Minimal Supervision, Maximum Accuracy: TabPFN for Microcontroller Performance Prediction

Residential load modeling with generative adversarial networks

Marco Castangia, Benedetta Giorgi, Stefano Quer, Lorenzo Bottaccioli and Edoardo Patti

***Abstract*—Precise residential load modeling is indispensable for crafting effective demand-side management strategies and simulating realistic household power consumption under diverse conditions. This paper introduces a novel generative framework, leveraging the power of Generative Adversarial Networks (GANs), to synthesize highly realistic daily activity patterns. By training on detailed Italian time-use data, the model captures nuanced behavioral statistics, reflecting the inherent variability of human routines. Furthermore, incorporating conditional generation based on the day of the week allows for contextually rich and adaptable simulations, capturing weekly lifestyle variations. Household power profiles are reconstructed by meticulously mapping the generated activities to the characteristic power signatures of common household appliances, resulting in simulations that exhibit strong concordance with empirical load data at both granular, appliance-level, and aggregated household levels. Critically, our GAN-based approach demonstrably accelerates simulation throughput compared to conventional Markov chain methodologies, enabling the efficient and scalable analysis of complex residential energy scenarios, and opening avenues for real-time applications and large-scale urban energy studies.**

***Index Terms*—residential load modeling, residential electricity use, occupant modeling, load simulation, generative adversarial network, markov chain**

I. Introduction

Residential energy consumption plays a pivotal role in the global transition toward net-zero emissions, as outlined by governments worldwide over the past few decades. Residential buildings account for nearly one-third of global energy demand, making their decarbonization essential for achieving carbon neutrality [1]. The shift toward zero-carbon-ready buildings is a critical milestone in this process, requiring innovative solutions for energy efficiency and grid integration.

Load modeling is a powerful tool for analyzing and forecasting changes in residential energy demand due to evolving consumer needs. For example, the widespread adoption of electric vehicles is expected to alter household load profiles significantly, concentrating energy consumption during night-time charging hours [2]. Similarly, the growing use of Heating, Ventilation, and Air Conditioning (HVAC) systems, driven by the transition to heat pumps and electric boilers replacing natural gas-based systems, will further reshape energy demand [3]. Additionally, rising global temperatures due to CO_2 emissions will likely increase the demand for cooling systems. These shifts and the growing integration of renewable energy sources underscore the need for strategic interventions in power distribution networks to ensure reliability and efficiency [4]. Beyond grid planning, load modeling plays a critical role in the analysis and design of demand response (DR) strategies. By simulating end-user behavior, it becomes possible to evaluate the most effective interventions to reduce energy demand in residential buildings. Importantly, the effectiveness of these strategies is closely related to user responsiveness, which can be enhanced by proposing behavioral adjustments that closely align with existing consumption patterns [5], [6].

A significant challenge in Demand Side Management (DSM) research is the limited availability of real-world load data. Most utilities internally develop their analytics using extensive data from customer smart meters. However, these datasets are often complex to access due to proprietary restrictions and privacy concerns. This lack of publicly available data hampers research progress. In this context, load modeling offers a privacy-preserving alternative by generating synthetic yet realistic power consumption profiles. By mitigating data scarcity, load simulations can accelerate advancements in DSM strategies and energy system planning [7].

Residential load modeling plays a crucial role in helping prosumers optimize the use of their photovoltaic systems and battery storage. By simulating realistic consumption patterns, it enables the scheduling of deferrable appliances to maximize self-consumption and minimize dependence on the grid [8]. The value of load modeling becomes even greater in energy communities, where the co-optimization of shared resources is essential for the coordinated management of multiple households [9], [10].

In recent years, Generative Adversarial Networks (GANs) have gained prominence as a powerful technique for synthetic data generation. Initially introduced for producing realistic images from training data [11], GANs have since been successfully applied to various data modalities, including images, videos, and audio signals [12]. This study investigates the efficacy of GANs as a paradigm shift in bottom-up residential load modeling, specifically for the simulation of daily user activities. While Markov chains have historically held sway, we posit that GANs offer a compelling alternative, achieving comparable fidelity in behavioral representation with a marked reduction in computational overhead. We present empirical evidence demonstrating this performance gain, advancing the argument for GANs as a more scalable and efficient methodology for large-scale residential load simulations.

The remainder of this paper is organized as follows: Section II reviews existing approaches to residential load modeling. Section III describes the datasets used to implement our simulation framework. Section IV details the calibration of GAN and Markov chain behavioral models and the deriva-

M. Castangia, S. Quer, L. Bottaccioli and E. Patti are with Politecnico di Torino, Turin, IT. E-mails: name.surname@polito.it.

tion of load profiles from daily activity patterns. Section V evaluates the accuracy of the simulated loads by comparing them with real-world power consumption at both the appliance and household levels. Finally, Section VI summarizes our key findings and outlines potential directions for future research.

II. Related works

Bottom-up simulation is a widely used load modeling approach that estimates electricity demand by simulating the daily activities of end users. This is typically achieved using a behavioral model calibrated with national time-use survey data from the country of interest [13]. In this framework, load modeling is performed by linking users' activities to the power consumption of household appliances. For example, the operation of a washing machine is triggered by the activity *"washing clothes"*. The aggregated household load curve is then obtained by summing the contributions of individual appliances [7].

The behavioral models that generate daily user activities in bottom-up simulations may vary across implementations. Foteinaki et al. [14] propose two probabilistic models to generate the daily activities of consumers based on the Danish time-use survey. Fischer et al. [15] introduce a stochastic model that generates household power profiles in Germany under various future scenarios. Their approach derives probability density functions from a comprehensive European time use survey, incorporating socioeconomic characteristics of participating families. Thorve et al. [16] implement an alternative behavioral model by training a random forest regressor on data from the American time-use survey. Given a set of socioeconomic factors as input, the random forest is trained to predict the time spent by the end user in different locations during the day. The simulated daily activities are sampled from the tree leaves based on the occupant's given socioeconomic factors. In our previous work [17], we introduced a multi-scale residential load model that employs a Non-Homogeneous Semi-Markov chain fitted to the Italian time use survey to simulate user behavior. Similarly, Osman et al. [18] calibrate a semi-Markov model using the Canadian time-use survey to model consumer behavior. Chen et al. [19] present a hybrid behavioral model combining Markov chains for activity scheduling with another probabilistic model for activity duration estimation. Their model is calibrated using the American time-use survey.

In contrast to bottom-up methods, top-down approaches simulate total household load directly by fitting generative models to accurate power consumption data. Generative Adversarial Networks (GANs) have been widely adopted for this purpose. Zhang et al. [20] propose a conditional GAN to generate the daily profiles of residential end users with a sampling frequency of 15 minutes. The GAN has been conditioned with the day of the week and the month to capture variations due to weather conditions and users' occupations. Gu et al. [21] implement a GAN-based approach to simulate weekly power curves for Irish households. They cluster weekly load profiles into ten representative consumer groups and then condition the GAN to generate power profiles based on the assigned group label. Hu et al. [22] propose to improve GANs by simultaneously generating groups of weekly loads with spatial and temporal correlations. Their method demonstrates the capability of generating individual loads that better resemble the distribution of real power profiles compared to the case of weekly curves generated separately. Liang and Wang [23] introduce a GAN architecture incorporating recurrent neural networks to capture temporal dependencies in load patterns better, demonstrating enhanced performance and stability compared to traditional GANs.

Modeling the daily activities of consumers is the most time-consuming component of bottom-up simulations, often posing a bottleneck to their scalability in large-scale scenarios. To address this limitation, we introduce a novel behavioral model based on Generative Adversarial Networks (GANs) that significantly accelerates execution time without compromising accuracy. In comparison to previous bottom-up approaches, the main contribution of our work lies in providing a more efficient and scalable behavioral model based on the use of neural networks. Unlike previous models based on Markov chains, which generate activities sequentially, our model simultaneously predicts entire daily schedules for thousands of users by leveraging the capabilities of parallel processing units. As reported in Table I, our methodology preserves the granularity of bottom-up approaches while exploiting the scalability of neural networks demonstrated in previous top-down methods. In this way, we enable large-scale simulations with thousands of agents, resulting in faster execution and more statistically robust outcomes.

TABLE I: Comparison of our load modeling framework with previously proposed approaches.

Reference	Approach	Model	Scalability	Granularity
[14]	Bottom-up	Probabilistic	Medium	High
[15]	Bottom-up	Probabilistic	Medium	High
[16]	Bottom-up	Probabilistic	Medium	High
[17]	Bottom-up	Markov Chains	Low	High
[18]	Bottom-up	Markov Chains	Low	High
[19]	Bottom-up	Markov Chains	Low	High
[20]	Top-down	GAN	High	Low
[21]	Top-down	GAN	High	Low
[22]	Top-down	GAN	High	Low
[23]	Top-down	GAN	High	Low
Our	Bottom-up	GAN	High	High

III. Dataset

In this section, we describe the data required to implement the simulator adopted in our methodology. National surveys and statistical datasets provide region-specific behavioral patterns and trends, though these are often constrained to particular historical periods (e.g., during increased remote work). However, because such data are widely available across many countries and are regularly updated to reflect evolving societal behaviors, the proposed framework can be readily adapted to simulate various geographic contexts or revised to incorporate recent behavioral shifts.

A. Time use survey (TUS)

The Italian Time Use Survey (TUS) was published in 2013 by the National Institute of Statistics (ISTAT) [24]. During this survey, the members of approximately 20,000 Italian families were asked to fill out a report of their daily activities at intervals of 10 minutes. The survey involved more than 41,000 individuals from different socioeconomic groups, work occupations, and family compositions. In this work, we used these data to train a behavioral model that simulates the activities of fictional end users during the day. Following the selection criteria of our previous work [17], we limited the user's activities to only those carried out inside the home and involving the use of electrical devices. In more detail, starting from an initial set of 146 different activities, we selected the following 11 activities: *studying*, *watching TV*, *using PC*, *reading*, *cooking*, *eating*, *dish washing*, *cleaning*, *ironing*, *being in bathroom* and *washing clothes*. Notice that national TUS data usually follows a standardized format to report information on the daily activities of end users [25]. Therefore, the proposed methodology can be easily adapted to data from different countries without losing generality.

B. Use of energy

In 2013, ISTAT collected a survey involving more than 20,000 families to obtain information on their energy consumption habits [26]. This survey collects statistics on the presence of certain household appliances and their relationship with the number of family members. We used this information to determine the probability of having a specific device conditional on the family size of the simulated household. To this aim, Table II shows the probability of owning different appliances, given the number of individuals in the household. Generally, we can see that the likelihood of having an appliance increases with the family size. The washing machine and the oven are present in almost all households; on the contrary, the dishwasher is only owned by about 50% of the families. Furthermore, the dryer looks pretty uncommon among Italian families, being used by less than 10% of the families, probably due to the temperate climate, which does not require its use most of the time.

TABLE II: Probability of appliance ownership as a function of the number of family members.

Number of family members	Dishwasher	Oven	Washing machine	Dryer
1	0.25	0.84	0.94	0.01
2	0.41	0.93	0.98	0.02
3	0.50	0.95	0.99	0.04
4	0.55	0.97	0.99	0.07
5	0.56	0.97	0.99	0.07

The survey on energy use also provides valuable information on the weekly activations of common deferrable appliances such as dishwashers, washing machines, and ovens [26]. Table III reports the number of weekly operations for each family size up to five members. As expected, the number of operations generally increases with family size. This information is used during load modeling to correctly associate the appliance activation with the specific task that triggers it. For example, the task *"dishwashing"* does not always correspond to turning on the dishwasher because the user may load the dishwasher multiple times before using it. The same reasoning can be applied to the activity *"washing clothes"*. Likewise, the activity *"cooking"* very rarely corresponds to the activation of the oven. Therefore, we used these parameters to determine these devices' activation probability whenever the user performs a triggering activity. In particular, the probabilities of each device are obtained by dividing the number of weekly activations by seven. This probability determines the likelihood of using the specific device each day of the week, provided the corresponding activity has been performed.

TABLE III: Number of weekly operations as a function of the number of family members (up to five).

Number of family members	Dishwasher	Oven	Washing Machine
1	2.14	1.10	1.64
2	3.11	1.38	2.69
3	4.11	1.55	4.02
4	4.80	1.65	5.30
5	5.32	1.83	6.19

C. Census dataset

The data from the 15$^{\text{th}}$ Italian census provides information on the distribution of family sizes in the Italian population [27]. This information has been used to generate realistic simulations with the correct distribution of family compositions. Table IV shows the distribution of Italian households based on the number of family members reported in the census dataset.

TABLE IV: Distribution of family compositions in the Italian population.

Number of family members	Probability
1	35%
2	30%
3	18%
4	8%
5	1%

D. Appliance power signatures

To reconstruct the daily power consumption of the simulated households, we leveraged the sub-metered power signatures from available public datasets. In particular, we selected a set of power signatures from the UK-DALE dataset [28]. UK-DALE provides the sub-metered power signatures of multiple electrical devices from five households monitored in the UK between 2012 and 2017. The power measurements are collected with a sampling rate of 6 seconds. We extracted the power profiles of six home appliances: Fridge, dishwasher, washing machine, dryer, oven, and iron. Figure 1 reports the power signatures of these devices. During load modeling, when the user carries out an activity associated with one of these appliances, we insert the corresponding power footprint into the house's total load in the activity's specific time slot. More straightforward power profiles, such as TV, PC, lights,

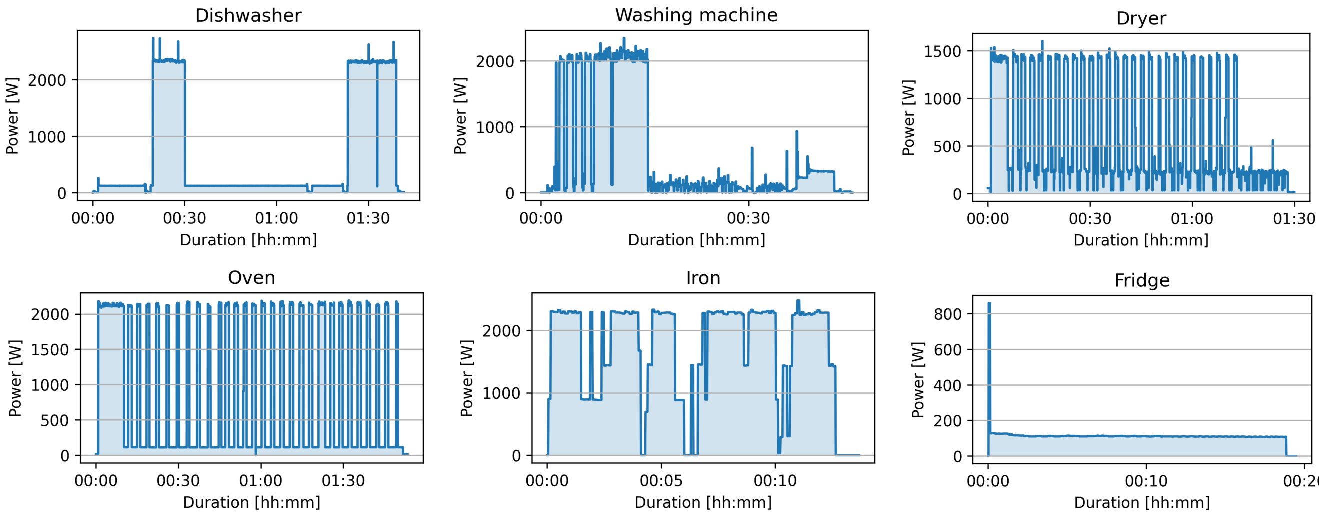


Fig. 1: Appliance power signatures from UK-DALE

and devices on standby, have been approximated using a constant power level. In more detail, based on the information collected in [29], we used a fixed consumption of 125W for each television set and 276W for each computer. For the lighting system, we considered an overall consumption of 150W for the whole house. Finally, standby power consumption is randomly drawn for each home from a uniform distribution between 10W and 50W.

E. Solar irradiance dataset

Leveraging the solar irradiance dataset, this study reconstructed seasonal patterns of artificial lighting consumption, illuminating the direct correlation between environmental conditions and electrical demand. The dataset provides the Global Horizontal Irradiance (GHI) measured in Turin (Italy) during 2015. The GHI is expressed in W/m^2 and represents the total energy received from the sun per square meter. Even though we used a proprietary dataset for GHI, the same data can be easily retrieved from other data sources providing the estimated historical GHI values for most of the world [30]. Figure 2 shows the typical cycles of solar irradiance during a clear sky day (orange) and a cloudy day (blue) in Turin. Bear in mind that in the event of overcast weather conditions, solar irradiation can reach lower values, potentially triggering the use of the lighting system even in the central hours of the day.

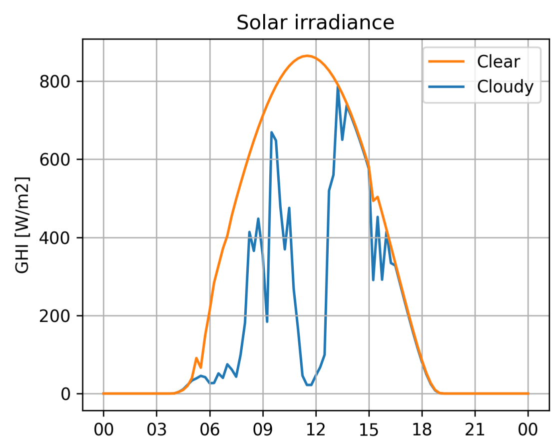


Fig. 2: Solar illumination trends in Turin during a clear sky and a cloudy day.

IV. Methodology

To demonstrate the effectiveness of the proposed behavioral model, we opted for a straightforward pipeline following the typical implementation of bottom-up load simulations. In particular, our procedure comprises two main steps as depicted in Figure 3. During the first step, i.e., the *activity generation*, we apply our behavioral model to generate the daily activities of the different households composing our population. During this stage, the activities of the various occupants are simulated separately and then merged to simulate the entire household.

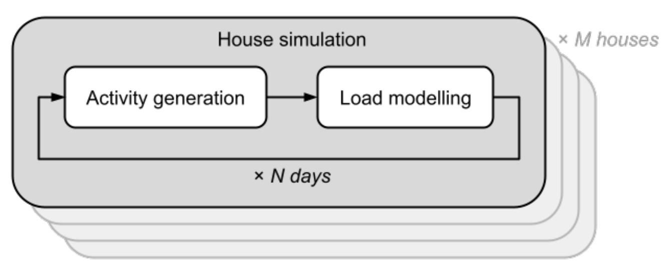


Fig. 3: The two stages of our residential load simulator.

In the second phase, i.e., the *load modeling*, we reconstruct the daily power consumption of the house by associating the activation of each typical electrical device with the activities generated in the previous phase.

In the remainder of this section, we first present the new GAN-based behavioral model for simulating daily activities. Then, we present an alternative behavioral model based on

Markov chains that we used for benchmarking our solution. Finally, we describe the load modeling stage that associates the user's daily activities with the appliance power signatures composing the total load of the house.

A. Generative adversarial networks for daily activities

Ian Goodfellow et al. [11] introduce GANs as a new framework for training neural networks via an adversarial approach. The aim was to capture and generate new samples from the underlying data distribution. In this novel framework, they alternately train two models. A generator model G is prepared to produce new realistic samples from some random noise z, and a discriminator model D is trained to distinguish actual training samples from those produced by G. Generally, G and D are implemented using neural networks as they are universal function approximators. The loss function of this two-player game between G and D is defined by the following expression:

$$\min_G \max_D \mathbb{E}_{x\sim P_r}[\log D(x)] + \mathbb{E}_{\tilde{x}\sim P_g}[\log(1 - D(\tilde{x})], \tag{1}$$

where x is a realization of the real data distribution P_r and $\tilde{x} = G(z)$ is a realization of the generator distribution P_g conditioned by the random noise z. In their standard formulation, GANs are very effective in generating realistic and novel samples but are also challenging to train. The training process is precarious, and converging requires a constant synchronization between G and D. In particular, if the discriminator D is not trained correctly, the generator G tends to "cheat" by consistently producing the same sample for different values of z (i.e., mode collapse). This problem makes it difficult to test different neural architectures because a different parametrization of the training process would likely be needed. To solve the problem of training instability, Arjovsky et al. [31] use an improved cost function using the Wasserstein distance between the actual data P_r and the generated one P_θ. This new framework, named Wasserstein GAN (WGAN), demonstrates significantly better stability to the original GAN formulation and also provides a meaningful metric to evaluate the convergence of the training process. However, WGANs still suffer from optimization problems due to the weight clipping applied to the discriminator D to enforce the Lipschitz constraint. To avoid the use of weight clipping, Gulrajani et al. [32] proposed adding a gradient penalty term to the value function of WGAN as an alternative way to enforce the Lipschitz constraint on the discriminator. The loss function of the WGAN with gradient penalty (WGAN-GP) is the following:

$$E_{\tilde{x}\sim P_g}[D(\tilde{x})] - E_{x\sim P_r}[D(x)] + \lambda E_{\tilde{x}\sim P_{\tilde{x}}}[(||\nabla_{\tilde{x}} D(\tilde{x})||_2 - 1)^2]. \tag{2}$$

Alternatively, the generator output can also be conditioned with multiple class labels, as demonstrated by Mirza et al. [33]. This capability benefits us for conditioning the generator with additional information and parameterizing our simulation. Conditional GANs are trained in the same way as their original implementation. The only difference is that the generator receives the class labels in addition to the noise prior z, and the discriminator receives the same class labels together with the example x to be processed.

The proposed behavioral model uses the WGAN-GP to generate the daily activities of simulated occupants. We trained the WGAN using the Italian TUS results described in Section III. First, we pre-processed the daily activities of TUS by converting them into a more suitable format that could be fed into the neural network. In more detail, we collected each individual's reported daily activities. Then, for each day of the survey, we generated a matrix X with 11 rows and 144 columns, where each row represents a potential activity performed by the user, and each column corresponds to a 10-minute time slot of the day. Each element $X_{i,j}$ of the matrix is initialized to zero, and it is set to one if the user performs the activity i at the time slot j. Finally, we associated a binary label c to each matrix X, indicating whether the sample was collected during a weekday or the weekend. This variable has been introduced to demonstrate the possibility of conditioning the generator's output on external factors affecting the simulation. At the end of pre-processing, we obtained 41227 training samples of pairs (X, c) corresponding to each person's reported daily activities in the Italian TUS.

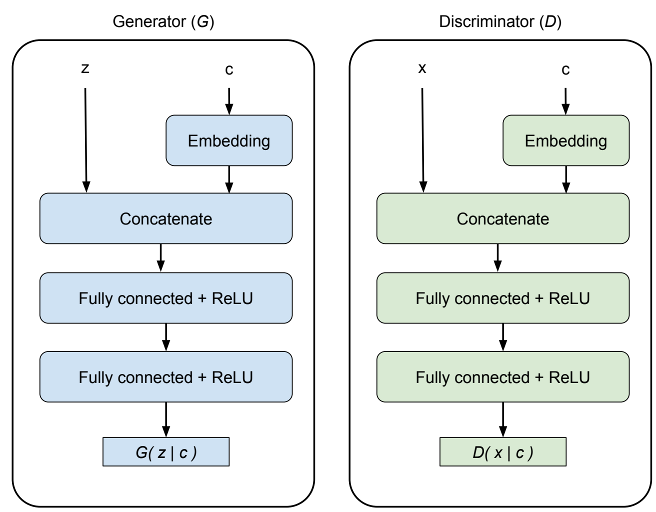


Fig. 4: The architecture of our conditional WGAN-GP.

Figure 4 describes the architecture of the conditional WGAN-GP we implemented in this work. The generator has two inputs: The noise prior z and the weekend indicator c. The noise vector z comprises 32 random numbers extracted from a normal distribution with a zero mean and unit variance. First, we pass the conditioning variable c through an embedding layer, which maps the label to a new hidden representation of 32 units. Then, we concatenate the prior noise by embedding the conditioning variable to obtain a new vector of size 64. This vector is finally passed through two additional fully connected layers with 128 units each and a rectified linear activation function. The last fully connected layer produces an output vector with shape 11×144, which contains the generated daily activities of the occupant, conditioned by the weekend occurrence. The architecture of the discriminator mirrors almost exactly the one of the generator, except that the input x is a vector with shape 11×144, and the output is a single unit with no activation function.

The hyperparameters of the WGAN-GP training phase are reported in Table V. Following the original implementation of the gradient penalty, we used a λ of 10 in the loss function. The generator and discriminator have been trained with the Adam optimizer with a learning rate 0.001 for 200 epochs, using a batch size of 64 samples. As suggested in the original WGAN paper, for each generator iteration, we trained the discriminator five times [31]. The number of epochs has been set by looking at the quality of the distributions of the generated daily activities at the different stages of the training process.

TABLE V: Training hyperparameters.

Hyperparameter	Value
λ	10
Optimization algorithm	Adam
Learning rate	0.001
Batch size	64
Number of epochs	200

B. Markov chain for daily activities

To benchmark the accuracy of the proposed generative neural network, we implemented a second behavioral model based on nonhomogeneous Markov chains, following the implementations of previous works adopting a bottom-up simulation approach [17], [19], [18]. Unlike the base definition, nonhomogeneous Markov chains use multiple transition matrices, one for each time step of the stochastic process. This property provides more flexibility in simulated daily activity distribution. Nonhomogeneous Markov chains can be fully described by two sets of parameters: a time-dependent transition matrix $P(n) \in \Re^{(12,12)}$ to determine the next activity given the current state at time n, and a vector of initial probabilities $\pi(0) \in \Re^{12}$ to set the initial state of the simulation. Note that we inserted an additional empty state to account for time steps where there is no activity from the user, who can be away or sleeping. Given the daily activities reported in the Italian TUS, we can easily estimate the parameters of the Markov chain. Each element $p_{i,j}(n)$ of the transition matrix $P(n)$ is computed with the following expression:

$$p_{i,j}(n) = \frac{a_{i,j}(n)}{a_i(n)}, \tag{3}$$

where $a_{i,j}(n)$ represents the number of co-occurrences of activity i followed by activity j, and $a_i(n)$ is the total number of occurrences of activity i at the time step n. The initial state's probabilities $\pi(0)$ are computed by finding the likelihood of performing each activity at the first time step of the day (i.e., midnight). To condition the simulated activities based on the day of the week, we estimated the parameters of two different Markov chains for the activities performed on weekdays and weekends. Once the Markov chain parameters have been fully determined, we can simulate the daily activities of an occupant using algorithm 1.

Algorithm 1 Markov chain simulation

Sample the initial state X_0 from $\pi(0)$
$n = 0$
while $n < 144$ **do**
    Sample the next state X_{n+1} from $P(n)$ given state X_n
    $n = n + 1$
return $\{X_0, X_1, ..., X_{143}\}$

C. Load modeling

Load modeling consists of distributing the appliance power signatures into the different time slots of the day according to the users' simulated daily activities. As reported in Table VI, specific daily activities directly trigger the operations of typical shiftable appliances (i.e., dishwasher, washing machine, dryer, oven, TV, and PC). Other devices instead operate independently during the day (i.e., fridge and standby), and their operations are not linked to any user's action. Finally, any user's activity at night, or without adequate sunlight during the day, can trigger the lighting system.

TABLE VI: Relation between daily activities and appliances.

Activity	Appliance
washing dishes	dishwasher
washing clothes	washing machine, dryer
cooking	oven
ironing	iron
using PC	PC
watching TV	TV

The following describes in more detail the rules we used to trigger the operations of the simulated devices.

1) Dishwasher: The activity *"dishwashing"* triggers the dishwasher. Since the activation of the dishwasher does not always correspond to this action, we enforced a limit of one activation per day. In addition, to limit the number of weekly operations, we extract a random number from a uniform distribution $U(0, 1)$ and only activate this appliance if this number is lower than the expected number of weekly operations divided by seven.

2) Washing machine: The action *"washing clothes"* activates the machine. As for the dishwasher, the activation of the washing machine does not have a one-to-one relation with this activity. We enforced a limit of one operation per day. Similarly to the dishwasher, we also pick a random number from a uniform distribution $U(0, 1)$ and activate this appliance only if this number is lower than the expected weekly operations divided by seven.

3) Dryer: The dryer is only activated after a previous washing machine run. In more detail, the dryer's power signature is placed two hours after the washing machine's activation to wait for its completion.

4) Oven: The activity *"cooking"* triggers the oven. We limit the number of operations to one per day, given that there can be multiple cooking sessions within the same day. To respect the expected number of weekly usages, we pick a random number from $U(0, 1)$ and compare it with the expected weekly operations divided by seven.

5) Iron: The iron is activated by the action *"ironing"* without any ambiguity on its usage. The duration of the iron power signature is truncated to the number of time slots of the ironing session.

6) PC: The PC is considered active in all time slots where the activity *"using PC"* is true. Its duration is equal to the time slot, i.e., 10 minutes.

7) TV: The TV follows the same rules as the PC and is triggered by the activity *"watching TV"*.

8) Fridge: The fridge operates throughout the day and is distributed according to its periodicity, which is the elapsed time between two consecutive activations. The starting time of the first operation is extracted randomly from a uniform distribution between zero and the fridge periodicity.

9) Lights: Lights are activated only when the global horizontal irradiance (GHI) is lower than 100 W/m^2. However, low GHI values can also be registered during the day in very cloudy weather conditions, which explains the use of lights.

10) Standby: Standby power consumption is present during the whole day and is added to the daily load of the house.

V. Results

In this section, we evaluate the activity generation and load modeling stages, comparing the performance of GANs and Markov chains in these tasks. First, we validate our approach by comparing the distributions of the generated activities with those reported in the Italian TUS. Then, we compare the simulated loads with those from the real-world measurements, assessing the consistency between the reconstructed and real power profiles. Finally, we evaluate the performance of the behavioral model in terms of its execution times, showing its benefits concerning previous methods.

A. Validation of activity generation

To verify the authenticity of the daily activities generated by our neural network, we compared the distributions of the simulated activities with those reported in the Italian TUS. In more detail, we simulated as many individuals as those interviewed during the survey, generating 14245 samples for the weekdays and 26982 samples for the weekends. We also ran the same simulation with the Markov chains to verify the correctness of the benchmark model. To compare the simulated and actual distributions of daily activities, we used the Kullback–Leibler divergence, which also provides a quantitative comparison between the proposed behavioral model and the benchmark. The Kullback–Leibler divergence for discrete distributions is defined as follows:

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}, \tag{4}$$

where $P(x)$ and $Q(x)$ are the probability mass functions of the reference and simulated distributions, respectively. Figure 5 shows the probability distributions of the 11 daily activities simulated by our behavioral models. We can see that both GAN and the Markov chain closely follow the distributions of the time-use survey, demonstrating the generated samples' authenticity. Regarding the Kullback–Leibler divergence reported in Table VII, the two models are equally accurate for weekday and weekend simulations. Therefore, we can say that GANs can be used as an alternative to Markov chains, providing equally accurate simulation results.

TABLE VII: Kullback–Leibler divergence computed on the daily activities distributions generated by GAN and Markov chain (MC) for weekdays and weekends.

Activity	Weekday		Weekend	
	MC	GAN	MC	GAN
Studying	0.013	0.010	0.013	0.011
Watching TV	0.013	0.004	0.012	0.003
Using PC	0.009	0.019	0.011	0.009
Reading	0.014	0.028	0.014	0.015
Cooking	0.018	0.008	0.017	0.008
Eating	0.015	0.004	0.016	0.003
Dish washing	0.024	0.015	0.024	0.009
Cleaning	0.020	0.012	0.017	0.005
Ironing	0.016	0.021	0.020	0.024
Being in bathroom	0.010	0.009	0.009	0.007
Washing clothes	0.032	0.053	0.016	0.024

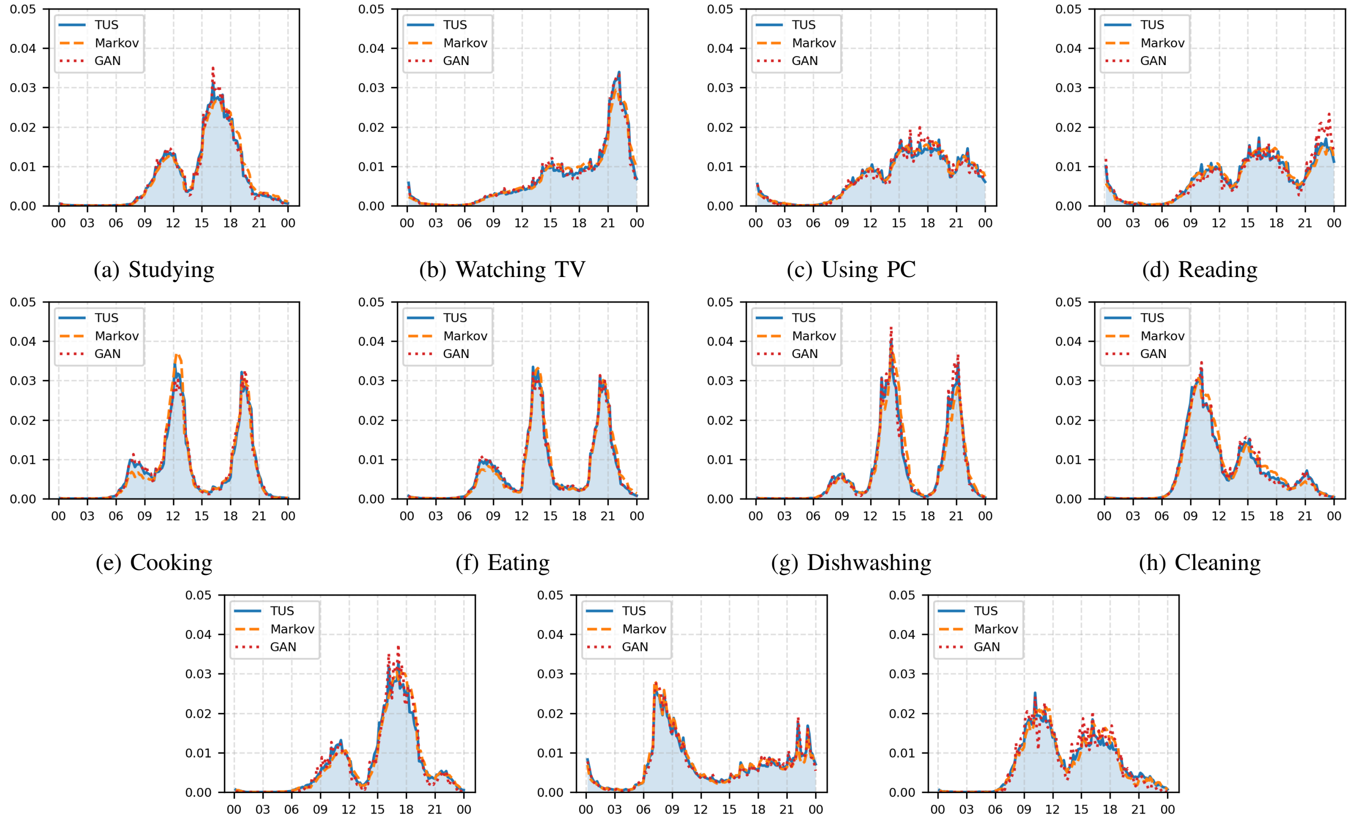
B. Validation of load modeling

The synthetic curves generated for the validation of load modeling were obtained by simulating 1000 families over two weeks. For every electrical device described in Section IV-C, we reconstructed its daily power consumption for each house at a six-second resolution. We first resample the devices' power measurements to get the synthetic curves by computing their average hourly power consumption. Then, we further average each appliance's hourly power consumption across all simulated households, separately for weekends and weekdays. At the end, we obtain a sequence of 24 values for each appliance and the whole house, representing the average hourly power consumption of the simulated appliances. These curves are finally compared with their real counterparts to assess the authenticity of the reconstructed power profiles. We used the mean bias error and Pearson correlation coefficient to measure the difference between the generated and accurate loads. The mean bias error measures the average difference between the two curves and indicates whether we overestimate or underestimate the actual power consumption on average. On the other hand, Pearson correlation measures the temporal consistency of load shapes without considering differences in their scales. The bias and the Pearson correlation are defined as follows:

$$bias = \frac{\sum_{i=1}^{n} (x_i - y_i)}{n}, \tag{5}$$

$$correlation = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 (y_i - \overline{y})^2}}, \tag{6}$$

where x is the vector of real power measurements and y is the vector of simulated loads. Both vectors have a length n



(a) Studying (b) Watching TV (c) Using PC (d) Reading (e) Cooking (f) Eating (g) Dishwashing (h) Cleaning (i) Ironing (j) Being in bathroom (k) Washing clothes

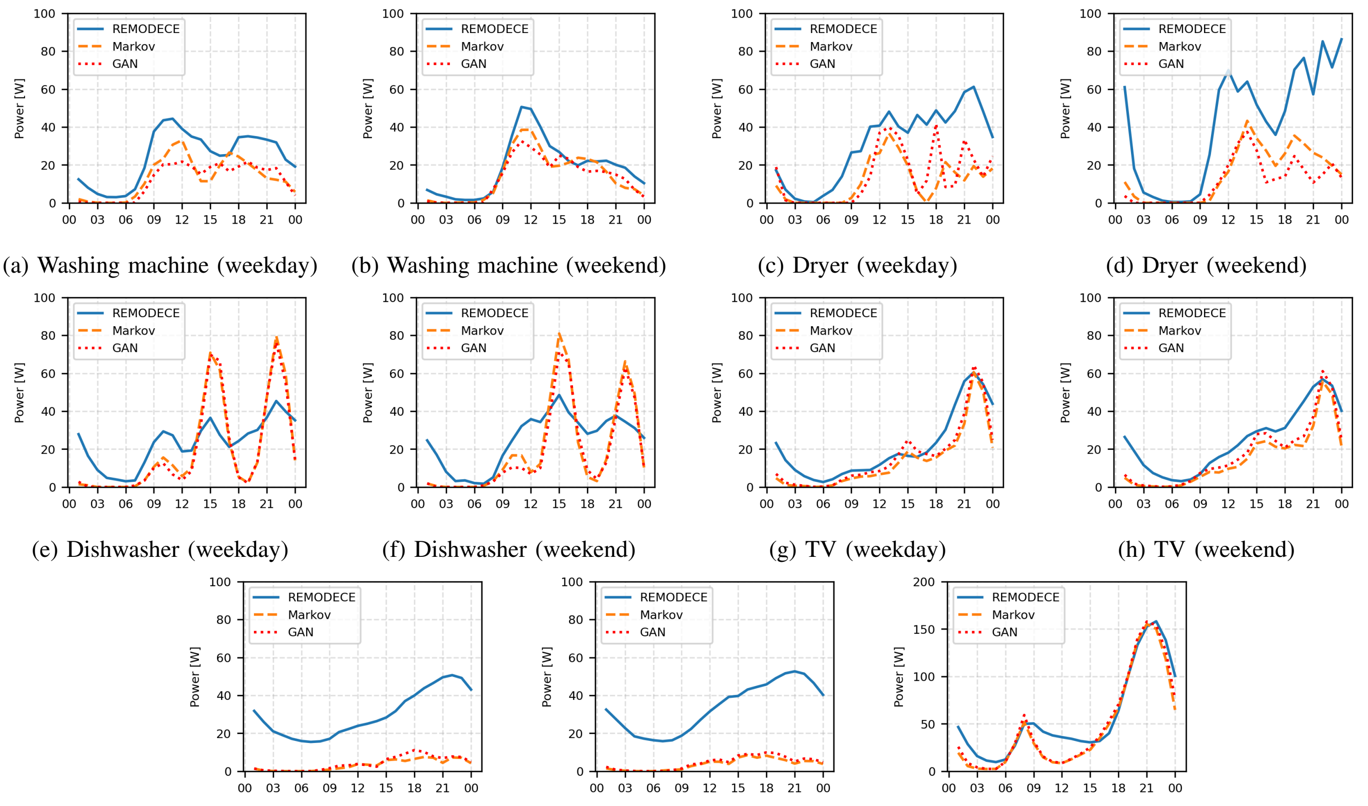
Fig. 5: Daily activity distributions generated by GAN and Markov chain compared to the Italian TUS.

equal to 24, where each element contains the average power consumption of the meter during the previous hour.

The simulated appliance loads have been compared with the average power measurements of the devices monitored in the REMODECE project [29]. In more detail, the REMODECE project collected the energy consumption of thousands of electrical devices by tracking more than 1000 households in 12 different European countries for two weeks. For each house owning the monitored device, REMODECE reports the average hourly power consumption of the appliance computed across the two-week monitoring period. To conform with the methodology of REMODECE, we computed the average hourly consumption of each device for a population size of 1000 households and a simulation length of two weeks.

REMODECE validated the simulated load profiles of five different electrical devices, including the washing machine, dryer, dishwasher, TV, and PC. In addition, we also evaluated the simulation of the house lighting system. Table VIII shows the mean bias error and the Pearson correlation of each appliance assessed separately for the weekdays and the weekends. Furthermore, we differentiated loads derived from the synthetic activities of GANs and those obtained from the Markov chain model. As confirmed also in Figure 6, all simulated loads approximate reasonably well the shapes of the actual appliance power profiles. As expected, the behavioral models produced very similar load modeling results, proving the interchangeability of generative adversarial networks and Markov chains in bottom-up simulations. Table VIII also shows a general negative bias error across all simulated devices, indicating an overall underestimation of the average hourly power consumption. Given that both behavioral models are correct for the Italian TUS (see Figure 6), we think that negative biases are primarily due to differences between European countries in the frequency of use of some appliances. External factors such as climate conditions can influence the use of specific devices (e.g., dryers). However, based on the correlation coefficients reported in Table VIII, we can state that the proposed behavioral model successfully provides coherent daily activities for scheduling the different appliance operations, which is fundamental for producing realistic bottom-up simulations.

We also evaluated the simulation results at the building level by using a study conducted by the Italian company *Research on Energy Systems s.p.a.* (RSE), reporting the typical daily loads of Italian families at the whole house level [34]. In particular, the RSE dataset contains the average aggregated loads of 1200 Italian families monitored for two years (2011-2012) by utility companies owning second-generation electricity meters. As reported in Table VIII, the total simulated loads show a high correlation with the real curves for weekdays and weekends, independent of the adopted behavioral model. Similarly to the appliance level results, the total simulated loads also have a significant negative bias concerning the actual curves, revealing a constant negative gap during the



(a) Washing machine (weekday) (b) Washing machine (weekend) (c) Dryer (weekday) (d) Dryer (weekend)

(e) Dishwasher (weekday) (f) Dishwasher (weekend) (g) TV (weekday) (h) TV (weekend)

(i) PC (weekday) (j) PC (weekend) (k) Lights

Fig. 6: Comparison between simulated appliance loads and REMODECE.

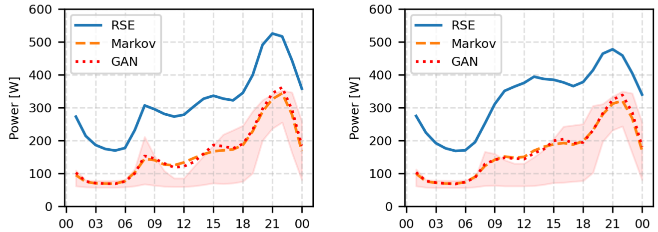
day of roughly 150 W, as shown in Figure 7. Since the gap is also present at night, we can presume that the causes should be searched in devices that generally operate during the day. For example, we did not simulate essential devices such as electric boilers and HVAC systems, which typically account for a good percentage of the total load where they are present. In addition, we only considered houses with a single fridge, whereas some families may easily own more than one refrigerator unit. Despite these simplifications, the high correlation indicates that the proposed behavioral model can correctly simulate the overall dynamics of power consumption, placing peaks and valleys at the correct times of the day. Furthermore, the interquartile range of the simulated households (red areas in Figure 7) demonstrates that our simulations can cover various behaviors, highlighting the generated loads' originality and diversity.

TABLE VIII: Evaluation of bias and correlation obtained by GAN and Markov chain (MC) in the load modeling task.

Loads	Bias [W]		Correlation	
	MC	GAN	MC	GAN
washing machine (weekdays)	-10.97	-12.38	0.88	0.93
washing machine (weekends)	-4.71	-5.79	0.96	0.95
dryer (weekdays)	-19.46	-16.60	0.66	0.68
dryer (weekends	-26.06	-30.15	0.77	0.74
dishwasher (weekdays)	-2.29	-4.01	0.74	0.72
dishwasher (weekends)	-4.87	-6.58	0.72	0.70
TV (weekdays)	-7.44	-5.15	0.91	0.92
TV (weekends)	-6.69	-7.36	0.91	0.92
PC (weekdays)	-26.61	-25.75	0.85	0.82
PC (weekends)	-29.95	-29.18	0.83	0.87
Lights	-11.15	-8.36	0.96	0.97
Total (weekdays)	-149.43	-143.43	0.98	0.98
Total (weekends)	-164.17	-163.85	0.93	0.91

C. Execution times

Table IX compares the execution times of the behavioral models based on GANs and Markov chains with increasing population sizes. In all cases, neural networks demonstrated to take significantly lower than Markov chains to simulate the daily activities of end users, demonstrating better scalability to more significant simulation scenarios. Contrary to the Markov chain generation procedure, neural networks can concurrently generate the daily activities of thousands of individuals in a single call, thanks to GPU parallelization of matrix operations. For this purpose, Table IX also shows the GPU memory use of the generator neural network when run on an NVIDIA QUADRO P2200 with different batch sizes (i.e., number of occupants). The results demonstrate that generating an entire day of activities for 100,000 users requires less than one second on a customer-grade GPU. GPU memory does not represent a bottleneck for running more extensive simulations. In fact, we need to split the simulated individuals in multiple batches for large population sizes and run the network numerous times. Overall, the results highlight the importance of GANs in bottom-up load modeling, demonstrating that



(a) weekday (b) weekend

Fig. 7: Comparison between total simulated loads and RSE.

their adoption can significantly speed up simulations of more significant scenarios.

TABLE IX: Execution times of GAN and Markov chain (MC) for different numbers of simulated occupants.

Population size	Execution time MC	Execution time GAN	GPU memory
10^2	223 ms	2 ms	23 MB
10^3	2.31 s	12 ms	57 MB
10^4	23 s	130 ms	430 MB
10^5	3 min	991 ms	3.26 GB

VI. Conclusion and Future Works

This paper presents a novel residential load simulation framework, centered on conditional Generative Adversarial Networks, designed to synthesize realistic daily activity patterns of end-users. The generator's architecture permits conditioning on specific days of the week, enabling the exploration of diverse behavioral scenarios reflecting weekly lifestyle variations. To rigorously validate the methodology, we conducted comparative analyses at both the individual appliance and aggregated household levels, juxtaposing simulated load profiles with real-world power consumption data. Our findings reveal a strong correlation between the generated and empirical load curves, providing compelling evidence that the model accurately captures daily activities' temporal dynamics and household appliances' operational timing. Crucially, the proposed GAN-based behavioral model significantly enhances computational efficiency compared to traditional Markov chain methodologies. This substantial reduction in execution time underscores the model's suitability for integration into future bottom-up simulation frameworks, particularly those requiring the analysis of large-scale residential energy datasets and complex scenario evaluations.

As future work, we plan to extend our simulations by integrating HVAC systems through the combination of our behavioral model with mathematical representations of thermal dynamics. Additionally, incorporating electric boilers and electric vehicles will further narrow the gap between simulated and real household loads. More advanced generative approaches such as diffusion models should be explored to further enhance the behavioral modeling of consumers [35], potentially capturing a broader range of realistic and diverse daily activity patterns.

References

[1] IEA, "Buildings," July 2023. [Online]. Available: https://www.iea.org/energy-system/buildings#tracking

[2] M. Gilleran, E. Bonnema, J. Woods, P. Mishra, I. Doebber, C. Hunter, M. Mitchell, and M. Mann, "Impact of electric vehicle charging on the power demand of retail buildings," *Advances in Applied Energy*, vol. 4, p. 100062, 2021.

[3] IEA, "The future of heat pumps," July 2022. [Online]. Available: https://www.iea.org/reports/the-future-of-heat-pumps

[4] B. Qiu, S. Mao, H. Xiao, and Z. Zhang, "Power-aware user association and renewable energy configuration for the optimized on-grid energy in hybrid-energy heterogeneous cellular networks," *IEEE Systems Journal*, 2023.

[5] J. Wang and S. Matsumoto, "An economic model of home appliance replacement: application to refrigerator replacement among japanese households," *Environmental Economics and Policy Studies*, pp. 1–20, 2021.

[6] P. K. Saha, N. Chakraborty, A. Mondal, and S. Mondal, "Intelligent real-time utilization of hybrid energy resources for cost optimization in smart microgrids," *IEEE Systems Journal*, 2024.

[7] X. Kang, J. An, and D. Yan, "A systematic review of building electricity use profile models," *Energy and Buildings*, vol. 281, p. 112753, 2023.

[8] C. Ammari, D. Belatrache, B. Touhami, and S. Makhloufi, "Sizing, optimization, control and energy management of hybrid renewable energy system—a review," *Energy and Built Environment*, vol. 3, no. 4, pp. 399–411, 2022.

[9] D. S. Schiera, L. Barbierato, A. Lanzini, R. Borchiellini, E. Pons, E. Bompard, E. Patti, E. Macii, and L. Bottaccioli, "A distributed multimodel platform to cosimulate multienergy systems in smart buildings," *IEEE Transactions on Industry Applications*, vol. 57, no. 5, pp. 4428–4440, 2021.

[10] L. Barbierato, D. S. Schiera, M. Orlando, A. Lanzini, E. Pons, L. Bottaccioli, and E. Patti, "Facilitating smart grids integration through a hybrid multi-model co-simulation framework," *IEEE Access*, 2024.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27. Curran Associates, Inc., 2014.

[12] C. Donahue, J. McAuley, and M. Puckette, "Adversarial audio synthesis," *arXiv preprint arXiv:1802.04208*, 2018.

[13] E. Proedrou, "A comprehensive review of residential electricity load profile models," *IEEE Access*, vol. 9, pp. 12 114–12 133, 2021.

[14] K. Foteinaki, R. Li, C. Rode, and R. K. Andersen, "Modelling household electricity load profiles based on danish time-use survey data," *Energy and Buildings*, vol. 202, p. 109355, 2019.

[15] D. Fischer, A. Surmann, and K. B. Lindberg, "Impact of emerging technologies on the electricity load profile of residential areas," *Energy and Buildings*, vol. 208, p. 109614, 2020.

[16] S. Thorve, Y. Y. Baek, S. Swarup, H. Mortveit, A. Marathe, A. Vullikanti, and M. Marathe, "High resolution synthetic residential energy use profiles for the united states," *Scientific Data*, vol. 10, no. 1, p. 76, 2023.

[17] L. Bottaccioli, S. Di Cataldo, A. Acquaviva, and E. Patti, "Realistic multi-scale modeling of household electricity behaviors," *IEEE Access*, vol. 7, pp. 2467–2489, 2018.

[18] M. Osman, M. Ouf, E. Azar, and B. Dong, "Stochastic bottom-up load profile generator for canadian households' electricity demand," *Building and Environment*, vol. 241, p. 110490, 2023.

[19] J. Chen, R. Adhikari, E. Wilson, J. Robertson, A. Fontanini, B. Polly, and O. Olawale, "Stochastic simulation of occupant-driven energy use in a bottom-up residential building stock model," *Applied Energy*, vol. 325, p. 119890, 2022.

[20] C. Zhang, S. R. Kuppannagari, R. Kannan, and V. K. Prasanna, "Generative adversarial network for synthetic time series data generation in smart grids," in *2018 IEEE international conference on communications, control, and computing technologies for smart grids (SmartGridComm)*. IEEE, 2018, pp. 1–6.

[21] Y. Gu, Q. Chen, K. Liu, L. Xie, and C. Kang, "Gan-based model for residential load generation considering typical consumption patterns," in *2019 IEEE power & energy society innovative smart grid technologies conference (ISGT)*. IEEE, 2019, pp. 1–5.

[22] Y. Hu, Y. Li, L. Song, H. P. Lee, P. Rehm, M. Makdad, E. Miller, and N. Lu, "Multiload-gan: A gan-based synthetic load group generation method considering spatial-temporal correlations," *IEEE Transactions on Smart Grid*, 2023.

[23] X. Liang and H. Wang, "Synthesis of realistic load data: Adversarial networks for learning and generating residential load patterns," in *Tackling Climate Change with Machine Learning 2022*. Neural Information Processing Systems (NIPS), 2022, pp. 1–8.

[24] ISTAT, "Multiscope on families: Use of time—microfates for public use." 2013. [Online]. Available: https://www.istat.it/it/archivio/202531

[25] Eurostat, "Harmonised european time use surveys (hetus) 2018 guidelines," 2019.

[26] ISTAT, "Survey on energy consumption of families: Microdates for public uses," 2013. [Online]. Available: https://www.istat.it/it/archivio/203344

[27] ——, "General census of population and housing at 2011—sample at 1 percent." 2013. [Online]. Available: https://www.istat.it/it/archivio/196129

[28] J. Kelly and W. Knottenbelt, "The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes," *Scientific data*, vol. 2, no. 1, pp. 1–14, 2015.

[29] A. Almeida *et al.*, "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.

[30] U. Pfeifroth, J. Drücke, S. Kothe, J. Trentmann, M. Schröder, and R. Hollmann, "Sarah-3–satellite-based climate data records of surface solar radiation," *Earth System Science Data*, vol. 16, no. 11, pp. 5243–5265, 2024.

[31] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International conference on machine learning*. PMLR, 2017, pp. 214–223.

[32] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," *Advances in neural information processing systems*, vol. 30, 2017.

[33] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[34] S. Maggiore, "Impatto su comportamenti e consumi delle famiglie di un sistema di prezzi biorari dell'energia elettrica," *RSE, Milano*, 2012.

[35] A. Kotelnikov, D. Baranchuk, I. Rubachev, and A. Babenko, "Tabddpm: Modelling tabular data with diffusion models," in *International Conference on Machine Learning*. PMLR, 2023, pp. 17 564–17 579.

Marco Castangia earned his Ph.D. in Computer Engineering from Politecnico di Torino in 2024. He is currently a research fellow in the Electronic Design Automation (EDA) group at the same institution. His research focuses on leveraging machine learning techniques to improve the efficiency of energy systems. During his Ph.D., he developed an innovative non-intrusive load monitoring framework in collaboration with an Italian start-up. His work also includes the study of advanced machine learning methods for short-term forecasting of renewable energy sources.

Benedetta Giorgi is currently working as a Data Scientist at Enel Global Services, contributing to AI-driven projects that enhance and support Enel's business operations. She holds a degree in Computer Engineering with a focus on Data Analytics from Politecnico di Torino, Italy, where she developed a strong foundation in data-driven problem solving and analytical techniques. Passionate about artificial intelligence and data, she is always eager to explore innovative ways to leverage AI for impactful solutions and optimizing business processes.

Stefano Quer received an M.S. in Electronic Engineering in 1991, and a Ph.D. in Computer Engineering in 1996. He has been a Visiting Faculty in the Department of Electronic Engineering and Computer Science at the University of California in Berkeley, an intern with the "Advanced Technology Group" at Synopsys Inc., and with the "Alpha Development Group" at Compaq Computer Corporation. He has been a Compaq Computer Corporation consultant. He is currently a professor with the Department of Control and Computer Engineering at Politecnico di Torino, Italy. His main research interests include systems and tools for CAD for VLSI, formal methods for hardware and software systems, and the development of sequential and concurrent algorithms that can achieve acceptable solutions with limited resources.

Lorenzo Bottaccioli is an Associate Professor at the Interuniversity Department of Regional and Urban Studies and Planning and at the Energy Center Lab of the Politecnico di Torino. His work emphasizes a multidisciplinary approach, integrating urban planning, energy systems, digital technologies, and policy frameworks to address complex challenges in urban energy management.

Edoardo Patti is currently an Associate Professor with the Department of Control and Computer Engineering, Politecnico di Torino, Torino, Italy. His research interests include ubiquitous computing, Internet of Things, software architectures with particular emphasis on infrastructure for ambient intelligence, machine learning for energy data analysis, and cosimulation techniques for multienergy systems integration in smart cities and energy communities.