



# Development of a methodology of validation for urban building energy models and application to French residential consumption

Lucie Lefort, Romain Bonabe de Rougé, Pascal Schetelat, Thomas Berthou, Peter Riederer, Bruno Duplessis, Eric Peirano

## ► To cite this version:

Lucie Lefort, Romain Bonabe de Rougé, Pascal Schetelat, Thomas Berthou, Peter Riederer, et al.. Development of a methodology of validation for urban building energy models and application to French residential consumption. uSIM2020, Nov 2020, Edenbourg (on line), United Kingdom. hal-03202728

**HAL Id: hal-03202728**

**<https://hal.science/hal-03202728>**

Submitted on 20 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Development of a methodology of validation for urban building energy models and application to French residential consumption

Lucie Lefort<sup>1,2,3</sup>, Romain Bonabe de Rougé<sup>1</sup>, Pascal Schetelat<sup>2</sup>, Thomas Berthou<sup>3</sup>,  
Peter Riederer<sup>2</sup>, Bruno Duplessis<sup>3</sup>, Eric Peirano<sup>1,3</sup>

<sup>1</sup>Efficacity, Champs-sur-Marne, France

<sup>2</sup>CSTB, Paris, France

<sup>3</sup>Mines ParisTech, PSL Research University, Paris, France

### Abstract

Cities need to drastically reduce their energy consumption in order to fight climate change. To help them implement efficient energy saving measures, urban building energy models (UBEM) have been developed. A large amount of input data is required to model buildings at an urban scale and that information is often missing. To overcome this lack of model input data, data enrichment approaches including archetypes and calibration processes have been used, resulting in errors in the generated data which are not always well quantified. The aim of this paper is to develop a methodology of validation for UBEM that takes into account a data enrichment process. The proposed approach is to split the validation process into several parts based on UBEM use cases. UBEM outputs and enriched data to be validated are selected for each use case and the fitness for purpose of the model is then verified against measured data. This validation method has been tested on a specific use case: the prediction of yearly electricity and gas consumptions of dwellings aggregated at the district scale. It has been applied to two different UBEM tools: Smart-E and Dimosim. Smart-E uses its own hypothesis for data completion while a specific tool, Qiriél, is used to infer missing data for Dimosim. The simulation of one hundred districts (90 000 dwellings) located in the French department of Seine-et-Marne and the comparison to measured data enabled to study the error distribution of both UBEM tools. Further work should be led at smaller spatial and time scales.

### Introduction

With an urgent need to act against climate change, more and more cities set ambitious environmental targets. In order to achieve such goals, reducing energy consumption in buildings seems to be a keystone. In France, for example, tertiary and residential sectors represented 45% of the national final energy consumption in 2015 (ADEME, 2015). To assist in lowering these consumptions at the city scale, urban building energy models (UBEM) have been developed over the last fifteen years (Reinhart and Cerezo Davila,

2016). UBEM compute the energy consumption of every building in a district and take into account interactions among them, like, for example, solar shading. The aim of UBEM is to provide cities with guidance to reduce final energy consumption not only for some specific buildings but also at the scale of a whole district.

However, the benefits brought by UBEM are balanced with certain challenges that may jeopardize their reliability.

A first challenge is the data enrichment process. It is much more difficult to gather data required for the simulation of one hundred buildings than for a single one. Information is often missing and assumptions must then be made. This practice of data enrichment creates an error that is not yet well-quantified. A second challenge is to take into account new phenomena that appear at the district scale, like the phenomenon of urban heat islands.

Finally, although a building energy model (BEM) has to be precise for each building it simulates, this is not necessarily the case for a UBEM. It may be acceptable for a UBEM to be statistically correct at the scale of a district but to be rather inaccurate at the scale of a single building. The challenge is then to identify the spatial and time scales on which the UBEM is reliable.

A thorough validation process would enable to verify the reliability of UBEM and would give them more credits as decision-support tools for cities. However, such validation process is not yet implemented within the UBEM research community and is therefore needed.

The aim of this paper is to present a validation methodology designed specifically for UBEM. This methodology is described in the next section before being applied to two UBEM tools in a second section. Results obtained by the two UBEM tools throughout the validation process are then being discussed.

### Development of a validation methodology for urban building energy models

Most UBEM stem from building energy models (BEM). However, at the difference with the BEM,

there is no consensus yet on their validation. In this section, we review the existing validation attempts that can be found in the literature and we draw a list of requirements for a validation methodology. Then, we describe a first proposal of validation methodology based on UBE M use cases.

### Existing attempts of validation

#### *BEM validation*

Building energy model can be seen as a cornerstone of UBE M. Therefore it seems natural to look at the existing validation process implemented for BEM.

A lot of research has been carried on BEM validation among which were the BESTEST and the PASSYS project.

In the BESTEST (Judkoff and Neymark, 2013), several sets of perfectly described buildings are provided and inter-comparison of models are computed on these well-known buildings. In addition, some empirical validation has been proposed with measurements on test cells.

Jensen (1995) gives a precise description of the methodology developed within the PASSYS project. The proposal is to split up the validation process into two main parts: the validation of single processes and the validation of the whole model. Jensen encourages namely to use analytical verification for single process validation and to use empirical validation for whole model validation.

The validation of thermal models used within UBE M through one of these two validation methods would not suffice to consider a UBE M tool as properly validated for several reasons:

- Additional physics and additional energy systems are modelled at the urban scale;
- An enrichment process is often used as a prior to UBE M and needs to be validated as well;
- It may be acceptable for UBE M to be accurate only at certain spatial scales, larger than the building scale.

However, some principles from these methodologies should be used as guidance for the development of a UBE M validation methodology. For example, separating single process validation and the whole model validation as advised by Jensen (1995) is of great interest for UBE M as well. Furthermore, in the same idea as the PASSYS project, the aim of the present work is to develop a validation methodology rather than validating a single tool.

Regarding validation attempts on UBE M, two approaches have been observed so far: comparative validation and empirical validation.

#### *Comparative validation*

Inter-model comparison has started within two international research projects, the IEA Annex 60 (Wetter, 2017) and the IBPSA Project 1 (Saelens et al., 2019) which is still ongoing. In both projects, the

taken approach is quite similar to the BESTEST (Judkoff and Neymark, 2013). A synthetic district has been designed to serve as a District Energy Simulation Test (DESTEST). Several UBE M tools are being tested on this DESTEST and their results are compared. This setup enables to control all the simulation parameters and thus proceed with the single process validation as advocated by Jensen (1995). However, since the district used is not real, it is not possible to compare the obtained results with measured data nor to test the enrichment process used by each UBE M tool.

#### *Empirical validation*

The aim of empirical validation is to ensure that UBE M tools represent the real-life urban projects accurately. However, such process is made difficult due to a lack of data at district and urban scales. Some examples can be found in the literature like in Sokol et al. (2017) or in Fonseca and Schlueter (2015). These attempts give a first idea of UBE M performances on real projects. However, datasets used in these validation processes are rarely publicly available thus preventing other research teams from going through the same validation process. Furthermore, there is no agreement so far on the metrics to be used to measure the error on UBE M outputs. This prevent the reader from comparing results obtained in different papers.

Nonetheless, some requirements can be drawn from these existing attempts of UBE M validation.

#### *Requirements for a UBE M validation methodology*

Some lessons can be learned from this brief literature review and it can be summarized as follows. A validation methodology dedicated to UBE M should comply with the following requirements:

- Specifying the spatial and time scales on which a given UBE M tool is reliable;
- Defining metrics to properly measure the UBE Ms accuracy throughout the validation process;
- Identifying the error generated by the model setting;
- Ensuring applicability and replicability of the methodology in different contexts (country, building sector etc.);
- Proposing test cases to apply the developed validation method.

A first proposal of methodology is presented hereafter based on these requisites.

### A validation method based on UBE M's use cases

The methodology developed here offers to inventory UBE Ms use cases and to identify, for each use case, critical UBE M outputs. The aim is to verify that the outputs of interest for each use case are accu-

rate enough to provide a real aid to the decision-maker. Such outputs of interest are called key performance indicators (KPI) in the rest of the paper. Once use cases and KPI have been identified, validation datasets are used in order to study, for each use case, the error distribution on the KPI. This methodology is described in Figure 1.

In an ideal case, several batches of measured data would be available, enabling to obtain an error distribution for each KPI of each use case and UBEM tool. The main idea behind this search for an error distribution is to provide the UBEM user with a confidence interval on the KPI, depending on the use case. UBEM use cases and corresponding KPI for validation were identified through a literature review that is not presented in this paper. Although, the main use cases are listed below:

- Development of renewable energies at the district scale;
- Energy flexibility;
- Performance measurements and verification;
- Buildings retrofitting;
- Study of the urban heat island effects and its impact on indoor comfort;
- Future energy use;
- Urban energy planning.

Future works should properly define metrics for each use case and for each KPI in order to measure the error on the KPI accurately.

A drawback brought by the proposed validation methodology is the need for large validation datasets when such datasets are often hard to find. For example, if the aim of the study is an error distribution at the district scale at least hundred districts with measured data would be needed to obtain this distribution. To tackle this issue, it is proposed to start the validation process with use cases for which additional data are available. Then, we could study in what extend such use cases could give some bounds for the error distribution of use cases with less validation data. For instance, if the annual energy consumption of a building is accurate, it does say something about its annual load profile. Furthermore, more and more energy disclosure laws are being implemented in western countries, giving the opportunity to access large buildings energy consumptions databases. Examples of these energy disclosure laws can be found in France but also in other countries like in the USA (Kontokosta and Tull, 2017).

## Application of the proposed methodology to two UBEM tools

### Validation framework

A first use case was chosen to apply and test the proposed validation methodology. This use case can be described as territorial energy diagnosis. It consists in being able to give an accurate outlook of a district

energy consumption. In this study, the annual electricity and gas consumptions of residential buildings were studied at the district scale. This use case may be seen as a preliminary study for other use cases since a first step for UBEM is to provide accurate representation of existing consumptions before evaluating different scenarios like the retrofitting of buildings or the development of renewable energies. The choice of this use case was also driven by the publicly available datasets.

This use case was applied on the French department of Seine-et-Marne at the scale of the IRIS. The IRIS is an administrative division of the French territory. Each IRIS gathers around 2000 inhabitants. All residential buildings located in about a hundred of IRIS were simulated over a year and their annual electricity and gas consumptions aggregated at the IRIS scale were compared to measured data. Simulations were completed with two different UBEM tools, Dimosim (Riederer et al., 2015) and Smart-E (Berthou et al., 2015), both relying on automated data enrichment processes, and results were compared. This case study deals only with the residential sector since less databases are available for the simulation of the tertiary sector in France. Datasets used for the simulation and for the validation are presented hereafter.

### Simulation datasets

Three databases were used for simulation:

- BD TOPO® (IGN, 2019) This database provides a map of the whole French territory where each building is described by a footprint and a height. Additional information is also given per building such as:
  - Building function (residential, commercial, etc.)
  - Construction year
  - Number of floors
  - Number of dwellings

Though the footprint and the height are given for every building, the additional information is not always available. This additional information is obtained through a joint between tax record files and the map. Information is missing when the joint has not been successful. The BD TOPO® is not publicly available yet but it can be granted for free for any research project. It should become completely public by 2022.

- PHEBUS (Conseil National de l'Information Statistique, 2013): This database is built upon a survey completed in 2013 for 5000 households and for each household, their appliances and their energy bills were reported. Following this first step, energy performance audits were completed in 2000 out of these 5000 households.
- Fichiers Détails logements (FDL) (INSEE, 2016): This database originates from the population census. Each row corresponds to 1 to

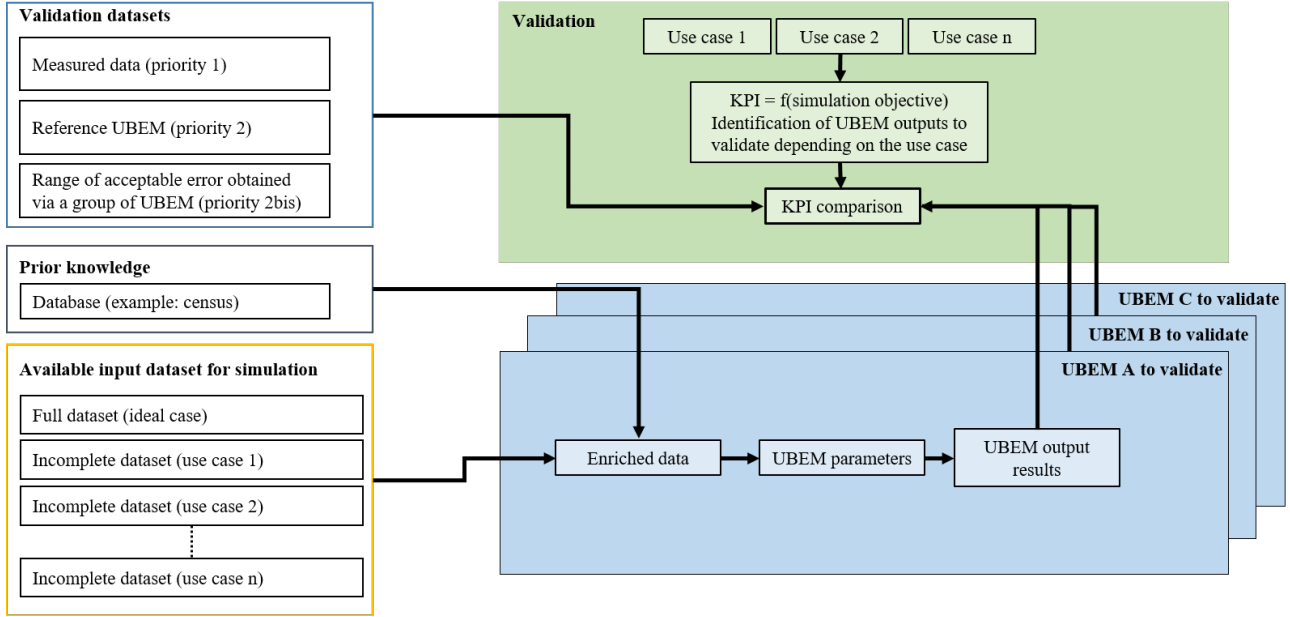


Figure 1: Proposed validation methodology

5 dwellings with identical properties. The precise location of these dwellings is unknown but the zipcode of the IRIS where they are located is provided on each row. Information contained in this database includes:

- Dwelling’s number of occupants;
- Socio-professional status of the household;
- Surface category of the dwelling;
- Type of dwelling (individual housing or collective housing);
- Construction year of the building;
- Main heating energy;
- Main heating system type (centralized or decentralized).

#### Validation dataset

The validation dataset employed for this use case is the Données Locales de l’Energie database (SDES, 2020). This database originates from a French law for the ecological transition that makes compulsory for energy suppliers to disclose annual energy consumptions data aggregated at the IRIS scale and detailed by sector. Annual electricity and gas consumption of the residential sector are thus available at the IRIS scale. These data are publicly-available and are updated every year. The related disclosure law was recently amended compelling energy suppliers to publish energy consumption data at the building scale, given some privacy threshold for the residential sector. These datasets requires some preprocessing but are of great interest for the validation of UBEM.

#### Two different UBEMs, one validation process

The proposed validation method aims to be applicable to a large range of UBEM. Two first UBEM tools were used here to test this method: Dimosim

(Riederer et al., 2015) and Smart-E (Berthou et al., 2015). These tools are very different from each other. A first difference comes from the data enrichment process. Smart-E uses internal hypotheses to complete missing data while Dimosim is coupled with a tool designed for the enrichment process: Qiriél (Ansanay-Alex et al., 2016).

A second difference comes from the datasets used as inputs data and as parameters. Dimosim takes as input data the geometry of the buildings retrieved from BD TOPO® and then tries to assess thermal properties and household characteristics thanks to PHEBUS and FDL. Conversely, Smart-E takes as input data the description of households provided by FDL that includes the surface of the dwellings but not the geometry of the buildings and then tries to allocate façade areas and roof areas thanks to BD TOPO®. BD TOPO® and FDL both give information on dwellings but since FDL comes from the census, it is anonymized to prevent identifying the households and the buildings. Therefore, the link between these two databases is not straightforward and assumptions must be made. Furthermore, information provided by these databases is not always consistent. For example, the number of dwellings recorded in BD TOPO® and FDL is different.

A third difference is the number of thermal zones defined for the simulation. Several thermal zones per building can be defined in Dimosim but for the sake of simplicity only one thermal zone per building was used in this case. Smart-E uses two thermal zones per dwelling by default, one for the heated space and one for the non heated space.

Booth et al. (2012) identified the most influential parameters on housing stock modelling uncertainties. The mean value of these parameters distributions

used for each UBEM tool in this study are summarized in Table 1. These values are the average on all the simulated dwellings weighted by the dwellings surfaces. These mean values have been displayed to illustrate the difference of parameterization between the two tools. For example, the share of dwellings heated with gas and the share of dwellings heated with electricity vary from an IRIS to another and are often slightly different between Dimosim and Smart-E due to different enrichment processes.

Table 1: Influential parameters on results uncertainty

Parameters	Dimosim	Smart-E
Fraction of space heated	100 %	77 %
Heating temperature set-point (C)	19.6	19.6
Share of dwellings with gas as heating energy	48.3 %	50.6%
Share of dwellings with electricity as heating energy	38.2 %	35.7 %
Efficiency of electric heaters	1.0	1.0
Efficiency of gas boilers	0.78	0.69
Window U-value (W/m <sup>2</sup> K)	3.21	2.69
Wall U-value (W/m <sup>2</sup> K)	0.96	1.38
Window-to-wall ratio	0.25	0.15
Air-change rate (including infiltrations and ventilation) (vol/h)	0.6	0.52
Number of simulated dwellings	89876	93717
Average surface per dwelling (m <sup>2</sup> )	105	92

## Results & Discussion

### KPI and metrics

The proposed methodology was tested on 93 IRIS (around 90000 dwellings), all located in the French department of Seine-et-Marne. The studied KPI in this case were the annual electricity and gas consumptions at the IRIS scale but also the annual thermosensitivity for gas and electricity at the IRIS scale. The thermosensitivity is defined as the increase of electricity or gas consumption due to the decrease of the outdoor temperature. Its estimation for each tool was done following the methodology provided with the DLE (ENEDIS et al., 2018). Measured data were available in the DLE for each of these KPI and for each IRIS. The error on annual consumption and thermosensitivity was measured through a percentage of error (PE) defined in Equation 1.

$$PE_X = \frac{X_{Simulated} - X_{measured}}{X_{measured}} \cdot 100 \quad (1)$$

The results are presented in Figures 2 and 4.

### Analysing the error distribution with respect to input data

The error distribution obtained with Dimosim on annual consumptions is shifted toward the positive values, meaning that Dimosim mainly overestimates the electricity and gas consumptions. Conversely, Smart-E tends to underestimate these consumptions. In absolute values, Dimosim seems to generate a greater error than Smart-E on this particular use case. This performance gap could come from the fact that Smart-E is calibrated in order to fit national electricity consumptions while there is no similar process in Dimosim.

Regarding the error on the thermosensitivity, it seems that Dimosim also overestimates this phenomenon for both gas and electricity. An hypothesis is then that Dimosim overestimates the heating needs in general. When comparing inputs parameters between Smart-E and Dimosim, the fraction of space heated looks like a good candidate to help reducing this error. In order to confirm this hypothesis, simulations should be run again with a different space-heated ratio and results should then be compared.

When taking a closer look at the error generated by Smart-E, the error on the electrical thermosensitivity seems surprisingly high compare to the error on the annual electricity consumption. In order to get a better understanding of all these error distributions and to try to explain phenomenon like the one observed on the thermosensitivity for Smart-E, an analysis of the correlation between the error and a large set of inputs parameters have been completed.

Correlations were estimated through the Pearson correlation coefficient. This coefficient is defined in equation 2 for two variable X and Y, where  $cov(X,Y)$  is the covariance of X and Y,  $\sigma_X$  is the standard deviation of X and  $\sigma_Y$  is the standard deviation of Y. A coefficient of 1 indicates a linear relationship between variables while a coefficient close to 0 indicates that there is no linear relationship (but a non-linear relationship is still possible between variables).

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} \quad (2)$$

Coefficient correlations between the percentage of error on annual gas and electricity consumptions, gas thermosensitivity and electrical thermosensitivity and inputs parameters are presented in Figure 3.

In general, the obtained coefficients are quite low, most of them being below 0.5 in absolute value. It is then difficult to draw influential parameters from this study. However, some trends can be identified. Thus, the two UBEM tools seem to behave quite differently in terms of error. Smart-E presents a coefficient of 0.64 between the percentage of error on the gas consumption and the one on the electricity consumption but such correlation does not appear for Dimosim.

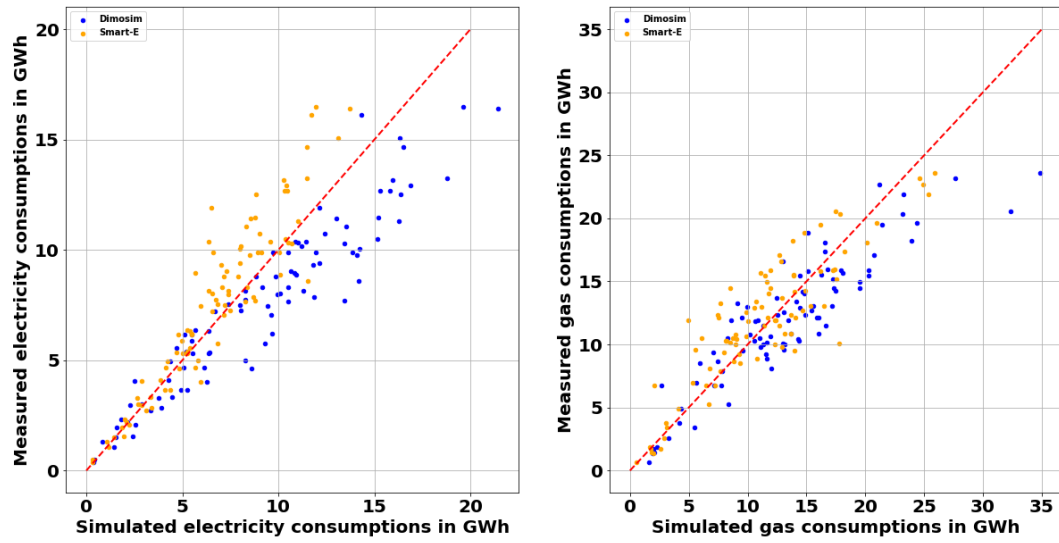
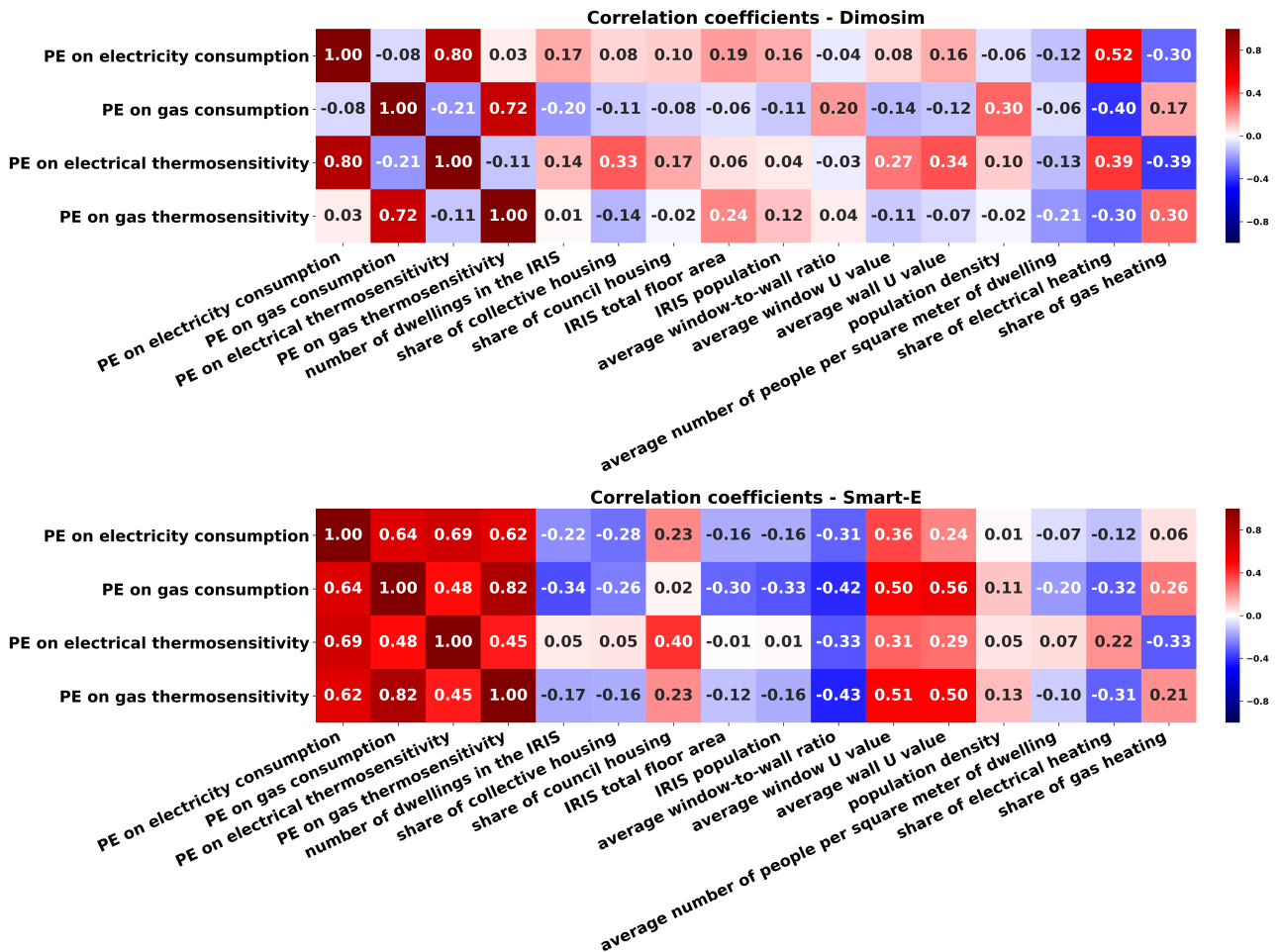


Figure 2: Annual measured and simulated consumptions of gas and electricity



Furthermore, the higher coefficients can be found for Smart-E between the percentages of error (gas and

electricity) and the window and wall U-values. In Dimosim, a relatively high correlation is observed be-

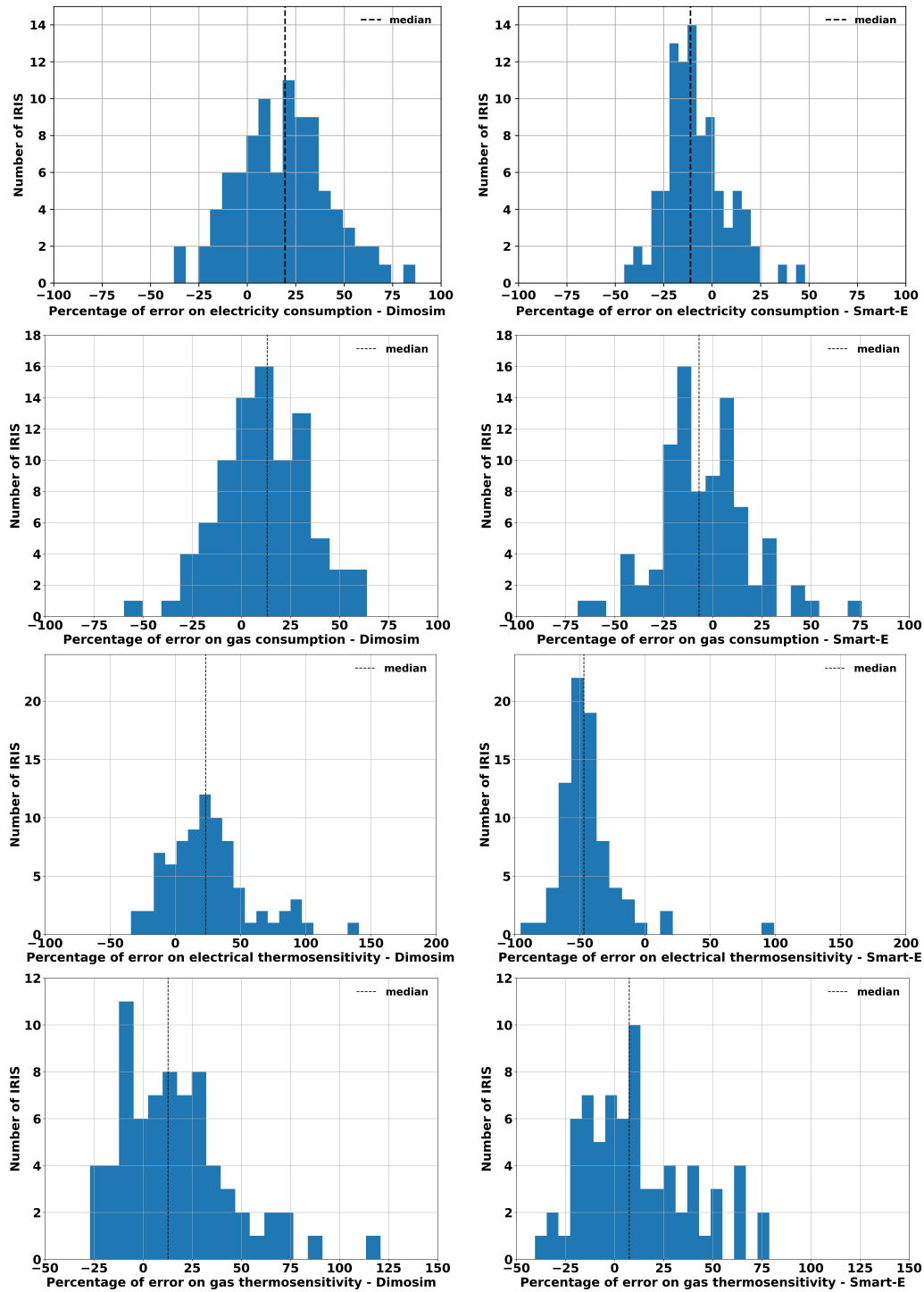


Figure 4: Percentage of error on annual energy consumptions and on thermosensitivity

tween the percentage of error on electricity and the share of dwellings using electricity as heating energy. Finally, a mild correlation between Dimosim and Smart-E errors was identified as well, questioning the quality of databases used for simulation and validation and the error they might bring in. Thus, this analysis enabled to identify levers to reduce the overall error of both tools.



## Conclusion

A methodology of validation designed specifically for UBEM was presented in this paper. This method aims to verify that UBEM are accurate enough to provide a real aid in the decision-making process. To do so, the validation procedure is broken down based on UBEM use cases and error distributions are studied for KPI of each use cases.

This method was tested with two first UBEM tools on a French test case. The analysis of the obtained error distributions enabled to identify weaknesses of both UBEM tools and hence improve their performances. The development of this methodology is still ongoing and future work should be carried on UBEM outputs at smaller spatial and time scales. The error analysis and the identification of critical inputs parameters in the output error generation should also be strengthen. Lastly, an important step would be to verify the applicability of the methods to a large range of UBEM tools. Since the available datasets for simulation influence greatly the enrichment process and UBEM behaviour in general, it would be of great interest to apply this methodology in other countries than France and with other UBEM tools as well. The authors would then be very interested in collaborating with other research teams on this topic.

## References

- ADEME (2015). Chiffres-clés Climat, Air et Énergie. Technical report.
- Ansanay-Alex, G., Y. Abdelouadoud, and P. Schetelat (2016). Statistical and Stochastic Modelling of French Households and their Energy Consuming Activities. In *12th REHVA World Congress (CLIMA)*.
- Berthou, T., B. Duplessis, P. Rivière, P. Stabat, D. Casetta, and D. Marchio (2015). SMART-E: A TOOL FOR ENERGY DEMAND SIMULATION AND OPTIMIZATION AT THE CITY SCALE. In *Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association, Hyderabad, India, Dec. 7-9, 2015*, pp. 1782–1789.
- Booth, A. T., R. Choudhary, and D. J. Spiegelhalter (2012). Handling uncertainty in housing stock models. *Building and Environment* 48(1), 35–47.
- Conseil National de l'Information Statistique (2013). Enquête Performance de l'Habitat, Équipements, Besoins et Usages de l'énergie (Phébus).
- ENEDIS, GRDF, and Agence ORE (2018). Estimation des données de thermosensibilité et de part thermosensible. Technical report.
- Fonseca, J. A. and A. Schlueter (2015). Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts. *Applied Energy*.
- IGN (2019). BD TOPO Version 30. - Descriptif de contenu.
- INSEE (2016). Logements ordinaires en 2016 - Recensement de la population - Fichiers détail.
- Jensen, S. Ø. (1995). Validation of building energy simulation programs: a methodology. *Energy and Buildings* 22(2), 133–144.
- Judkoff, R. and J. Neymark (2013). Twenty years on!: Updating the IEA BESTEST building thermal fabric test cases for ASHRAE standard 140. *Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association* (January 2013), 63–70.
- Kontokosta, C. E. and C. Tull (2017). A data-driven predictive model of city-scale energy use in buildings. *Applied Energy* 197, 303–317.
- Reinhart, C. F. and C. Cerezo Davila (2016). Urban building energy modeling - A review of a nascent field. *Building and Environment*.
- Riederer, P., V. Partenay, N. Perez, C. Nocito, R. Triggance, and T. Guiot (2015). Development of a simulation platform for the evaluation of district energy system performances. In *Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association, Hyderabad, India, Dec. 7-9, 2015*, pp. 2499–2506.
- Saelens, D., I. De Jaeger, F. Bünning, M. Mans, A. Vandermeulen, B. Van Der Heijde, E. Garreau, A. Maccarini, Ø. Rønneseth, I. Sartori, and L. Helsen (2019). Towards a DESTEST: a District Energy Simulation Test Developed in IBPSA Project 1. In *Proceedings of BS2019: 16th Conference of International Building Performance Simulation Association, Rome, Italy, Sept. 2-4, 2019*.
- SDES (2020). Données locales de consommation d'énergie.
- Sokol, J., C. Cerezo Davila, and C. F. Reinhart (2017). Validation of a Bayesian-based method for defining residential archetypes in urban building energy models. *Energy and Buildings*.
- Wetter, M. (2017). New Generation Computational Tools for Building & Community Energy Systems Annex 60 Final Report. Technical report, International Energy Agency.