



HAL
open science

A novel approach for electric load curve holistic modelling and simulation

Thibaut Barbier, Robin Girard, François-Pascal Neirac, Nicolas Kong,
Georges Kariniotakis

► **To cite this version:**

Thibaut Barbier, Robin Girard, François-Pascal Neirac, Nicolas Kong, Georges Kariniotakis. A novel approach for electric load curve holistic modelling and simulation. MedPower 2014 - The 9th Mediterranean Conference on Power Generation, Transmission Distribution and Energy Conversion, Nov 2014, Athènes, Greece. hal-01087253

HAL Id: hal-01087253

<https://minesparis-psl.hal.science/hal-01087253v1>

Submitted on 1 Dec 2014

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.

A novel approach for electric load curve holistic modelling and simulation

Thibaut Barbier, Robin Girard, François-Pascal Neirac, Nicolas Kong and Georges Kariniotakis

Abstract-- This paper presents a novel approach of an electric load curve simulator using a set of grey box models that results to an efficient trade-off between complete and complex physical models and fast simplified statistical models.

The input parameters are macroscopic data coming from large databases such as national census, DSO's client information and meteorological data such as temperature or irradiation data. The problem of matching between the different databases is investigated to assess comparable load curves. Validation is performed using load measurements at the medium voltage level. Once the model is calibrated it can be turned into a good prediction tool useful for planning studies since it permits easily to incorporate the evolution of usages, the characteristics of consumption devices, as well as the evolution of the building's characteristics.

Index Terms— automatic calibration, DSO, electric load curve modelling, electric load curve comparison, grey box models, large databases, power system planning, smartgrids, statistical modelling.

I. INTRODUCTION

A. Context of the Study and General Issues

THE majority of power systems actors today face important challenges related to climate change, fossil fuel price increase, security of supply, deregulation and other factors that characterize a changing environment. The need to take into account the ternary causes: sobriety, efficiency and renewable energy is here. In this context it is important to be able to understand actual electricity consumption and possible ways for saving electricity, not only in terms of energy but also in terms of instantaneous power, at different scales on the grid.

In this context it is important to be able to understand actual electricity consumption; first to target possible ways for saving energy, but also to understand temporal and spatial variation of instantaneous power.

Indeed, sizing the electricity grid, particularly the public distribution network, by predicting the electricity consumption spatially in term of maximum power is

becoming a challenge. The Distribution System Operators (DSO), have to find trade-offs between security of supply even under the worst scenario conditions and minimization of investments. For that purpose they have sized distribution networks by considering principally statistical parameters and past measurements (temperature, electrical consumption...) linked to the peak power consumption. However, this approach is becoming less and less adapted, with the ongoing evolution of the context:

- new uses and technical innovations leading to huge changes in electricity consumption (e.g. electric vehicles, passive houses);
- new possibilities of smart load control with a massive deployment of information and communication technology devices including smart meters.
- more and more decentralized production (e.g. photovoltaic panels, cogeneration);

Hence, the DSO looks for a development of electric loads models at different spatial scales able to account for the above mentioned evolutions. The spatial scales of interest range from the substation that gathers thousands of electricity customers down to single or small groups of individual customers. On the other hand it is important to have models with a fine temporal scale (down to the minute resolution), in order to be able to simulate as realistically as possible peak electricity consumption in the different scenarios of future demand.

This improved knowledge of the electricity demand concerns also other actors ranging from district and state policy makers, down to the electricity customers. For instance, a fine modeling of electricity demand can be an advisor for politicians to choose which suburb has to be renovated in terms of energy needs in priority and how (e.g. heat network or individual heat pumps).

This research topic is not recent: for decades several authors have looked into forecasting models for short and long term electric demand, this later being of interest here. Historically, there are two mainstream approaches for modelling the electric energy demand: top-down and bottom-up models. This classification is based on the data used as input to the models, according to Swan and Ugursal [1]. If this data is "macroscopic", dealing with a housing stock for instance, it will be a top-down model. However, if input is "microscopic", and the calculation is done consumer by consumer and then there is an extrapolation to a total consumer stock, the model is bottom-up. In order to have a physical sense, and a high level of detail, even in the absence of energy demand measurements, the bottom-up models are preferentially used, but their requirement for high computation time can be an issue. On contrary, top-down

This work was supported in part by ERDF and Cap Energies.

T. Barbier, R. Girard, F-P. Neirac and G. Kariniotakis are with MINES ParisTech, PSL - Research University, PERSEE - Centre for Processes, Renewable Energies and Energy Systems, 06904 Sophia Antipolis Cedex, France (e-mails respectively: thibaut.barbier@mines-paristech.fr, robin.girard@mines-paristech.fr, francois-pascal.neirac@mines-paristech.fr, georges.kariniotakis@mines-paristech.fr).

N. Kong is with ERDF, Paris La Défense, France (e-mail: nicolas.kong@erdf.fr).

models are useful when the need is to generate rapidly a trend evolution function of some general information (i.e. unemployment rate, Gross Domestic Product) on a large scale stock of electricity consumer, but they cannot simulate well the impacts of technical breaks or new consumption devices which have not been measured [2]. Yet, some authors have tried to take advantage of both models, developing hybrid models.

Indeed, macroscopic data is becoming more and more precise and available: population census are in many countries public data going down to a scale lower than cities, including detailed databases of buildings supported by satellite images. Also, the DSOs are instrumenting more and more their grid especially with the large-scale deployment of smart meters. They can have history of several years of electric power measurements of the major part of their feeders (representing around 1000-10000 low voltage electricity customers), and some low-voltage lines (around 100-1000 low voltage electricity customers).

B. Objectives of the Simulator

To benefit from this huge amount of data as input to an electric load curve simulator, models must be able to consider both microscopic and macroscopic types of data as input; in this sense we denote that a model has to be “holistic”. The data has to be accessible and understandable by every actor linked to the electricity grid (district manager, DSO, electric consumer...). When it comes to the microscopic data where information can be available at the level of a client or a device, each contributing element (consumption device) of the electric load curve must be

simulator we propose in this paper is to be able to simulate with very short computation time (i.e. a few tens of seconds for a year at one hour time step resolution) the individual electric consumption of thousands of customers (both in residential and tertiary sectors). This involves simulating whole cities. At the same time we should be able to modify inputs and perform sensitivity analysis quite rapidly. The simulation has to be done with a variable time step for a maximum flexibility to deal with different actors, and with possibilities of access to large databases for the input data. Microscopic inputs have to be considered, like local specificities on climate or buildings.

The simulator must include a physical modelling of the different electric consumption devices and buildings, gathering both efficient calculation time, physical sense, and enabling the development of an automatic, fast and adaptive statistical method in big dimensions to benefit from the different databases.

To answer these objectives and requirements, we have proposed to split the study into four parts illustrated Fig.1. This figure is a summary of the view of our electric load curve simulator, each square block will be detailed in the following sections.

We have proposed a structure of a load curve simulator which is presented in detail in Section II. Section III shows how the outputs are evaluated, by comparison with a reference load curve. Section IV presents the way the different databases are matched in order to compare load curves with same inputs. Section V investigates possible answers to make an automatic calibration of models parameters. Finally in Section VI are presented the conclusions of this work.

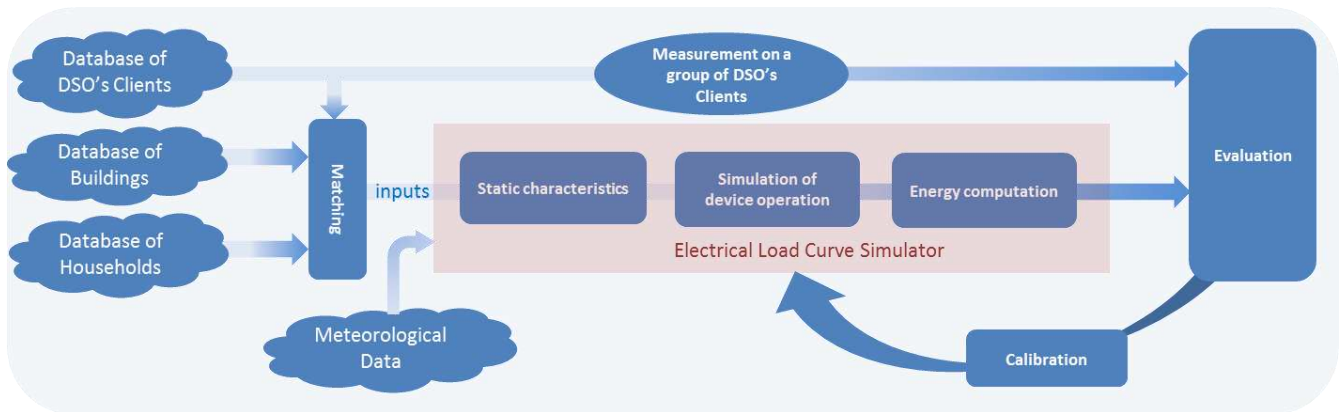


Fig.1. Illustration of the approach used for electric load curve simulation

identified spatially and its temporal profile should be known. The level of details of the physical models for each consumer has to be though the best compromise between representation of reality and impact on the simulation time when this model is considered within the overall load simulator. All electric consumers, from residential to tertiary and industries have to be taken into account to benefit from the DSO's electric power measurements. As models have many parameters, a calibration from the input data and literature results is necessary. Finally, the simulator and the involved models have to be as generic and open as possible to account for evolutions in the context and be able to integrate new models (i.e. for consumption devices etc).

The technical objective we have set for the load curve

II. STRUCTURE OF THE SIMULATOR

A. Introduction

The developed simulator implements the following models for devices of the residential sector: TV, washing machine, dishwasher, computer, light, swimming pool, dryer, oven, cooktop, refrigerator, freezer, water heater, electric heating, heat pump, air conditioner and miscellaneous (standby).

The consumption devices are modeled by elementary models such as block profiles of power. Figure 2 depicts the example of the model proposed for a washing machine: the first power block corresponds to heating the water with a resistance, the second block models the washing cycle with a

constant power of motor, and the third block models the spinning with a constant power of motor.

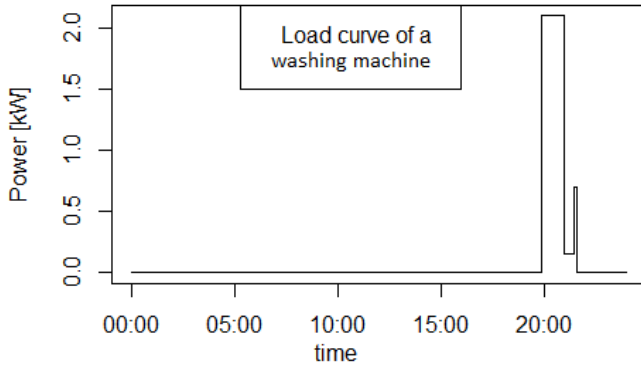


Fig.2. Example of load curve of a washing machine modeled by a constant block of power.

The simulator simulates the residential electrical load curve consumer by consumer. A consumer in the residential sector can be a house or an apartment, with usual occupation conditions or used as secondary residence.

For simulating residential load, the inputs listed below are required:

- the building characteristics such as its surface, height, age, thermal inertia and resistance, type (apartment, house...);
- the household characteristics: number of people living in the housing of the consumer, ownership rate of the different consumption devices;
- the devices characteristics (e.g. power of the washing machine).

The simulator also implements tertiary sector electricity consumers, with the categories: offices, teaching/research, shops, shops with food, restaurants, hotels and charging stations for electric vehicles (EV).

Some examples of the input needed to simulate these tertiary consumers are:

- the building characteristics such as the surface, height, age, thermal inertia and resistance;
- the consumer characteristics: e.g. number of clients of a hotel or a restaurant, ownership rate of the different consumption devices (e.g. presence of electric heating, air conditioner...);
- devices characteristics (e.g. installed power of light)

We can notice that the number of inputs of the developed simulator is very high rising up to around a thousand of parameters. The values of many of these inputs are set to default ones thanks to typical figures available at national level (e.g. ownership rate of consumption devices in residential sector), or in the literature (e.g. sizing of an electric heating with building assembly characteristics) and finally through measurements campaigns (e.g. through the REMODECE project results for the power of consumption devices).

Other inputs are set thanks to different existing databases for buildings characteristics (e.g. BDTOPO database from IGN institute in France), households (e.g. from INSEE institute in France) or electricity customers (e.g. from the French DSO ERDF).

B. Implementation of the Simulator

The technical objectives of simulator in mind, we have chosen an architecture with a computation core as fast and simple as possible implemented in a low level language: C++. This core allows to run rapidly physical models of consumption devices. A specificity is that the implemented models of consumption devices are the base of all the calculation, so fundamental models of power consumption of consumption devices (such as constant block of power seen previously) can be easily changed if needed.

In order to deal with big data, and to be able to make statistical computation, the computation core has been interfaced with R software, which is an open source numerical simulation platform.

C. Different parts of the simulator

The electrical load curve simulator is implemented into three blocks (static characteristics, simulation of device operation and energy computation), which are illustrated in Fig.1 and explained in the following sections.

1) Static Characteristics

This part of the simulator permits to populate the model of each consumption device with the necessary parameters. (i.e. power of the TV, repartition of duration (e.g. minimum and maximum duration for the daily use of a TV), etc. The values of the inputs are set as explained above either by default or by existing databases and measuring campaigns. Then, these consumption devices are grouped to simulate a consumer (e.g. a consumer has two TVs, a washing machine, some lights).

2) Simulation of Device Operation

At the beginning of this step, we consider that a consumer is equipped with consumption devices that can be launched during a chosen simulation period. To do this, a given profile of use for each consumption device is set (see example Fig.3). This profile is a combination of daily, weekly and monthly profiles, and is interpreted in the calculation core as a probability density for launching the given consumption device. A Monte Carlo simulation is performed where the number of sampling in the simulation period is chosen as an affine trend with number of inhabitants in the housing of the consumer. Let us remark that the start date and duration sampled in this launch phase are always independent from any time step: here we just memorize start dates, duration and power of operation (called "instance") of all consumption device generated in the previous step.

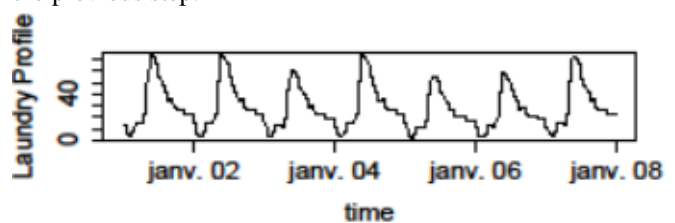


Fig.3. Example of density of use of a washing machine entered as input from R interface

We can note that consumption the devices, electric heating, heat pumps and air conditioners are simulated differently than what was previously described: they are linked to assembly thermal characteristics, outdoor

temperature, solar irradiation and temperature set point, modeled by a first order “RC” model.

3) Computation of Energy of each Consumer on a chosen Temporal Grid

This step is the moment when the temporal resolution is chosen: a temporal grid is created, with a given time step, and all the “instances” previously described are integrated in the chosen temporal grid (see Fig.4), making a final vector of energy at the chosen temporal resolution. The energy is simply calculated by multiplying the power time the duration of this power, at each time step. This specific approach permits to change time step without making again all the previous steps of simulation, as all the instances are memorized.

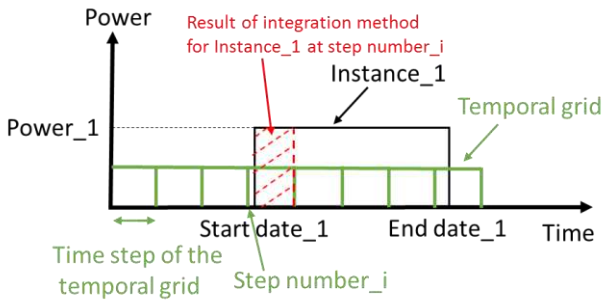


Fig.4. Illustration of the way the evaluation of energy is done in the simulator

III. EVALUATION OF RESULTS

For a specific test case the simulation is done in a bottom-up way, consumption device per consumption device, housing per housing, and tertiary consumer per tertiary consumer. This leads to a considerable amount of output results which are memorized and can be evaluated by a comparison with a reference case that can come from a group of DSO’s customers as illustrated Fig.1.

A. Partial evaluation on specific models

To our knowledge, often in the literature in electrical load curve simulation using physical bottom-up modeling, the evaluation of the results is performed on some aspects of the simulated load curves, which are compared to a limited number of measured load curves in specific cases and locations, always in the residential sector [2].

This kind of evaluation is useful for us to check some specific models and for setting by default the values of some inputs.

For example the annual energy consumed per consumption device was measured at large scale thanks in the frame of the REMODECE project (2008). This allows us to set by default the number of launch of a consumption device in order to get the same trend of yearly consumption by the same park of devices as the campaign measurement.

Another example is the measurement of hourly mean load curves per consumption device, and the comparison with simulated one, as depicted Fig.5 and Fig.6.

B. Methodology for the evaluation on the whole load curves

When results of simulation are obtained, they have to be evaluated. Here are investigated evaluation methodologies.

Some authors like [3] have developed what [2] calls “statistical adjusted engineering models”. They use automatic measured load curve learning, but it is only

relevant for short term forecasting as this learning is to reproduce the past power demand profiles.

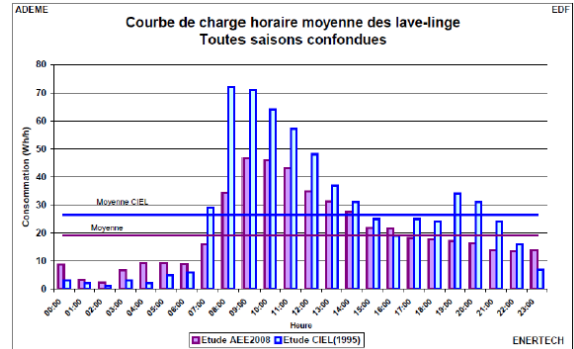


Fig.5. Example of a mean hourly load curve of washing machines obtained by measurements in hundreds of houses [5]

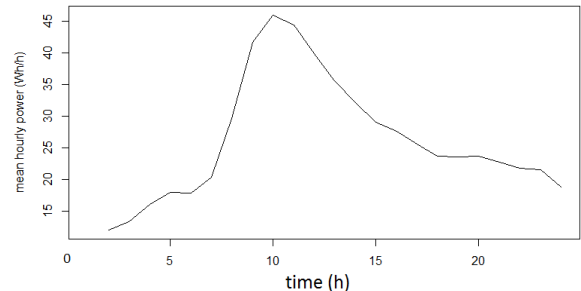


Fig.6. Example of a mean hourly load curve obtained by simulations of 1000 laundries for one year.

To our knowledge, the evaluation with large DSO’s load curve databases is new.

When we want to compare simulated and real load curves it leads to a problem in itself, which is tackled in the next section. Also authors have focused on residential sector modeling and DSO’s load curve measurements are the aggregation of residential, tertiary and industrial electricity consumption. The latter is less frequent as usually industrial customers are grouped in industrial zones, with their own electricity feeder. But at the scale of medium voltage feeders, the load curve is always a mix of residential and tertiary consumption.

Reference [4] proposes in its last part a review of the different mathematical tools used by authors to compare electrical load curves.

Some of them are statistical indicators such as mean, minimum, maximum values or root mean square values. Some of these indicators such as the mean value can be useful to determine for example if the standby model is well calibrated.

Other indicators take into account the temporal dimension in the load curves, adding the phase difference aspect in the comparison. The aspect of the peak correlation and gap is very important for DSOs which sizes its electric grid thanks to the peak value.

An interesting method developed in [4] consists in taking into account both amplitude and temporal differences, using the Dynamic Time Wrapping; it can be a useful tool to investigate the models parameters influence in the distance between the real and the simulated load curve.

We can also compare the frequency spectrum of two load curves depicted in Fig.10 in their temporal aspect. Fig.7 and Fig.8 are the plot of their frequency spectrum, and it can be useful to investigate if the different temporal repetition are present: for instance the daily shape (24h) of the load curves

leads to a peak more important in the simulated curve than in the measured one, whereas it is the opposite for the half daily shape (12h).

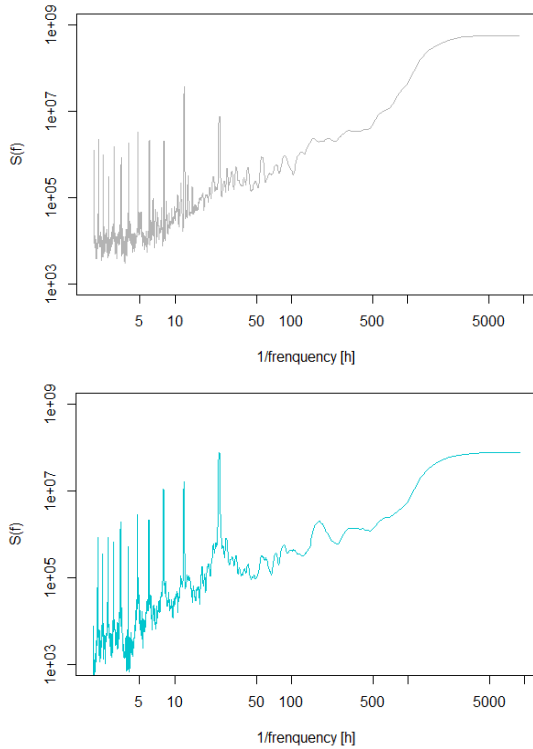


Fig.7. Spectrum of a load curve of a medium voltage feeder (upper grey curve) and the simulated load curve (lower turquoise curve), for a whole year.

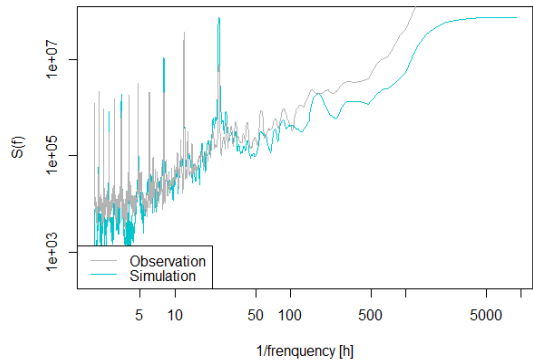


Fig.8. Spectrum of a load curve of a medium voltage feeder (grey curve) and the simulated load curve (turquoise curve), for a whole year.

Finally, comparison of sensibility with inputs parameters, such as the thermo-sensitivity can be compared, and are investigated in a further section.

IV. THE MATCHING PROBLEM

A. Introduction

To assess comparable load curves, we must have the same characteristics between the electricity customers that effectively consume power in the measured load curve and the simulated load curves. In general, the DSO knows the list of his customers that are connected to a feeder; it has for each customer their characteristics listed in Table 1. However, with this information we cannot fill out the inputs for the electric load curve simulator and we have to use others databases.

Indeed much useful information that can be converted into input in our simulator is found in various existing databases. In Table I we provide information on the main databases used in our electrical load simulator with their origin, the entities they describe as well as the characteristics of these entities.

TABLE I
LIST OF THE DIFFERENT DATABASES USED WITH THEIR ORIGIN, ENTITY AND CORRESPONDING ENTITY CHARACTERISTICS

Origin of the database	Database entity	Useful characteristics of the entities for the simulator
INSEE institute (France)	Housing	number of inhabitants, type (flat, house), surface given by intervals, number of rooms, age of the building, category (vacant, secondary residence, principal residence), type of heating (electric, fuel, gas, wood, other), IRIS ¹ mesh.
IGN institute (France)	Building	Text address, geometry of the building (polygon size).
ERDF company (France)	Electricity customers	Subscribed power, annual consumption, text address, IRIS ¹ mesh, type of customer (e.g. residential, tertiary artisan or shopkeeper).

Below are given the definitions of the different entities in order to use precise vocabulary:

-Housing: houses or dwelling places thought of as a group. A housing occupied as a principal residence has a household.

-Building: anything built on an area of land, having a roof and walls and intended to be kept at a well-identified location.

-Electricity customer (or customer): materialize a unique contract with the DSO who is paid in order to provide an electricity supply service at a point of common coupling where the customer belongs.

-Electricity consumer (or consumer): entity which is used in the electric load curve simulator that tries to benefit from all databases in order to simulate electrical load curve as accurate as possible from the chosen inputs.

The different databases provide complementary information. Their matching is important in order to fill in the simulator with the maximum of available input data. .

Two cases are investigated: the case where there are no common characteristics (third column of Table I) from a database to another and the case where there is at least a common characteristic.

B. Case where there are no common characteristics between two databases.

We will deal with this case with a concrete example: the matching between the housing database and the residential electricity customer database. Indeed, the final objective is

¹ The IRIS mesh is a mesh lower than can be lower than a city; it was defined by INSEE Institute; each INSEE mesh has a unique six figures identifier.

to compare a measured load curve of a feeder and a simulated load curve with the same housing characteristics as these of the residential electricity customers connected to the feeder. To do this, each housing can be matched to electric customers, and reciprocally.

We can define a criterion to assess in what extend the two different lists of entities are correctly matched.

Simply, this is done by the definition of a certain distance between comparable parameters linked to the entities.

In our work we propose to express with a model the characteristics of an entity. This models has as parameters the characteristics of another entity. A first case study considered in our work concerns modelling the annual electricity consumption based on some specific housing characteristics.

Figure 9 shows the obtained results using data coming from ERDF's and INSEE's databases for the French Brittany region and for year 2011.

In this case we know the list of housings and electricity customers per IRIS mesh, and we use this information to investigate for example the behavior of the total residential electricity customer annual consumption per municipality versus the total surface of housing per municipality. The fit with a linear model has a mean error of 11.4%. If we add another explicative variable such as the surface heated by electric heaters and the age of the building, we can decrease the mean error of a linear model down to 7%.

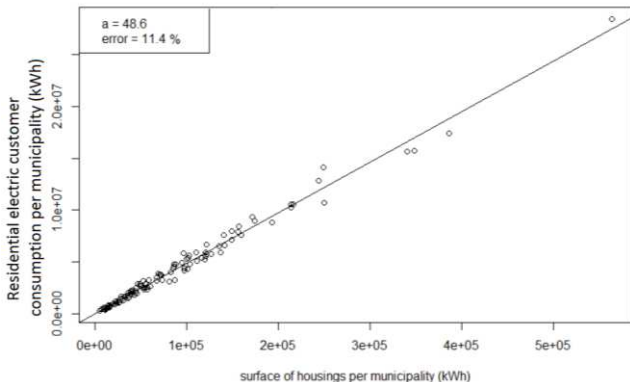


Fig.9. Fit of the residential electric customer annual consumption per municipality versus the surface of housings per municipality.

C. Case where there exist common characteristics between two databases.

The approach here is to match the different entities among them, by minimizing the global distance of the same characteristics between the matched entities.

Like in the previous section, we will present some results based on specific cases, and namely the same example as in the previous section, to illustrate the matching problem and its challenges.

As a problem set, we want to match the housing database (INSEE database) and the residential electricity customer database (ERDF database). We consider that the annual electricity consumption is known for the two different entities. To simplify the problem we consider that there are more housings than electricity customers, thus, an electric customer can gather more than a housing. However, a housing must be linked to an electricity customer (we consider that all the housings have a contract with a DSO for electricity supply, which is very close to the reality in France).

We define below the list of parameters to be matched from the two considered databases:

$x_{i \in \{1,2,\dots,n\}}$ represents the estimated electricity annual consumption of the housings, n is the number of housings in the considered IRIS mesh. x_i is the i^{th} housing.

$y_{j \in \{1,2,\dots,p\}}$: represents the residential annual consumption of the electricity customers. p is the number of customers in the considered IRIS mesh. y_j is the j^{th} electricity customer.

The problem of matching can be formulated as a distance (here chosen to be quadratic) minimization problem between the matched entities:

$$(\alpha_{ij})_{(i,j) \in \{1,2,\dots,n\} \times \{1,2,\dots,p\}} \min [\sum_j^p (\sum_i^n \alpha_{ij} x_i - y_j)^2] \quad (1)$$

S.t.:

$$\alpha_{ij} \in \{0,1\}, \forall (i,j) \in [1:n] \times [1:p] \quad (2)$$

$$\sum_j^n \alpha_{ij} \geq 1 \forall i \in [1:n] \quad (3)$$

$$\sum_i^n \alpha_{ij} = 1 \forall j \in [1:p] \quad (4)$$

where:

- $\alpha_{ij \in \{1,2,\dots,n\} \times \{1,2,\dots,p\}}$ is the matching unknown : $\alpha_{ij} = 1$ if x_i is matched with y_j , 0 else (2);
- (3) is the constraint that each housing is linked to at least an electricity customer;
- (4) is the constraint that each electricity customer is linked to a unique housing.

A simplified resolution algorithm of this optimization problem was tested on some data in order to obtain first results depicted in Fig.10. This Figure shows that the two curves present similar characteristics especially in term of peak correlation, which is a promising result. We can remark especially the model of standby that has to be calibrated in order to have similar mean consumption, and that the peak numbered 1 in the figure in the measurement is certainly due to reorganization, meaning that temporarily another feeder was connected to the measured feeder, for maintenance reasons.

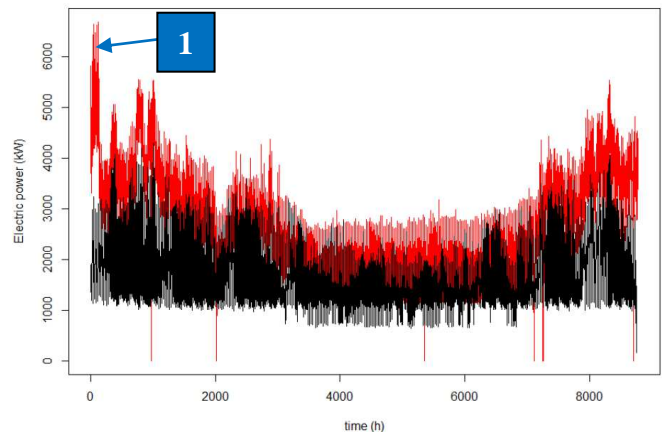


Fig.10. Measured load curve of a medium voltage feeder (red plot) and the simulated load curve (black plot), for a whole year, using as input the housing characteristics of the results of the simplified matching problem.

V. AUTOMATIC CALIBRATION

The automatic calibration step, depicted in Fig. 1, is based on a comparison between the simulation results simulation and reference values for the specific inputs chosen. This comparison permits to calibrate or re-adjust the different parameters of the simulator in order to minimize the gap of its output and the reality.

A. Case study on Thermo-sensitivity

We present here an example of what can be an automatic calibration using the DSO's load curve. We consider as parameter the thermal resistance of buildings and we simulate the repartition of this parameter in the buildings of an area. The way this repartition is made, represents a level of "thermo-sensitivity" for the ensemble of buildings. Then using the thermo-sensitivity of the observed load curves we can calibrate the model.

Many definitions exist for thermo-sensitivity. We chose here a simple definition which is the linear fit of consumption points in ordinate and temperature points on the abscissa, by only selecting points whose temperature are below a certain threshold of non-heating; we take in this example 15°C. We compare the thermo-sensitivity of the simulated and measured load curves represented in Figure 10. The results are given in Figures 11 and 12. The coefficients a and b are respectively the linear and intercept coefficients of the fit, R^2 is the coefficient of determination.

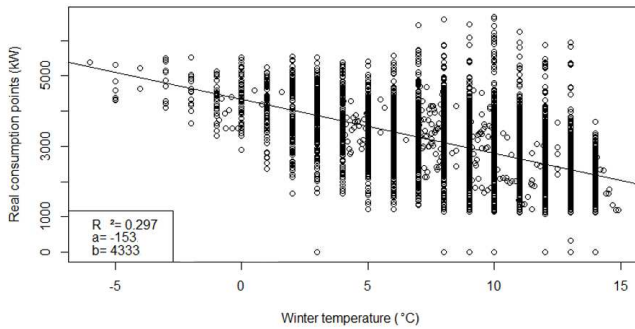


Fig.11. Thermo-sensitivity of the measured load curve of Fig. 10 where the temperature is the temperature of the cities of the customers weighted with the number of customers in the cities.

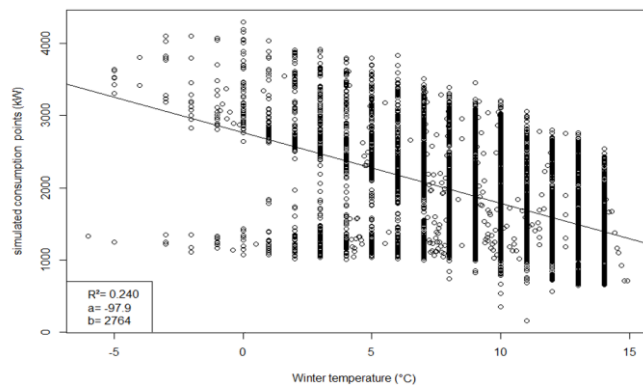


Fig.12. Thermo-sensitivity of the simulated load curve of Fig. 10 by taking a medium level of thermal resistance per buildings, where the temperature is the temperature of the cities of the customers weighted with the number of customers in the cities.

For the thermal model of buildings we chose a thermal resistance which is in a very simple model proportional to the surface of the consumer. We can look for the behavior of the different parameters a , b and R^2 when the surface thermal resistance or the target temperature are modified. We can

then plot the sum of the relative errors between a , b and R^2 simulated and from the measurements, function of the surface thermal resistance and the target temperature to find automatically the couple of inputs that minimizes the errors, as illustrated in Fig.13.

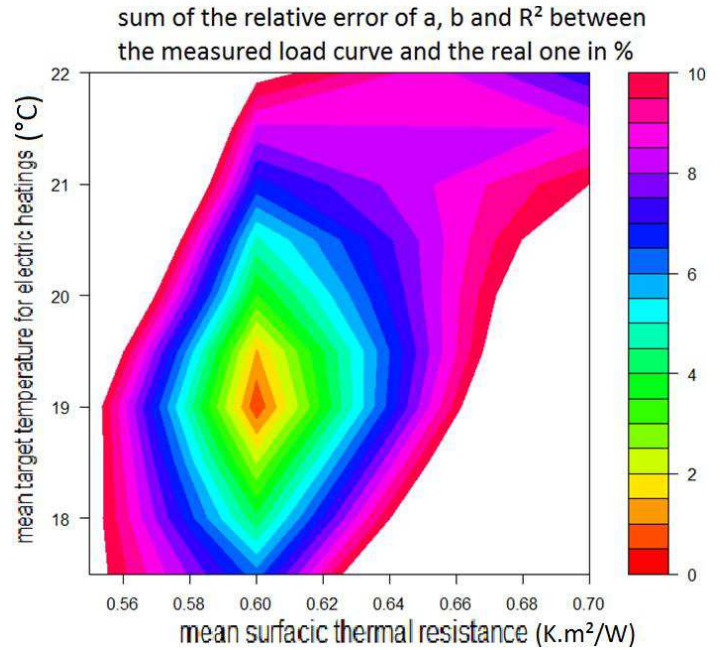


Fig.13. Sum of the relative errors between a , b and R^2 from simulated outputs of simulator and from the measurements, as a function of the surface thermal resistance and the target temperature.

VI. CONCLUSION

With current challenges, available data and simulation tools, we have shown that electrical load curve modelling is more than the objective of prediction of demand, to be an accessible, understandable and accurate tool for every actor around the electricity grid. To answer the previous objectives, we have presented in this paper an electric load curve simulator which aims to be as simple and generic as possible to be able to deal with computation time and big data available.

A novel kind of validation is proposed, which tries to take into account results in the available literature and DSO's power measurements. The paper has presented the approach proposed for matching different databases. The qualitative and quantitative results presented are promising. To benefit from these big databases, future work will focus on the automatic calibration of parameters of the simulator.

VII. ACKNOWLEDGMENT

ERDF is acknowledged for supporting the study and providing the data for this work.

VIII. REFERENCES

- [1] L. G. Swan, V. I. Ugursal. "Modeling of end-use energy consumption in the residential sector: a review of modeling techniques". *Renewable and Sustainable Energy Reviews*. 13(8) pp. 1819–35. 2009.
- [2] A. Grandjean, J. Adnot, G. Binet. "A review and an analysis of the residential electric load curve models". *Renewable and Sustainable Energy Reviews*. 16 pp. 6539–6565. 2012.
- [3] K. Train. "An assessment of the accuracy of statistically adjusted engineering (SAE) models of end-use load curves". *Energy*. 17(7) pp. 713–23. 1992.

- [4] A. Grandjean. "Introduction de non linéarités et non stationnarités dans les modèles de représentation de la demande électrique résidentielle". Ph. D. Thesis. January 2013.
- [5] Ademe, EDF, Enertech. "Campagne de mesures des appareils de production de froid et des appareils de lavage dans 100 logements". AEE2008 Project. Final report. April 2008.

IX. BIOGRAPHIES

Thibaut Barbier was born in Grasse in France in 1991. He graduated from Ecole Normale Supérieure de Cachan in both mechanical engineering, electronics and telecommunications (2012); he also graduated from Ecole Polytechnique, in the renewable energy sciences and technology bachelor program (2014). He passed in 2013 the Agrégation in Sciences Industrielles pour l'ingénieur option ingénierie électrique (French high school teacher competitive exam in electric engineering).

He has worked in 2014 for 6 months as a research internship in electric load curve modeling in PERSEE research Centre in MINES ParisTech and is since September 2014 a Ph.D student in PERSEE in electricity consumption modelling.

Robin Girard is a senior researcher at MINES ParisTech working on the integration of renewable energies into the electric system. He has a PhD in Applied Mathematics and high dimensional statistics from Université Josph Fourier (Grenoble, 2008). He has worked in research projects and has published scientific papers on several subject including forecasting, electric system planning, electricity market, and smart grids.

François-Pascal Neirac was born in 1956. He is graduated from Ecole Centrale (1979) and PhD of Mines ParisTech. He is today professor at Mines ParisTech and researcher at PERSEE. He is the director of the Post-Master programme "International Energy Management", delivering a double degree diploma with Tsinghua University in Beijing. His research activities deal with the analysis and the optimization of complex energy systems. He has coordinated many research projects. His recent activities focus on the modelling and the economic analysis of energy systems including distribution electricity grids, storage technologies, and distributed generation.

Nicolas Kong was born in Cambodia, in 1956. He is senior Consultant in the Technical Management of ERDF, Network Development Department. He graduated engineer from ENSMA (Ecole Nationale Supérieure de Mécanique et d'Aérotechnique) in 1981 and work for EDF since 1982. He is currently working on the knowledge, the load forecast of power electric system and the impact of new uses and decentralized production on the French network.

Georges Kariniotakis was born in Athens, Greece. He received his Eng. and M.Sc. degrees from Greece in 1990 and 1992 respectively, and his Ph.D. degree from Ecole des Mines de Paris in 1996. He is currently with the Centre PERSEE of MINES ParisTech as a senior scientist and head of the Renewable Energies and Smartgrids Group. He has been involved as participant or coordinator in more than 40 R&D projects in the fields of renewable energies and distributed generation. His scientific interests include among others timeseries forecasting, decision making under uncertainty, modelling, management and planning of power systems. He has authored more than 180 scientific publications in journals and conferences. He is Senior IEEE member and member of several groups of experts like the European Wind Energy Technology Platform (TPWind).