(b) Census Section Yearly Energy Consumption

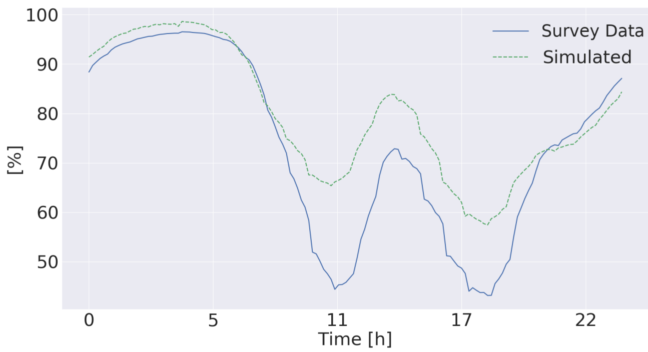
(c) Four Random Buildings - Daily Load Profile

(d) Four Random Census Sections - Daily Load Profile

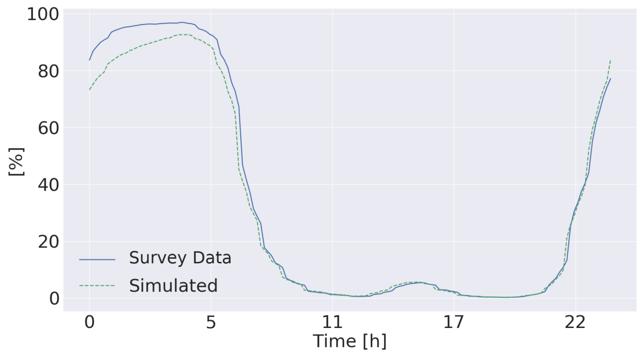
FIGURE 14: Model outcome at different spatial and temporal dimensions.



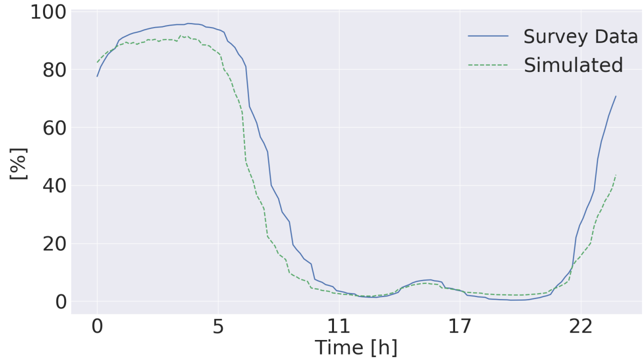
(a) Presence profile during weekdays



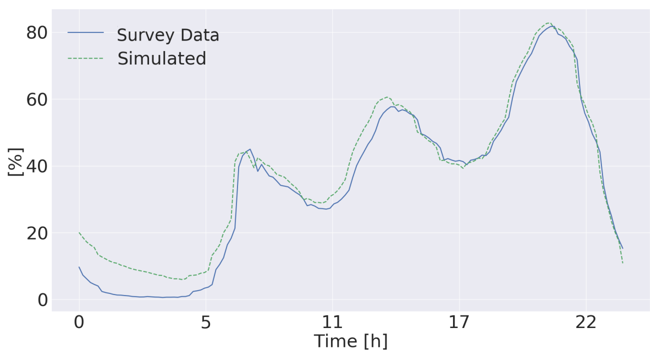
(b) Presence profile during weekend days



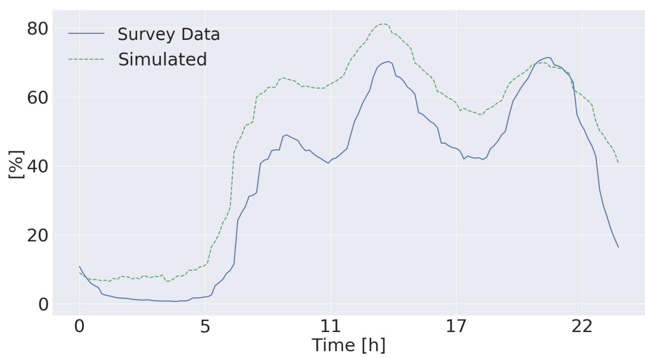
(c) Sleeping profile during weekdays



(d) Sleeping profile during weekend days



(e) Active profile during weekdays



(f) Active profile during weekend days

FIGURE 15: Presence, Active and Sleeping time profiles: Comparison between simulated results and *Time Use* data

TABLE 3: Activity time profiles: agreement between model and TUS data

Time profile	Bias	Correlation Coefficient	Index of agreement	MAE	RMSD
$Presence_{weekdays}$	0.00	0.99	1.00	0.02	0.03
$Presence_{weekends}$	0.05	0.95	0.92	0.06	0.09
$Sleeping_{weekdays}$	-0.01	1.00	1.00	0.02	0.04
$Sleeping_{weekends}$	-0.04	0.98	0.99	0.05	0.09
$Active_{weekdays}$	0.03	0.99	0.99	0.05	0.06
$Active_{weekends}$	0.08	0.96	0.96	0.11	0.13

As previously mentioned, to validate the final results of our *multi-scale model*, we compared the household load profiles of around $1'000$ simulated families with the values reported by RSE [47] and REMODECE [48], respectively. Figure 16 shows the plot of these load profiles, where x-axes represents the hours of the day and y-axes the mean aggregated value of household loads of the considered families over the one year period, normalised within a $0 - 1$ range. As shown by the plot, the three trends (respectively our model, RSE and REMODECE) have reasonably similar behaviours, with higher agreement with RSE than with REMODECE. In particular, we can observe corresponding peaks of our simulated and real data profiles in the morning and evening, with some discrepancies especially with REMODECE. Few minor fluctuations occur again between 10:00 and 17:00. The most significant gap between simulated and real profiles is registered at night times, which might be due to several reasons. There could be discrepancies in the sleeping times of the end-users, with tremendous impact on load estimation in that period of the night. On top of that, some appliances that are currently not included in the model due to lack of data (e.g. air conditioning systems) might be particularly relevant at night time. Nonetheless, the overall match between our curve and the reference curves is still reasonably high, in that it is comparable with the match between the two references themselves.

More in-depth analysis can be made by looking at Table 4, where we report the quantitative assessment of the similarity between our simulation and, respectively, REMODECE and RSE aggregated profiles. For reference we also show values of the comparison between REMODECE and RSE reference curves. From the values reported in Table 4, we can draw the following considerations. i) our model shows a reasonably good similarity with both RSE and REMODECE. Again, we can see a better agreement with RSE aggregated values than with REMODECE. A possible explanation is that, while RSE reports measurements taken in Italy, which is consistent with the data-sources of our model, REMODECE aggregated values refer to European measurements, maybe showing behavioural patterns that are slightly different than the Italian ones. ii) the last line of the Table shows that REMODECE and RSE profiles do not match perfectly, which confirms the observations made before. For example, MAE and RMSD values between REMODECE and RSE are quite comparable with values between our simulation and RSE.

TABLE 4: Aggregated simulation: agreement with reference data-sets.

	Bias	Correlation Coefficient	Index of agreement	MAE	RMSD
Simulation vs Remodece	-0.04	0.84	0.91	0.09	0.13
Simulation vs RSE	-0.02	0.91	0.95	0.06	0.09
Remodece VS RSE	-0.02	0.96	0.98	0.05	0.07

To deepen our analysis, we also compared desegregated appliances load profiles of our simulated families with the same values measured by REMODECE [48], either in weekdays and weekends. Figure 17 and Figure 18 show the plots related to *Non-shiftable appliances* and *Shiftable appliances*, respectively (see the categorisation introduced in Section IV-B). In these figures, x-axes represents the times of the day and y-axes reports values of appliance loads, normalised within a $0 - 1$ range.

As shown by the figures, simulated load profiles of both *Non-shiftable appliances* and *Shiftable appliances* follow with few fluctuations the measured load profiles. Nonetheless, the match between the trend of simulated and measured curves is still acceptable. These trends report the usage of appliances through their load consumption, intrinsically reflecting the behavioural patterns of the end-users at home. The similarity of simulated and real life-profiles proves that the model provides a realistic picture of energy consumption in residential houses either in weekdays and weekends.

To provide a quantitative assessment of our results, we show in Table 5 indexes measuring the agreement between the simulation single appliances and REMODECE's measurements. As we can gather from the Table, the values show a good similarity of our model and REMODECE for most of the appliances. Only exceptions are i) the use of the dryer during weekends, ii) the use of dishwasher and vacuum cleaner in either weekdays and weekends. While the reduced agreement during the weekends is not surprising, as the behavioural patterns of the end-users are intrinsically less predictable during non/working days, it is important to highlight again that REMODECE and TUS (which is the main data-source of our model) are uncorrelated datasets. More specifically, TUS data represent daily activities of Italian users, while REMODECE data result from the monitoring of electricity load profiles of European residential customers. On this basis, the discrepancies with our model are reasonably low.

TABLE 5: Single appliances simulation: agreement with REMODECE

	Bias	Correlation Coefficient	Index of agreement	MAE	RMSD
$WashingMachine_{weekdays}$	-0.18	0.85	0.85	0.20	0.24
$WashingMachine_{weekends}$	0.02	0.95	0.97	0.08	0.10
$Dryer_{weekdays}$	-0.13	0.90	0.90	0.15	0.19
$Dryer_{weekends}$	-0.05	0.70	0.83	0.18	0.25
$Dishwasher_{weekdays}$	-0.26	0.77	0.73	0.26	0.32
$Dishwasher_{weekends}$	-0.27	0.74	0.71	0.27	0.33
$Lightingsystem_{weekdays}$	-0.05	0.95	0.96	0.10	0.11
$Lightingsystem_{weekends}$	-0.03	0.89	0.94	0.12	0.15
$PC_{weekdays}$	-0.22	0.90	0.79	0.22	0.25
$PC_{weekends}$	-0.15	0.91	0.86	0.19	0.22
$TV_{weekdays}$	-0.08	0.92	0.93	0.10	0.13
$TV_{weekends}$	-0.10	0.94	0.94	0.11	0.14
$Vacuumcleaner_{weekdays}$	-0.18	0.52	0.67	0.24	0.32
$Vacuumcleaner_{weekends}$	0.02	0.53	0.73	0.16	0.26

Finally, Figure 19 shows the auto-correlation trends of two mean weekly load profiles belonging respectively to a simulated house (dashed-line) and to a real-world house measured by IREN (solid-line). These profiles were obtained by averaging the energy consumption over one year by each day of the week. The y-axes of the plot reports the auto-correlation coefficient values, which provide the degree of similarity between a time series and a lagged version of itself over successive intervals, while the x-axes reports the time lag. As shown in Figure 19, both our simulation and IREN trends have high auto-correlation values due to non-stationary data, which demonstrates the non-randomness of the simulation. This consideration is confirmed by the cyclical patters of the two profiles over the 24 hours, which also reflects a high correlation between the simulated and the real load profile. Hence, we can conclude that our *multi-scale model* is able to provide realistic weekly residential load profiles.

VI. CONCLUSION

In this paper, we proposed a *multi-scale model* that simulates residential load consumption with different spatial-temporal

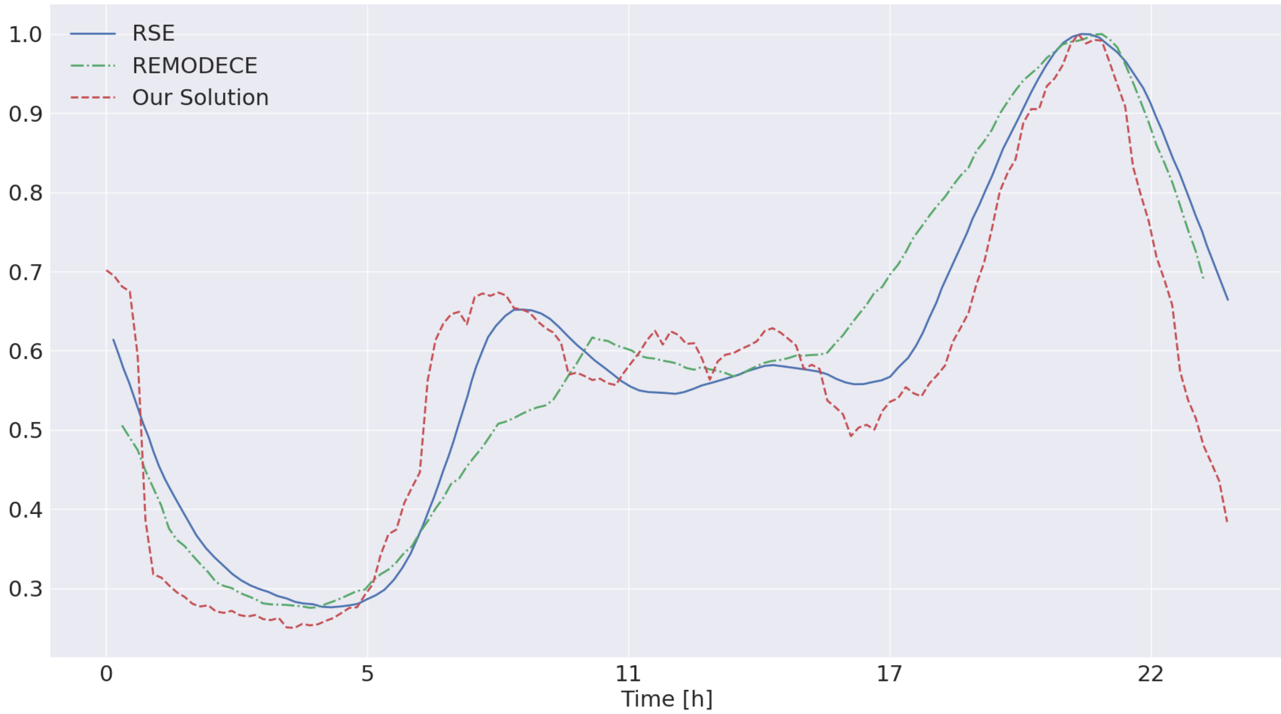


FIGURE 16: Aggregated load trends: comparison between our simulated results, RSE and REMODECE (normalised values)

resolutions following a *Bottom-up* approach. In its core, our model exploits Monte Carlo Non Homogeneous Semi-Markov chains i) to simulate the activities and behaviours of each household inhabitant (or end-user) and ii) to associate load consumption trends of the used appliances to each activity. The proposed methodology takes into account the probability to perform an action during the day as well as its duration in time. To model end-user's activities and to estimate distributions of heterogeneous families and corresponding appliances, we analysed the national survey on *Time Use* and other available statistical information on population. The final outcomes of our model are residential load profiles, grouped by weekdays and weekends, with multi-level aggregation in time and space. These outcomes can be used to design and test novel services for smart grids, such as NILM, State Estimation, Demand Response and Demand Side Management, even in the context of distributed co-simulation frameworks for complex-system analysis of power networks [50].

To assess the goodness of our model, we compared the activity patterns provided by our simulations with the results of the survey on *Time Use*. Then, we validated the load curves obtained with our model, both aggregated and desegregated, with real load trends provided by RSE, REMODECE and IREN, respectively. Our experimental analysis demonstrated that the proposed *multi-scale model* generates realistic residential load profiles, providing detailed information on energy consumption in residential houses in either weekdays and weekends.

As already discussed, the performance of our simulations is currently limited by lacking data on specific appliances that have a significant impact on household energy consumption. More specifically, the model would majorly benefit from additional information on i) energy load profiles of appliances belonging to different energetic classes and ii) distribution of such appliances in different families. This information is currently not provided by any national surveys but might be available very soon. On these premises, we believe that the margins for improvement in the near future are very large.

REFERENCES

[1] United Nations, FCCC, "Adoption of the Paris Agreement. Proposal by the President," 2015, Available: http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf.

[2] L. Dusonchet and E. Telaretti, "Economic analysis of different supporting policies for the production of electrical energy by solar photovoltaics in western european union countries," Energy Policy, vol. 38, no. 7, pp. 3297–3308, 2010.

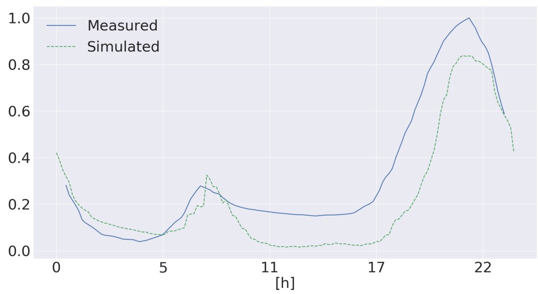
[3] International Energy Agency., "Tracking clean energy progress 2016," 2016. [Online]. Available: http://www.iea.org/publications/freepublications/publication/TrackingCleanEnergyProgress2016.pdf

[4] A. Bahmanyar, S. Jamali, A. Estebsari, E. Pons, E. Bompard, E. Patti, and A. Acquaviva, "Emerging smart meters in electrical distribution systems: Opportunities and challenges," in Electrical Engineering (ICEE), 2016 24th Iranian Conference on. IEEE, 2016, pp. 1082–1087.

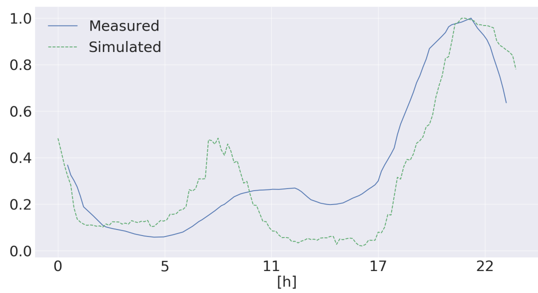
[5] M. A. Mengistu, A. A. Girmay, C. Camarda, A. Acquaviva, and E. Patti, "A cloud-based on-line disaggregation algorithm for home appliance loads," IEEE Transactions on Smart Grid, 2018.

[6] B. Zhao, L. Stankovic, and V. Stankovic, "On a training-less solution for non-intrusive appliance load monitoring using graph signal processing," IEEE Access, vol. 4, pp. 1784–1799, 2016.

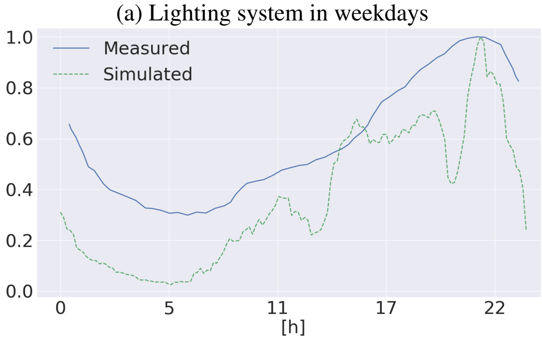
[7] M. Pau, E. Patti, L. Barbierato, A. Estebsari, E. Pons, F. Ponci, and A. Monti, "A cloud-based smart metering infrastructure for distribution grid services and automation," Sustainable Energy, Grids and Networks, 2017.



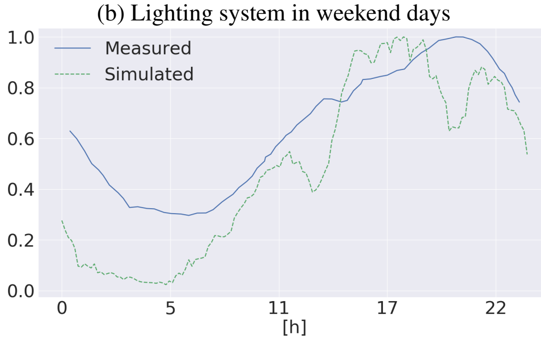
(a) Lighting system in weekdays



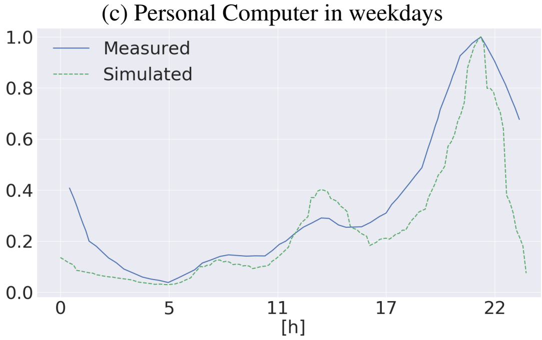
(b) Lighting system in weekend days



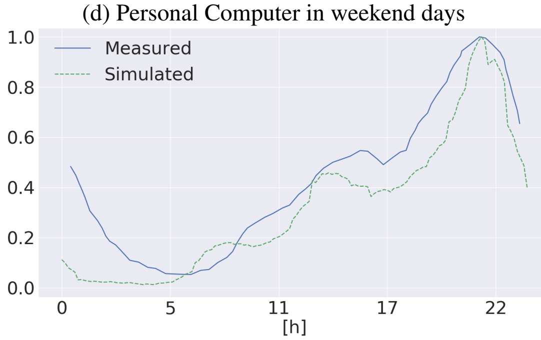
(c) Personal Computer in weekdays



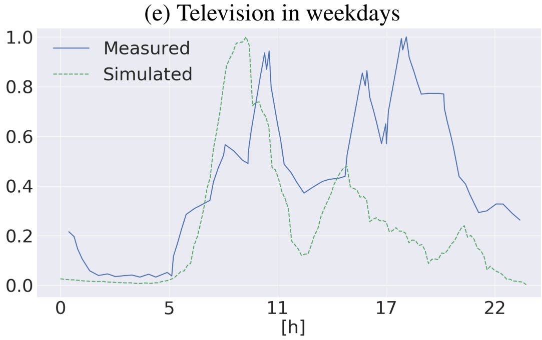
(d) Personal Computer in weekend days



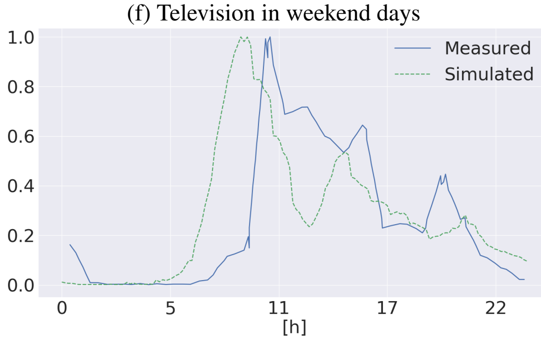
(e) Television in weekdays



(f) Television in weekend days

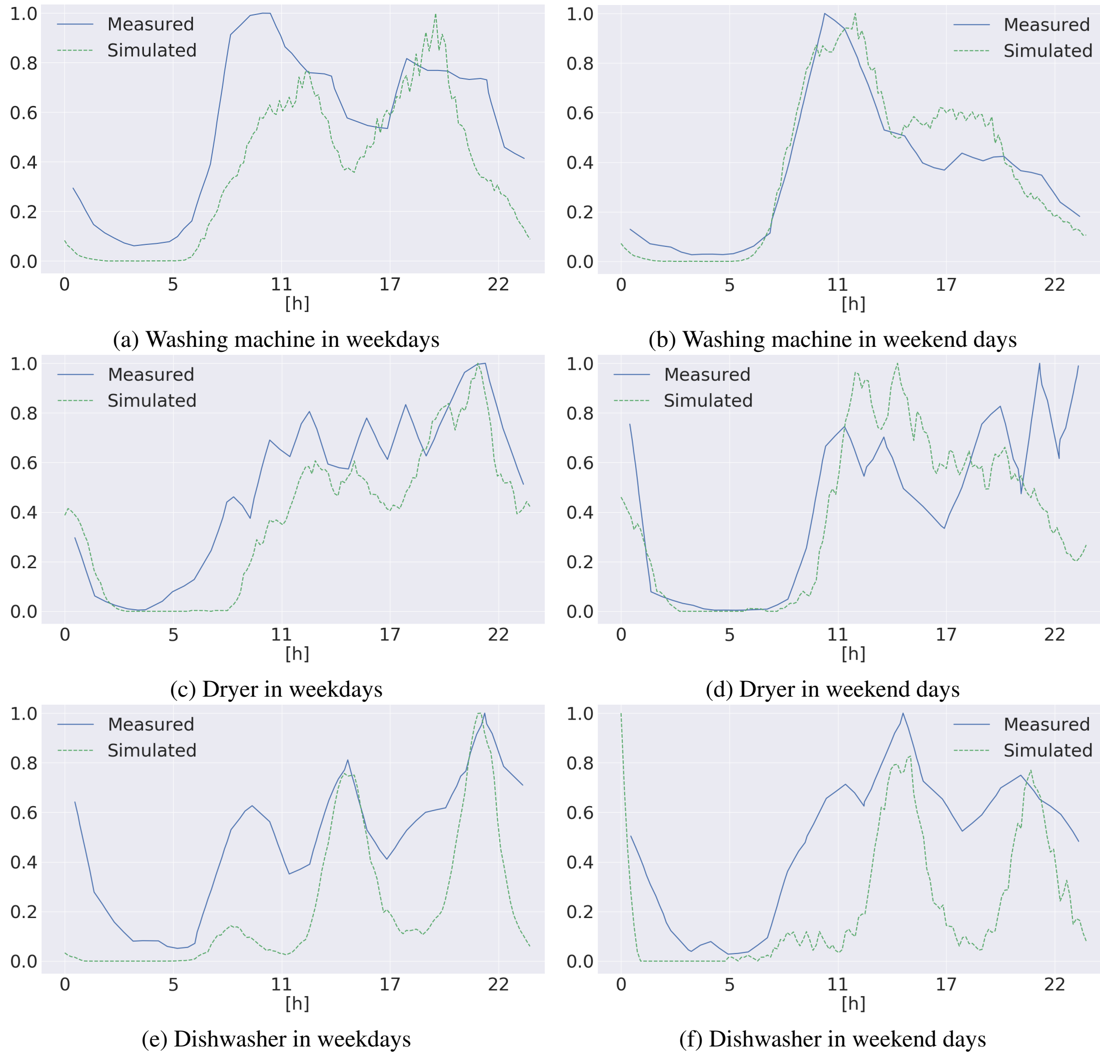


(g) Vacuum-cleaner in weekdays



(h) Vacuum-cleaner in weekend days

FIGURE 17: *Non-shiftable* appliance load profiles: Comparison between simulated results and REMODECE (normalised values)



(a) Washing machine in weekdays
(b) Washing machine in weekend days
(c) Dryer in weekdays
(d) Dryer in weekend days
(e) Dishwasher in weekdays
(f) Dishwasher in weekend days

FIGURE 18: *Shiftable* appliance load profiles: comparison between simulated results and REMODECE (normalised values)

[8] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 570–582, 2015.

[9] C.-C. Lin, D.-J. Deng, W.-Y. Liu, and L. Chen, "Peak load shifting in the internet of energy with energy trading among end-users," IEEE Access, vol. 5, pp. 1967–1976, 2017.

[10] A. Monticelli, "Electric power system state estimation," Proceedings of the IEEE, vol. 88, no. 2, pp. 262–282, 2000.

[11] A. Chiş, J. Rajasekharan, J. Lundén, and V. Koivunen, "Demand response for renewable energy integration and load balancing in smart grid communities," in Signal Processing Conference (EUSIPCO), 2016 24th European. IEEE, 2016, pp. 1423–1427.

[12] G. Hoogsteen, A. Molderink, J. L. Hurink, and G. J. Smit, "Generation of flexible domestic load profiles to evaluate demand side management approaches," in Energy Conference (ENERGYCON), 2016 IEEE International. IEEE, 2016, pp. 1–6.

[13] M. Nijhuis, M. Gibescu, and J. Cobben, "Bottom-up markov chain monte carlo approach for scenario based residential load modelling with publicly available data," Energy and Buildings, vol. 112, pp. 121–129, 2016.

[14] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Development of low-voltage load models for the residential load sector," IEEE Trans. Power Syst, vol. 29, no. 5, pp. 2180–2188, 2014.

[15] J. Widén and E. Wäckelgård, "A high-resolution stochastic model of domestic activity patterns and electricity demand," Applied Energy, vol. 87, no. 6, pp. 1880–1892, 2010.

[16] M. Bladh and H. Krantz, "Towards a bright future? household use of electric light: A microlevel study," Energy Policy, vol. 36, no. 9, pp. 3521–3530, 2008.

[17] A. Wright and S. Firth, "The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations," Applied Energy, vol. 84, no. 4, pp. 389–403, 2007.

[18] W. Kong, Z. Dong, G. Chen, and Y. Jia, "A rule based domestic load profile generator for future smart grid," in Power Engineering Conference (AUPEC), 2014 Australasian Universities. IEEE, 2014, pp. 1–5.

[19] A. Grandjean, J. Adnot, and G. Binet, "A review and an analysis of the residential electric load curve models," Renewable and Sustainable Energy Reviews, vol. 16, no. 9, pp. 6539–6565, 2012.

[20] M. Muratori, M. C. Roberts, R. Sioshansi, V. Marano, and G. Rizzoni, "A highly resolved modeling technique to simulate residential power demand," Applied Energy, vol. 107, pp. 465–473, 2013.

[21] G. Strbac, "Demand side management: Benefits and challenges," Energy policy, vol. 36, no. 12, pp. 4419–4426, 2008.

[22] A. Capasso, W. Grattieri, R. Lamedica, and A. Prudenzi, "A bottom-up approach to residential load modeling," IEEE Transactions on Power Systems, vol. 9, no. 2, pp. 957–964, 1994.

[23] E. Carpaneto and G. Chicco, "Probabilistic characterisation of the aggregated residential load patterns," IET Generation, Transmission & Distribution, vol. 2, no. 3, pp. 373–382, 2008.

[24] J. K. Gruber and M. Prodanovic, "Residential energy load profile generation using a probabilistic approach," in Computer Modeling and Simula-

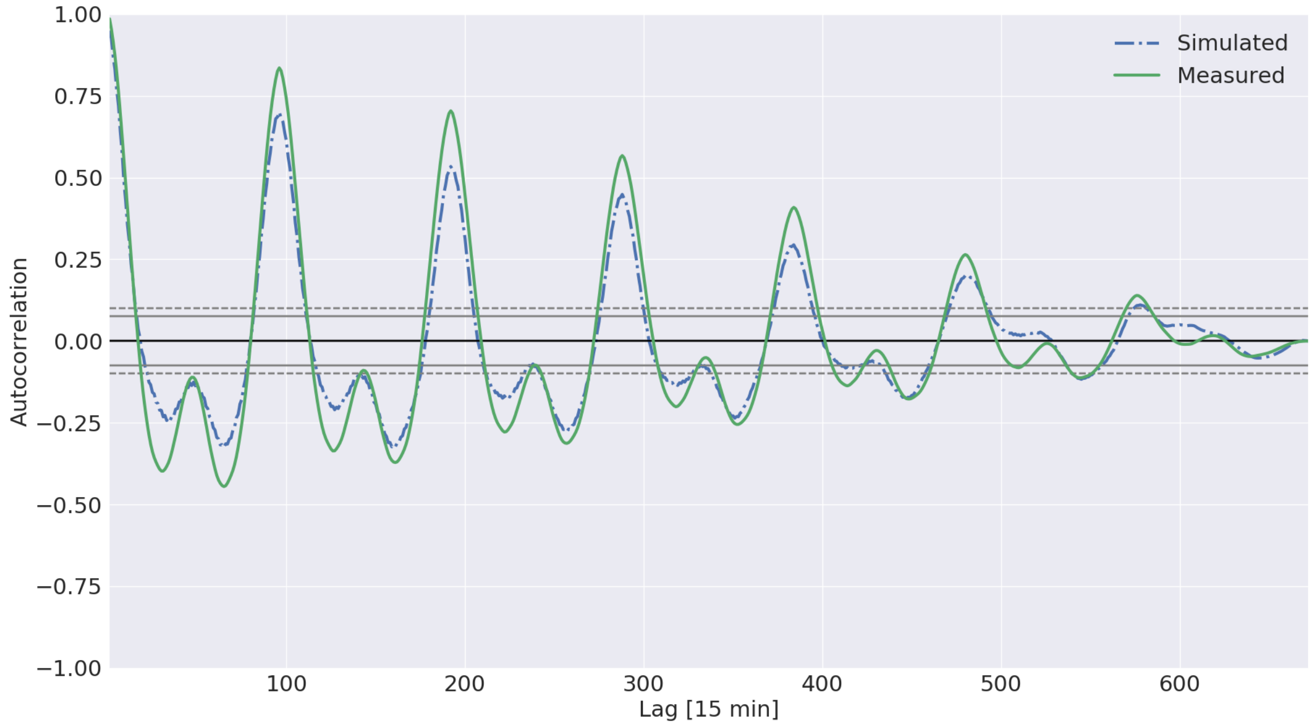


FIGURE 19: Auto-correlation of our simulations (dashed-line) compared with a real-world house measured by IREN (solid-line)

tion (EMS), 2012 Sixth UKSim/AMSS European Symposium on. IEEE, 2012, pp. 317–322.

[25] Y. Lan, J. Wu, and Z. Tang, "Generation of domestic load profiles using appliances' activating moments," in Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES. IEEE, 2014, pp. 1–5.

[26] S. Klein, W. Beckman, J. Mitchell, J. Duffie, N. Duffie, T. Freeman, J. Mitchell, J. Braun, B. Evans, J. Kummer et al., "Trnsys version. 16," Solar Energy Laboratory, University of Wisconsin, 2004.

[27] J. M. Gonzalez, E. Pouresmaeil, C. A. Canizares, K. Bhattacharya, A. Mosaddegh, and B. Solanki, "Smart residential load simulator for energy management in smart grids," IEEE Transactions on Industrial Electronics, 2018.

[28] M. N. Iqbal and L. Kütt, "End-user electricity consumption modelling for power quality analysis in residential building," in 2018 19th International Scientific Conference on Electric Power Engineering (EPE). IEEE, 2018, pp. 1–6.

[29] B. J. Johnson, M. R. Starke, O. A. Abdelaziz, R. K. Jackson, and L. M. Tolbert, "A method for modeling household occupant behavior to simulate residential energy consumption," in Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES. IEEE, 2014, pp. 1–5.

[30] J. Widén, A. M. Nilsson, and E. Wäckelgård, "A combined markov-chain and bottom-up approach to modelling of domestic lighting demand," Energy and Buildings, vol. 41, no. 10, pp. 1001–1012, 2009.

[31] J. Widén, M. Lundh, I. Vassileva, E. Dahlquist, K. Ellegård, and E. Wäckelgård, "Constructing load profiles for household electricity and hot water from time-use data-modelling approach and validation," Energy and Buildings, vol. 41, no. 7, pp. 753–768, 2009.

[32] I. Richardson, M. Thomson, D. Infield, and C. Clifford, "Domestic electricity use: A high-resolution energy demand model," Energy and buildings, vol. 42, no. 10, pp. 1878–1887, 2010.

[33] I. Richardson, M. Thomson, and D. Infield, "A high-resolution domestic building occupancy model for energy demand simulations," Energy and buildings, vol. 40, no. 8, pp. 1560–1566, 2008.

[34] I. Richardson, M. Thomson, D. Infield, and A. Delahunty, "Domestic lighting: A high-resolution energy demand model," Energy and Buildings, vol. 41, no. 7, pp. 781–789, 2009.

[35] R. McKenna, L. Hofmann, M. Kleinebrahm, and W. Fichtner, "A stochastic multi-energy simulation model for uk residential buildings," Energy and Buildings, vol. 168, pp. 470–489, 2018.

[36] A. Sancho-Tomás, M. Sumner, and D. Robinson, "A generalised model of electrical energy demand from small household appliances," Energy and Buildings, vol. 135, pp. 350–366, 2017.

[37] ISTAT - Italian National Institut of Statistics, "Multiscope on families: use of time - microfates for public use," 2013, available at https://www.istat.it/it/archivio/202531.

[38] ISTAT, Italian National Institut of Statistics, "Survey on energy consumption of families: microdates for public uses," 2013, available at https://www.istat.it/it/archivio/203344.

[39] ISTAT (Italian National Institut of Statistics), "General census of population and housing at 2011 - sample at 1 percent," 2013, available at https://www.istat.it/it/archivio/196129.

[40] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, "On the accuracy of appliance identification based on distributed load metering data," in Sustainable Internet and ICT for Sustainability (SustainIT), 2012. IEEE, 2012, pp. 1–9.

[41] Weather Underground, http://www.wunderground.com.

[42] J. Janssen, A. Blasi, and R. Manca, "Discrete time homogeneous and non-homogeneous semi-markov reliability models." in Proceedings of the Third International Conference on Mathematical Methods in Reliability, Trondheim, Norway. Citeseer, 2002.

[43] J. Janssen and R. De Dominicis, "Finite non-homogeneous semi-markov processes: Theoretical and computational aspects," Insurance: Mathematics and Economics, vol. 3, no. 3, pp. 157–165, 1984.

[44] J. Torriti, "A review of time use models of residential electricity demand," Renewable and Sustainable Energy Reviews, vol. 37, pp. 265–272, 2014.

[45] J. Rolink and C. Rehtanz, "Estimation of the availability of grid-connected electric vehicles by non-homogeneous semi-markov processes," in PowerTech, 2011 IEEE Trondheim. IEEE, 2011, pp. 1–7.

[46] R. Diao, S. Lu, M. Elizondo, E. Mayhorn, Y. Zhang, and N. Samaan, "Electric water heater modeling and control strategies for demand response," in Power and Energy Society General Meeting, 2012 IEEE. IEEE, 2012, pp. 1–8.

[47] S. Maggiore, "Analisi di impatto dellintroduzione della tariffa bioraria obbligatoria. prot: 13000580 rse spa. 2012."

[48] A. e. a. Almeida, "Final report remodece-residential monitoring to decrease energy use and carbon emissions in europe (ieea program funded project)," University of Coimbra, Dep. of Electrical Engineering, Pólo II, vol. 3030, p. 290, 2008.

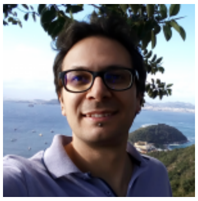
[49] Iren, 2018, available at https://www.gruppoiren.it/home.

[50] L. Bottaccioli, A. Estebsari, E. Pons, E. Bompard, E. Macii, E. Patti, and A. Acquaviva, "A flexible distributed infrastructure for real-time co-simulations in smart grids," IEEE Transactions on Industrial Informatics, vol. 13, no. 6, pp. 3265–3274, 2017.

LORENZO BOTTACCIOLI is Assistant Professor at Politecnico di Torino. He received the B.Sc. in Mechanical Engineering in 2011 at University of Perugia; the M.Sc. in Energetic and Nuclear Engineering in 2014 at Politecnico di Torino; and the PhD, cum laude, in Computer Engineering at Politecnico di Torino. His main research interests concern Smart Energy, Smart City and Smart Communities with particular attention on software solutions i) for planning, analysing and optimising smart energy system and ii) for spatial representation of energy information.

SANTA DI CATALDO received her Biomedical Engineering degree (summa cum laude) from Politecnico di Torino, Italy, in 2006. In 2007 she joined the Department of Control and Computer Engineering (DAUIN) in Politecnico di Torino as a Research Assistant. She holds a Ph.D. in Systems and Computer Engineering from the same university since April 2011. She is currently Assistant Professor (tenure track). Her main research interests are biomedical image processing, machine learning and heterogeneous data integration, including techniques for pattern recognition, image segmentation, feature quantification and classification with special regards to medical and biological applications.

ANDREA ACQUAVIVA is Associate Professor at Politecnico di Torino, Italy. He received the Ph.D. degree in electrical engineering from the University of Bologna, Italy, in 2003. In 2003, he became an Assistant Professor with the Computer Science Department, University of Urbino, Italy. From 2005 to 2007, he was a Visiting Researcher with the Ecole Polytechnique Federale de Lausanne, Switzerland. In 2006, he joined the Department of Computer Science, University of Verona, Italy. He has been with the Interuniversity Department of Regional and Urban Studies and Planning, Politecnico di Torino. His research interests focus i) on parallel computing for distributed embedded systems such as multicore and sensor networks; ii) on software solutions for Smart Cities; iii) on simulation and analysis of biological systems using parallel architectures. In the fields above, he has authored over 200 scientific publications (between 2000 and 2018).



EDOARDO PATTI is Assistant Professor at Politecnico di Torino. He received both M.Sc. and Ph.D. degrees in Computer Engineering at Politecnico di Torino in 2010 and 2014, respectively. During the Academic Year 2014/2015, he was Academic Visiting at the University of Manchester. His research interests concern: i) Ubiquitous Computing; ii) Internet of Things; iii) Smart Systems and Cities; iv) Software architectures with particular emphasis on infrastructure for Ambient Intelligence; v) Software solutions for simulating and optimising energy systems; vi) Software solutions for energy data visualisation to increase user awareness. In the fields above, he has authored over 70 scientific publications (between 2011 and 2018).

• • •