Impact of Model Granularity on Residential Occupant-Dependent Electrical Load Modelling

Jérôme Le Dréau

Impact of Model Granularity on Residential Occupant-Dependent Electrical Load Modelling

Jérôme Le Dréau[1*]

[1] LaSIE UMR CNRS 7356, La Rochelle Université, France
[*jledreau@univ-lr.fr](mailto:jledreau@univ-lr.fr)
ORCID: 0000-0002-6804-4160

RESUME. Une modélisation précise des courbes de charge est indispensable pour modéliser l'interaction bâtiments-réseaux, qui peut être matérialisée à l'échelle du bâtiment par une production locale ou un signal prix dynamique. Ce type de modélisation repose généralement sur des modèles « bottom-up », qui nécessitent de nombreux paramètres d'entrée (par exemple, des paramètres socio-économiques, d'activités, les taux d'équipement et l'efficacité énergétique). L'objectif principal de cette étude est de déterminer le niveau de détail requis pour qu'un générateur de courbes de charge puisse capturer avec précision l'intensité d'utilisation et la variation temporelle d'équipements spécifiques. Différents modèles ont été développés, prenant en compte les variations dans la modélisation de la présence et/ou des activités des occupants, ainsi que dans la modélisation des courbes de charge unitaires. Un jeu de données de validation (Elecdom, 2023) est utilisé pour comparer les résultats des modèles. Différents indicateurs de performance sont ensuite définis pour analyser les résultats, en se concentrant sur des objectifs spécifiques d'interaction bâtiments-réseaux. Concernant l'utilisation du lave-linge, l'analyse de ces résultats révèle que les courbes de charge unitaires mesurés améliorent la diversité en termes d'intensité d'utilisation. La diversité temporelle est améliorée par les modèles plus avancés, mais reste cependant limitée par le nombre restreint de catégories d'occupants.

MOTS-CLÉS : charge électrique, modèle d'occupants, bottom-up, stochastique, niveau de détail.

ABSTRACT. Accurate modelling of the electrical load is required when modelling building-to-grid interaction in order to properly capture the dynamics and evaluate the match with local production or dynamic price signals. Such modelling usually relies on bottom-up models, which require many input parameters (e.g. socio-economic, activity-, appliance- and energy efficiency-related parameters). The main objective of this study is to determine the level of detail required for an electrical load model to accurately capture the intensity of use and the temporal variation of specific equipment. Different models are developed, accounting for variations in the modelling of occupant presence and/or activities and in the modelling of the cycle power profiles. A validation dataset (Elecdom, 2023) is used to compare the models' outputs. Different KPIs are then defined to compare the results, focusing on specific building-to-grid objectives. For the use of the washing machine, the analysis of these results reveals that measured cycle profiles improve the diversity in terms of intensity of use. The temporal diversity is improved with person-based models but still suffers from limitations due to the limited number of categories for occupants.

KEYWORDS: electric-load, occupant model, bottom-up, stochastic, level-of-detail

1. INTRODUCTION

Buildings can play an active role in the transition to a new energy system by producing energy locally (e.g. with a photovoltaic system) and providing flexibility services to the grid by modulating their energy usage. This new role for buildings requires modelling tools that can properly capture diversity in terms of both time and intensity of use. Such precision is necessary in order to capture dynamics and evaluate the match with local production or dynamic price signals. In this article, we will focus on the modelling of appliances that are occupant-dependant, such as washing machines or television. According to (Carlucci et al. 2020), there are five categories of model: deterministic, statistical, stochastic/probabilistic, data-driven and agent-type. Of these, generalised linear and stochastic models are the most widely used in building energy modelling, particularly those based on Markov chains (Richardson et al. 2010). This type of modelling requires many input parameters related to socio-economics, activities, appliances and energy efficiency. Although there is a wealth of literature on tool development, there is limited research on the comparison of modelling procedures and the ability of tools to generate coherent diversity. (Bottaccioli et al. 2019) developed a non-homogenous semi-Markov model that can generate statistically consistent populations of heterogeneous families. (Yamaguchi et al. 2019) proposed a cross-analysis of four existing methods for modelling household appliance use. They noted the limitations of existing tools in replicating intra- and inter-household variations, and emphasised the various parameters influencing load simulation. (Sanderson et al. 2020) compared three modelling techniques for wet appliances and demonstrated that models driven by the presence of occupants performed best in terms of modelling load distribution over time. (Barsanti et al. 2024) challenged the common modelling assumption that socio-demographic factors are sufficient to characterise energy demand heterogeneity and compared three modelling methods. Their alternative approach, which centred on patterns of appliance use, showed improvements in modelling diversity.

The state-of-the-art shows that there is considerable variation in appliance use modelling methodology, and that only a limited number of articles compare different techniques. While most stochastic models can accurately reproduce the average load of a group of houses, their ability to model usage diversity has not been extensively evaluated. There is no consensus on the most suitable method and the level of detail required for accounting for diversity. This article aims to evaluate the impact of modelling techniques and assumptions on the modelling of diversity (in terms of intensity and temporality) in stochastic bottom-up models. In Section 2, the different models developed are presented, together with the relevant key performance indicators. In Section 3, some results are analysed, first focusing on the influence of the cycle power profiles, and then on the modelling of occupants. The washing machine case study is used to test different modelling strategies.

2. METHODS

Different occupant models have been developed to test the influence of model granularity on residential occupant-dependent electrical load modelling. In this work, the models used are first-order non-homogenous Markov chains. This type of model has been preferred to regression-based or agent-based models as they are quite widely used for occupant modelling. The modelling work has been performed in Python, using an updated DEMOD library with parameters adapted to the French context and developing specific models (Le Dréau 2025). The modelling time-step is 10 minutes.

The overall procedure can be broken down into three steps as shown in Figure 1. This type of modelling relies on different databases, which are used to describe the activities of the population (time-use surveys TUS), the characteristics of households (census data) and the energy use of equipment (cycle profiles). The procedures under investigation in this work are highlighted in orange in Figure 1 and deal with the following two topics:

- the level of granularity of the TUS data and their integration into Markov-chain-type occupant models;
- the level of detail on the cycle power profile data.

This work uses the concept of level of detail (LoD) to characterise the differences between modelling approaches, with LoD 1 representing the simplest approach. This concept was introduced by (Malik et al. 2022) and (Wu et al. 2023) in the context of occupant modelling.

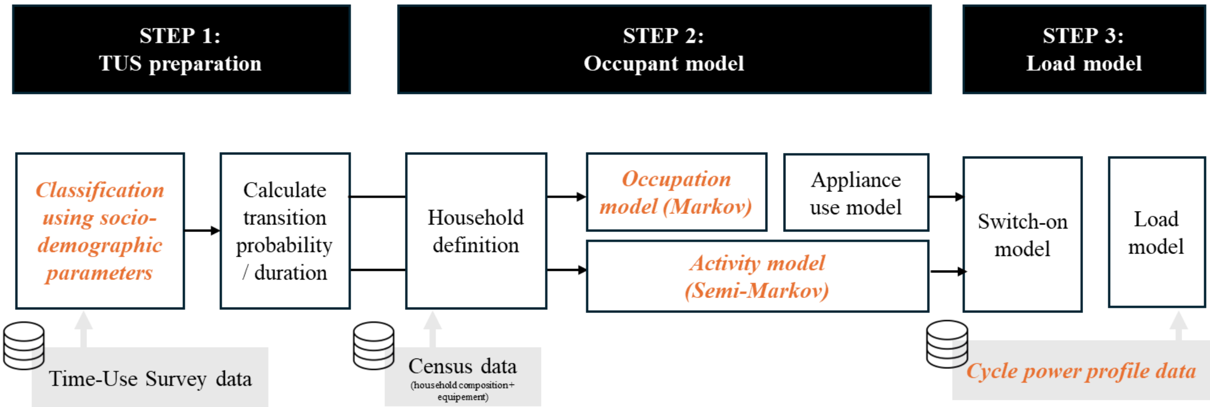

Figure 1 : Overview of the procedure for modelling appliance use
(in orange are the procedures under investigation in this article).

2.1. Step 1 and 2: Modelling occupants from time-use survey data

First, the Time-Use Survey (TUS) data (INSEE 2010) must be processed to calculate the transition probabilities and durations for the Markov chain models. Three classifications have been tested to evaluate the influence of the subgroup granularity: one with two typologies (adult/student), one with three typologies (man/woman/student), or one with 9 typologies (man/woman/student, single-person/multi-person household, worker/retired). For the more detailed typology, we assume that both family and job status influence much the activities performed by occupants. Weekends and weekdays data have also been split, leading to 25 different groups of TUS data. Then, the parameters of the Markov and semi-Markov chain models are calculated for each of these subgroups. The second type of model differs from the first in that it calculates the sojourn time for each state, thereby avoiding the frequent changes in activities that can be encountered with the classic Markov chain. The number of states is also defined differently, with either 4 states[1] for the Markov chain or 13 states[2] for the semi-Markov chain. The first model is based on the classic Richardson model (Richardson et al. 2010), whereas the second one is based on the model developed by (Bottaccioli et al. 2019). For the simplest model, the appliance usage model must also be defined based on the average activity-related usage profile of the related population. In a second step, the different types of households are defined based on census data related to typical household composition and characteristics of household equipment. Average national statistics are used (e.g. in a two-person household, the probability of the individuals being retired or active, and of the household being a couple or divorced, is determined based on INSEE census data (INSEE 2022)). Kids are assumed to have the same schedule as students. It should be noted that appliance ownership is assumed to be constant for all subgroups. In total, four different levels of detail have been defined for the categorisation of TUS and the type of occupant model:

- occ_LoD 1: an occupancy-based model, using a 4-state Markov chain model based on two subgroups (adult/kid)

[1] *'away from home and active', 'away from home and asleep', 'at home and active', 'at home and asleep'*
[2] *'away', 'cleaning', 'cooking', 'dishwashing', 'electronics', 'ironing', 'laundry', 'music', 'other', 'self_washing', 'sleeping', 'studying', 'watching_tv'*

- occ_LoD 2: an activity-based model, using a 13-state semi-Markov chain model based on:
    - occ_LoD 2-1: two subgroups (adult/kid)
    - occ_LoD 2-2: three subgroups (man/woman/kid)
    - occ_LoD 2-3: nine subgroups (man/woman/kid, single/couple, active/retired)

2.2. STEP 3: MODELS FOR POWER CYCLE LOADS

The last step consists of converting the activities to energy consumption, assuming a proportional relationship between the related activity and the switch-on of the equipment. In order to compare the diversity of the outputs, all simulations are set to have the same mean number of cycles for the pool of houses. This target is used to calibrate the probability of switch-on events in relation to the number of activity-related time steps. Ultimately, the power cycle load of the associated equipment can be incorporated into the load time series. A single type of power cycle load is assigned to the house. In total, four different levels of details have been defined for the power cycle loads model:

- *cycle_LoD 1*: constant power when switched-on
- *cycle_LoD 2*: simplified profile, with a peak of power at the beginning of the cycle
    - *cycle_LoD 2-1*: average energy class (A+)
    - *cycle_LoD 2-2*: randomly selected energy class (from A+++ up to B)
- *cycle_LoD 3*: randomly selected profile from the Tracebase database (Reinhardt 2017)

2.3. KEY PERFORMANCE INDICATORS FOR DIVERSITY EVALUATION

The objective of this work is to evaluate the influence of the model granularity on the diversity of the load profiles generated by a bottom-up simulator. The key performance indicators (KPIs) are thus related to the inter-household diversity and aim at characterising this diversity both in terms of intensity of use (i.e. how much) and of temporality of use (i.e. when). Some of the KPIs evaluated in this study have also been used by (Yamaguchi et al. 2019) and (Barsanti et al. 2024). The indicators are presented in Table 1.

Table 1 : Set of indicators used for comparing the modelling approaches.

	Description	Equation	Unit	Eq.
Total demand	Total energy consumption in the year	$E = \sum P(t)$	kWh	(1)
Maximum daily peak	Peak consumption of the average daily load profile	$P_{peak\ daily} = max(P_{av.daily}(t))$	Wh over 10 min	(2)
Entropy	Entropy of hour of switch-on, with pdf the probability distribution function at each time of day (10-min)	$H = -\sum pdf(t) \cdot ln(pdf(t))$	-	(3)
Morning-peak share	Ratio of energy consumed between 7am and 9am to the total energy	$\%_{morning\ peak} = \frac{\sum_{h\in[7-8]} P(t)}{\sum P(t)}$	%	(4)
Evening-peak share	Ratio of energy consumed between 5pm and 8pm to the total energy	$\%_{evening\ peak} = \frac{\sum_{h\in[17-20]} P(t)}{\sum P(t)}$	%	(5)
Self-production ratio	Ratio of energy supplied by a 3 kWp PV installation (south-orientation, La Rochelle) to the total consumption	$\%_{self-prod} = \frac{\sum P_{from\ PV}(t)}{\sum P(t)}$	%	(6)

The KPIs related to the intensity of use are the total demand (Eq. 1) and the maximum daily peak (Eq. 2). The other four KPIs are related to the temporality of use, with entropy (Eq. 3) representing the intra-household diversity of switch-on times. Lower entropy indicates a higher concentration of switch-on times within a limited timeframe. The morning-peak share, evening-peak share and self-production ratio aim to characterise building-to-grid interactions. The morning-peak share (Eq. 4) and evening-peak share (Eq. 5) evaluate the concentration of consumption during peak hours. The self-production ratio

(Eq. 6) represents the local consumption from a PV system. A 3 kWp system has been assumed, located in La Rochelle and facing south with a tilt of 20°. The PV output has been calculated using the *PVWatts* model from the *pvlib* library and the hourly solar irradiance model from *SARAH3*.

The main interest in these KPIs is to observe their range of variation within the simulated households, as this will characterise the models' ability to account for diversity. Therefore, the interquartile range (25-75%) is calculated and compared for these indicators.

3. RESULTS

To test and validate the methodology, this study focuses on energy use associated with washing machines. The diversity resulting from the different modelling types is evaluated using a sample of 100 households, which are composed as follows: 28 single-person households, 26 two-person households, 15 three-person households, 21 four-person households and 10 five-person households. This sample corresponds to the Elecdom panel (Ademe 2023) and the results are compared with these measurements. On average, this panel use the washing machine 3.2 times per week, for a total yearly consumption of 83 kWh.

3.1. INFLUENCE OF THE MODELLING OF THE CYCLE POWER PROFILES

Firstly, the influence of the cycle power profiles is evaluated, with four levels of detail being tested (see Section 2.2). The occupant model used corresponds to the 4-state Markov chain model (*occ_LoD 1*). The average duration of a washing machine cycle is 100 minutes. For *cycle_LoD 1*, the mean power is set to 220 W. For the simplified profile (*cycle_LoD 2*), the cycle has a peak of around 2 kW at the beginning (lasting 12 minutes). The distribution of the different KPIs is shown in Figure 2 and the interquartile range (IQR) are calculated in Table 2.

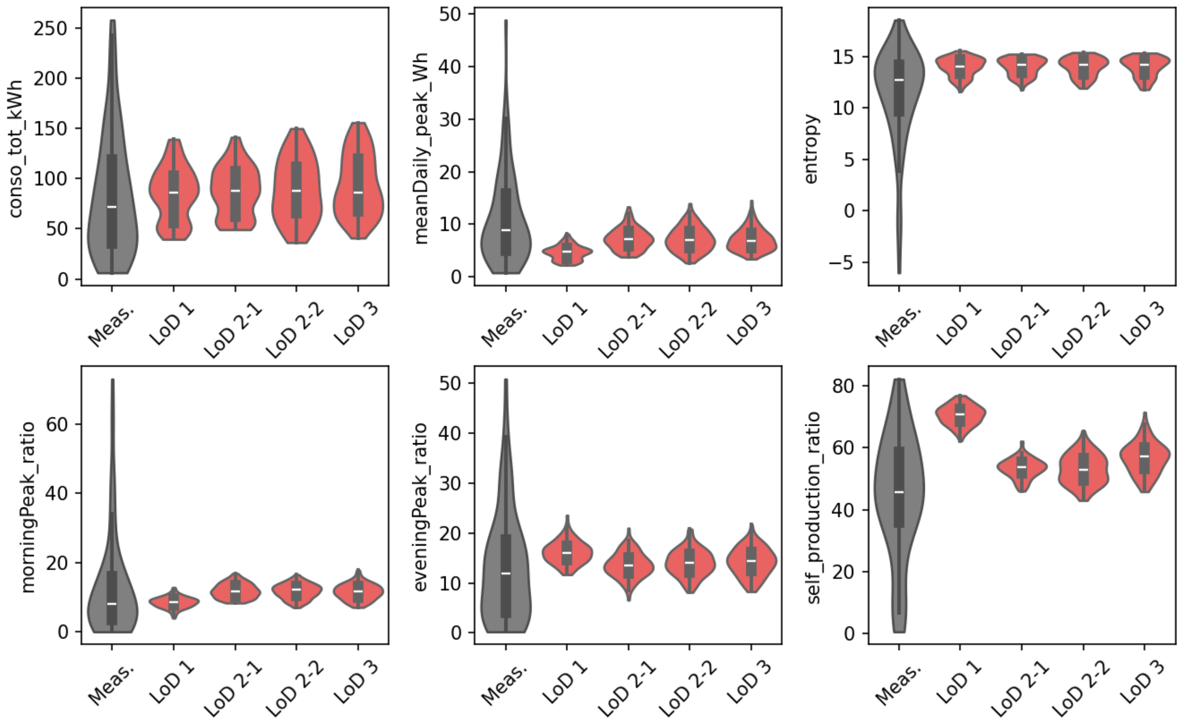

Figure 2: Violin plots of the KPIs for the measurements (in grey) and for various cycle power profile models (in red).

Table 2 : Characteristics of cycle power models and resulting interquartile ranges for some KPIs

	Profile	Energy class	IQR_{energy} [kWh]	$IQR_{dailypeak}$ [Wh]	$IQR_{entropy}$ [-]	$IQR_{self\text{-}prod}$ [-]
Measurements	-	-	83	10.9	4.5	22.8
Model *cycle_LoD 1*	Constant	Single (A+)	48	2.0	1.4	4.3
Model *cycle_LoD 2-1*	Simplified	Single (A+)	45	2.9	1.4	3.8
Model *cycle_LoD 2-2*	Simplified	Multiple (B > A+++)	46	**3.3**	1.4	**7.2**
Model *cycle_LoD 3*	Measured	Multiple	**52**	2.9	**1.5**	7.0

These results show that the level of detail in the cycle power profile is not of main importance for modelling the diversity. The variation in the number of cycles according to the number of persons in a household already creates a relatively large variation in consumption. Only the definition of energy classes improves diversity slightly, with a few more extreme intensity-of-use values. Using a simplified load shape (*cycle_LoD 2*) improves the modelling of the peak value and thus of the self-production by capturing the lost energy. Using of a measured cycle power profile does not deem necessary to improve the accuracy.

3.2. INFLUENCE OF THE MODELLING OF OCCUPANT PRESENCE AND/OR ACTIVITIES

Then, the influence of the modelling of occupants has been evaluated, with four levels of detail beeing tested (see Section 2.1). The measured cycle power profiles (*cycle_LoD 3*) were used to compare the models. The distribution of the different KPIs is shown in Figure 3 and the interquartile range (IQR) are calculated in Table 3.

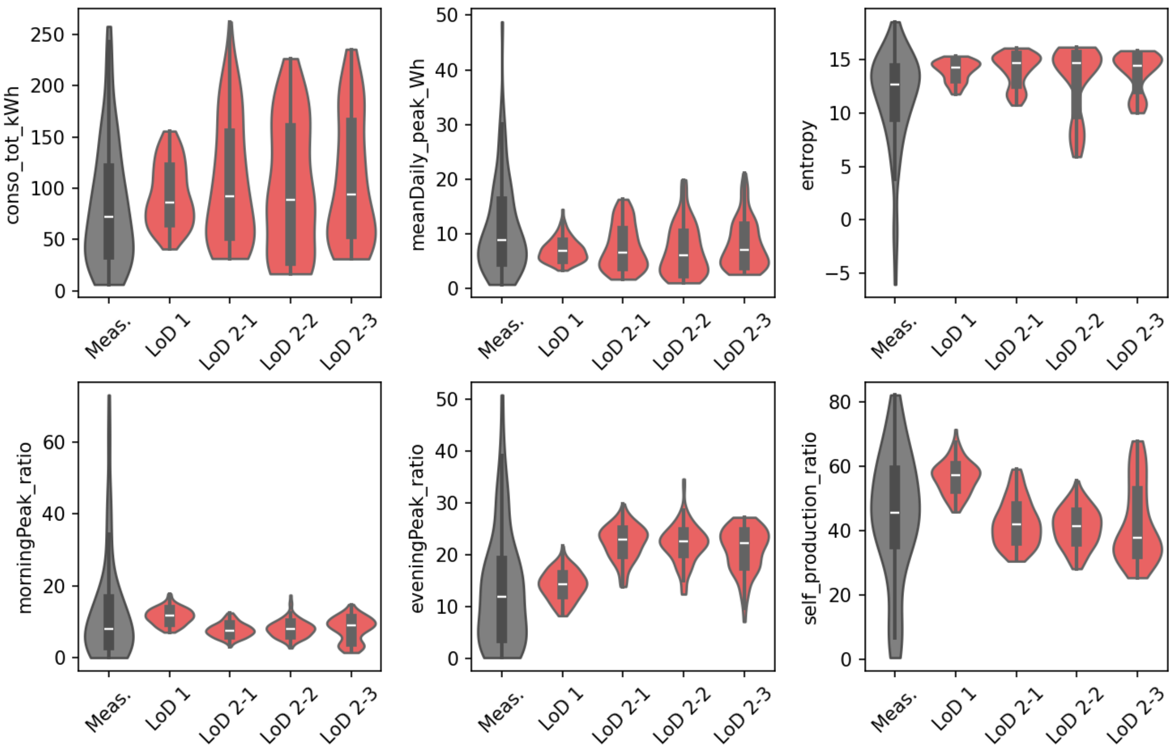

Figure 3 : Violin plots of the KPIs for the measurements (in grey) and for occupant models (in red).

Table 3: Characteristics of occupant models and resulting interquartile ranges for some KPIs (occ. = occupant-based / act. = activity-based).

	Occupant model		Nbr of TUS groups	IQR_{energy} [kWh]	$IQR_{dailypeak}$ [Wh]	$IQR_{entropy}$ [-]	$IQR_{self\text{-}prod}$ [-]
Measurements		-	-	83	10.9	4.5	22.8
Model *occ_LoD 1*	Occ.	Markov (4 states)	2 (adults/kid)	52	2.9	1.5	7.0
Model *occ_LoD 2-1*	Act.	Semi-Markov (13 states)	2 (adults/kid)	99	6.3	2.6	10.2
Model *occ_LoD 2-2*	Act.	Semi-Markov (13 states)	3 (man/woman/kid)	**128**	**7.2**	**5.5**	9.0
Model *occ_LoD 2-3*	Act.	Semi-Markov (13 states)	9 (man/woman/kid, single/couple, active/retired)	107	7.0	3.1	**19.2**

In terms of intensity of use (total consumption and mean daily profile), using an activity-based model rather than an occupancy-based model improves the representation of diversity, as demonstrated by the doubling of the IQR range for the two indicators. The distribution becomes also more skewed as the level of detail increases. In terms of the temporality of use, the different models still have limitations. Despite a slight improvement with the level of detail, more complex models struggle to capture households with regular schedules (characterised by low entropy) and those that use their washing machine outside of active occupancy (and overestimating the evening peak ratio). The evening peak ratio for occ_LoD 2 models is overestimated due to the mismatch between the related activity (laundry) and the switch-on time. The change to more subgroups in the TUS data only affects the representation of diversity for self-consumption and the tails of some distributions. This suggests that a large degree of diversity may already be present within a single subgroup (e.g. an active male worker). When comparing the average daily load profiles of measurements and simulations (Figure 4), the lack of temporal diversity in the modelling can be observed. Despite being stochastic, Markov chain models tend to reproduce similar patterns between households.

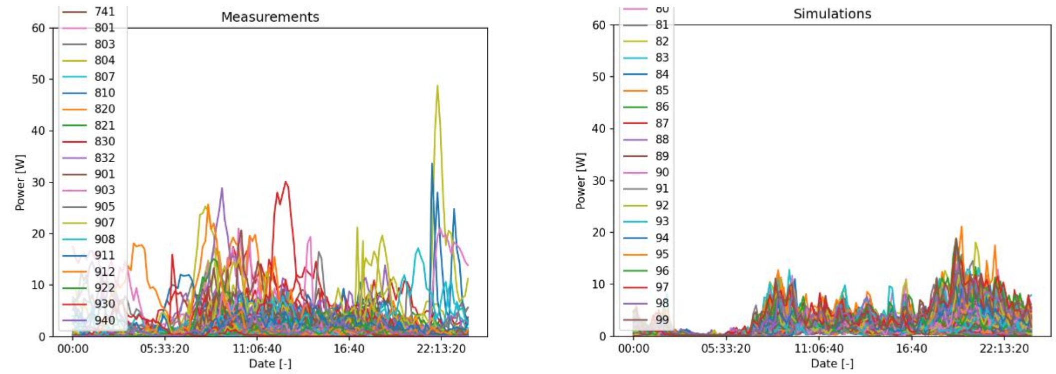

Figure 4: Average daily load profile of 100 households from the measurement (left) and from the LoD 2-3 model (right)

4. CONCLUSION

This work aimed at evaluating the impact of the level of detail on the modelling of diversity (in terms of intensity and temporality) in stochastic bottom-up models. The influence of cycle power profile models and occupant presence/activities model has been evaluated and compared to measured data on 85 washing-machines. This analysis showed that simplified power profiles were sufficient to capture equipment diversity, and that activity-based models performed better at representing behavioural diversity, with a wider range of intensity and temporal variability of use. However, limited temporal diversity in usage was observed, even with nine different subgroups of persons.

This analysis can inform different future research directions. Firstly, non-supervised classification techniques should be tested on the TUS data to create a wider range of subgroups. Then, other modelling techniques should be added to the benchmark, such as regression-based methods (e.g. StROBe) or agent-based methods (e.g. LPG or SMACH).

5. SUPPLEMENTARY MATERIAL

The code used to generate these results is available on https://gitlab.univ-lr.fr/jledreau/demodfr.

6. ACKNOWLEDGMENT

This work was carried out as part of the PEPR FlexTASE research project (ANR 22-PETA-0009).

7. BIBLIOGRAPHY

Ademe. 2023. « Panel usages électrodomestiques ». https://librairie.ademe.fr/changement-climatique-et-energie/4473-panel-usages-electrodomestiques.html.

Barsanti, Matteo, Selin Yilmaz, et Claudia R. Binder. 2024. « Informing targeted Demand-Side Management: Leveraging appliance usage patterns to model residential energy demand heterogeneity ». *Energy and Buildings* 321 (octobre): 114639. https://doi.org/10.1016/j.enbuild.2024.114639.

Bottaccioli, Lorenzo, Santa Di Cataldo, Andrea Acquaviva, et Edoardo Patti. 2019. « Realistic Multi-Scale Modeling of Household Electricity Behaviors ». *IEEE Access* 7: 2467-89. https://doi.org/10.1109/ACCESS.2018.2886201.

Carlucci, Salvatore, Marilena De Simone, Steven K. Firth, et al. 2020. « Modeling occupant behavior in buildings ». *Building and Environment* 174 (mai): 106768. https://doi.org/10.1016/j.buildenv.2020.106768.

INSEE. 2010. « Enquête Emploi du temps 2009-2010 ». https://www.insee.fr/fr/metadonnees/source/operation/s1362/presentation.

INSEE. 2022. « Population census ». https://www.insee.fr/en/metadonnees/source/serie/s1321.

Le Dréau, J. (DEMOD-FR). 2025. « GitLab ». https://gitlab.univ-lr.fr/jledreau/demodfr.

Malik, Jeetika, Elie Azar, Ardeshir Mahdavi, et Tianzhen Hong. 2022. « A level-of-details framework for representing occupant behavior in agent-based models ». *Automation in Construction* 139 (juillet): 104290. https://doi.org/10.1016/j.autcon.2022.104290.

Reinhardt, Andreas. 2017. « tracebase ». https://github.com/areinhardt/tracebase.

Richardson, Ian, Murray Thomson, David Infield, et Conor Clifford. 2010. « Domestic electricity use: A high-resolution energy demand model ». *Energy and Buildings* 42 (10): 1878-87. https://doi.org/10.1016/j.enbuild.2010.05.023.

Sanderson, Edward, Aikaterini Fragaki, Jules Simo, et Bogdan Matuszewski. 2020. « Comparing Bottom-Up Energy Consumption Models Using The Wasserstein Distance Between Load Profile Histograms ». *BSO-V 2020: IBPSA-England Building Simulation and Optimisation Conference 2020*. https://publications.ibpsa.org/proceedings/bso/2020/papers/bso2020_Sanderson.pdf.

Wu, Yi, Jingjing An, Mingyang Qian, et Da Yan. 2023. « Application-driven level-of-detail modeling framework for occupant air-conditioning behavior in district cooling ». *Journal of Building Engineering* 70 (juillet): 106401. https://doi.org/10.1016/j.jobe.2023.106401.

Yamaguchi, Y., S. Yilmaz, N. Prakash, S. K. Firth, et Y. Shimoda. 2019. « A Cross Analysis of Existing Methods for Modelling Household Appliance Use ». *Journal of Building Performance Simulation*, mars 4. World. https://www.tandfonline.com/doi/abs/10.1080/19401493.2018.1497087.