Features of residential energy consumption: Evidence from France using an innovative multilevel modelling approach

Fateh Belaïd
Paris-Est University
Scientific and Technical Center for Building
Email: fateh.belaid@cstb.fr /fateh.belaid@univ-littoral.fr

David Roubaud
Montpellier Business School (MBS)
Montpellier Research in Management (MRM)
Email: d.roubaud@montpellier-bs.com

Emilios Galariotis
Audencia Business School, France
Email: egalariotis@audencia.com

Abstract

Recent efforts to reduce residential energy consumption have renewed interest in investigating salient drivers of household energy use. This article contributes to this ongoing literature by developing a model to examine geographic effects on energy use. Using a new, rich, micro-level survey that combines various information about dwelling attributes, occupant characteristics, and behaviors, we suggest a combined a bottom-up and top-down statistical approach based on a multilevel regression model (MRM) and an innovative variable selection approach via the Adaptive Elastic Net Regularization technique (AdaEnetR). This provides the ability to extract geographic effects from the total variation in residential energy consumption as well as explain simultaneously the remaining variation with relevant explanatory variables and their interactions. The current model addresses several interrelated issues posed by the use of econometric methods to examine residential energy demand, including the risk of aggregation/disaggregation bias. Our empirical findings demonstrate MRM's ability to quantify effectively approximately 0.67% of geographic effects (aggregate level) and approximately 0.31% of household and dwelling effects (individual level). Further, the findings show that household attributes are important factors that influence residential energy consumption patterns.

Keywords: Multilevel regression; Residential energy consumption; Geographic effects; Energy efficiency
JEL Classifications: C01,C20, D12, Q40, Q50.

1. Introduction

Energy efficiency is one of the most important aspects of the European Union's low-carbon economy roadmap to reduce CO_2 emissions by 80% to below 1990 levels by 2050 and 40% by 2030. In addition, in response to the excessive fluctuations in energy prices in recent years, considerable attention has been paid to managing energy demand by improving efficiency. This paper focused on residential energy consumption and efficiency in France. The French residential sector consumes 30% of the total energy consumption, more than does the industrial sector and nearly a similar amount to that of the transportation sector. Moreover, as the residential sector also contributes to more than 20% of national CO_2 emissions, it is an important target for energy efficiency incentive measures to mitigate greenhouse gas emissions and address concerns about degradation of environmental quality. Because approximately 60% of existing French residential dwellings were built before 1975, when there were no energy regulation standards, and the renewal of the existing stock is very low (up to 1%), this particular group of housing units is the primary target of housing refurbishment programs.

Electricity and gas constitute the two main sources of energy households in France consume, and electricity used for space heating represents more than 60% of household energy consumption. In recent years, energy consumed for space heating has decreased the most significantly throughout France, by 33% since 1990, but at the same time, the amount of energy consumed for electrical appliances specifically has increased by 40% because of the use of numerous new home electric devices (Smartphones, tooth brushes, etc.). To pursue the global efforts to reduce household energy consumption and adopt incentive measures to do so, a recent policy framework has set the objective to reduce the amount of residential energy consumption by 40% by 2030[1].

Obviously, a thorough understanding of salient factors that affect domestic energy consumption is a major requirement for any residential energy policy scheme. Accordingly, studies and research works on causes of residential energy demand have become popular in recent years, both in policy and academia [Zheng et al. (2014), Hu et al. (2017) and references therein]. In this context, studying national energy consumption surveys that provide detailed information on households' characteristics and their effects on energy consumption has proven to be particularly relevant. Recent efforts to improve energy performance of existing dwelling stock have revived interest in household energy demand (Geng et al., 2017), and patterns and factors that shape household energy consumption have been the subject of intense debate, both in academia and policy (Belaïd, 2017; Schulte and

[1] Loi Relative à la Transition Energétique (2015).

Heindl, 2017). The first series of studies of household energy use/demand began in the late 1970s, following the first oil crisis (Mazur and Rosa, 1974; Sonderegger, 1978; Dubin and McFadden, 1984). The goals of subsequent research on the drivers of residential energy consumption has varied (demand estimation, prediction, driver examination, etc.), as have the methods used (top-down, bottom-up, etc.). In addition, concern on the part of various fields has resulted in an increasing amount of research designed to develop innovative energy modeling approaches and include insights from other fields (Belaïd, 2017). However, the approaches used to examine socioeconomic and housing influences in most studies of household energy use and its various determinants essentially have been limited to multiple linear regression models (Tso and Guan, 2014).

Compared to the literature available on residential energy use, our paper hypothesized that geographic factors join many other factors (household features, housing attributes, etc.) and play a substantial role in influencing residential energy use and energy efficiency. This theoretical assumption relies on the idea that, in addition to individual and socioeconomic factors, a wide range of structural factors can influence residential energy demand, including global economic welfare, energy sources available, climate conditions, and cultural habits. Moreover, the goal of our paper is to contribute to this nascent strand of research by examining the predominant factors in household energy consumption in France using a multilevel regression model (MRM) and an innovative variable selection approach via Adaptive Elastic Net Selection (AENS). Such a study will provide a better understanding of household energy consumption (Robinson, 2009) and help model and predict it with reference to relevant household characteristics.

Further, unlike classical multiple regression models, multilevel regression modelling has numerous advantages. A hierarchical regression model provides a framework to examine sources of variation within consumer groups, and factors that predict cluster-level variances can be identified by considering the clustered nature of the dataset (Carle, 2009). Obviously, working simultaneously with two hierarchical levels in a stratified dataset mitigates the risk of aggregation/disaggregation bias (or atomist error bias). This bias, known commonly as the Robinson effect (Robinson, 2009), is attributable to the fact that inferences derived from one hierarchy are not applicable to another. Nevertheless, despite the valuable contributions MRM may offer to improve residential energy consumption modelling, it has not been employed in the energy literature.

Thus, this research makes a four-fold contribution. First, it introduces a new dimension to explore the spectrum of residential energy consumption in which the energy debate is associated with household environment. Second, while many studies have investigated residential energy demand, their empirical treatments are limited in important ways. The hierarchical model we used in this study helps extract geographic and environmental effects from total energy consumption variations and highlights the residual variations with the interactions of multiple factors. Third, to our knowledge, empirical research on this issue in France is rather limited because of the lack of information and availability of disaggregated data on household energy use, and any new empirical investigation is welcome. Finally, the innovative variable selection method we employed in this study allowed us not only to select relevant factors that make the model easy to interpret, but to enhance the accuracy and stability of the predictors and avoid the so-called curse of dimensionality.

In summary, to further our understanding of household residential energy demand, this study addressed two sets of research questions. First, what are the features of residential energy demand in France and what are its predominant drivers? Second, how does the geographic and environmental context influence residential energy consumption?

Our empirical results showed MRM's ability to quantify effectively approximately 0.67% of geographic effects (aggregate level) and approximately 0.31% of household and dwelling effects (individual level).

Proportion of the overall explained variance proportion is about 76% compared to 62% using OLS model. In addition, the findings highlight that household and housing attributes are important factors that influence residential energy consumption patterns.

The remainder of this paper includes the following: Section 2 describes our research hypothesis and introduces the data and the modeling approach. Section 3 provides our empirical findings, and Section 4 draws conclusions and offers policy implications based on the empirical results of the model.

2. Data and Methodology

2.1. Research hypotheses

Motivated by a desire to increase energy security and reduce CO_2 emissions, research and energy policy discussions have focused increasingly on improving energy efficiency. In recent years, the adverse environmental effects of residential energy demand have led to significant concern about energy-saving behavior and household preferences for energy-efficient measures.

This study hypothesized that regional attributes have a significant effect on domestic energy use. Thus, our theoretical model integrated regional effects on households' energy demand overall using a new approach that creates a combined top-down and bottom-up model to reveal the complexities of residential energy consumption.

Based on the previous findings from the literature, the following research hypotheses were proposed:

H1: Environmental factors have a significant effect on household energy consumption, including heating degree days (HDD), regional energy prices; regional income; unemployment, and poverty rate.

H2: Household socioeconomic attributes play an important role in household energy consumption, including age of household responsible person (HRP); household size; income; employment status, and tenure type.

H3: Housing characteristics play a crucial role in determining residential energy consumption patterns. The principal control variables are: dwelling energy performance certificate; housing type; dwelling size, heating system, etc.

H4: Household behavior and lifestyle may influence household energy consumption. The primary explanatory variables used in our model were: heating temperature; appliance use rate; household preference between economy and comfort, and heating restriction.

The MR model developed in this study is a combined bottom-up and top-down model, and the study included the different dimensions cited above to shed light on the various factors in residential energy consumption.

2.2. Data

The data used in this study were drawn from the recent household energy consumption survey (PHEBUS)[2]. This is a new national survey implemented by INSEE (the French National Institute of Statistics and Economic Studies) that gathers energy-related information for the principal housing units occupied in France. PHEBUS includes two separate sections: in-person interviews with housing occupants selected randomly, with questions regarding their household equipment energy consumption, global energy consumption, and attitudes about energy-saving, and another section that comprises a diagnosis of the housing unit's energy efficiency. The main objective of PHEBUS is to provide a clear picture of household energy use within French metropolitan housing stock in 2013 (the most recent data available). The PHBUS dataset consists of observations taken from 2,356 housing units selected to represent the 27.6 million housing units that are occupied as a primary residence. The survey includes only housing units that correspond to an individual house, those located inside buildings, independent rooms inside buildings with a private entrance, and homes for the elderly. PHEBUS was conducted from April to October 2012, across 96 departments (DEP) and 12 regions (REG) within French metropolitan areas. The data were collected using an

[2] The database can be accessed freely from the Operation Manager of the survey: Service of Observations and Statistics (SOeS) under the direction of French Ministry of ecology, Sustainable Development and Ecology, subject to the prior agreement of the Committee on Statistical Confidentiality.

area-probability sampling scheme derived from national census data collection in France, and are representative of housing units across regions, climatic zones, housing type (insulated house or multi-unit housing) and year of housing construction.

2.3. Modelling approach

In this section, we specify the multilevel regression model (MRM) framework underlaying the analysis. To examine the salient variables in household energy demand, we developed a bottom-up statistical approach based on the MRM model and an innovative variable selection approach via the Adaptive Elastic Net Selection technique (AdaEnetR).

Fig. 1. illustrates the conceptual framework of our modelling approach, which involves four main steps.

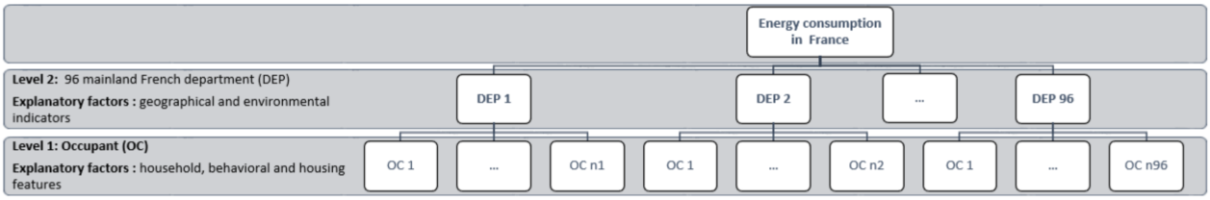

Fig. 1. The conceptual framework of our modelling approach

Step 1. Scaling sample weight: Following Carle (2009), we used two methods to assess the appropriate sample weights. In fact, PHEBUS includes unequal weights to ensure a sample that represents the population. However, it does not consider that the unequal sample weights in the standard hierarchical model may result in biased parameter estimates (Carle, 2009). The two design weights (A & B) were estimated as follows:

$$W_{ik}^{a}(Scaled\ weight\ A) = W_{ik}\left({n_k}/{\sum_i W_{ik}}\right) \quad (1)$$

$$W_{ik}^{b}(Scaled\ weight\ B) = W_{ik}\left({\sum_i W_{ik}}/{\sum_i W_{ik}^{2}}\right) \quad (2)$$

W_{ik} represents the unscaled design weight for person i in cluster k.

Step 2. Null model assessment: Prior to the analysis, it is necessary to determine whether hierarchical regression is required to quantify the grouping effects. In fact, multilevel datasets do not need a hierarchical modelling approach necessarily. Therefore, we must quantify the Interclass Correlation Coefficient (ICC), which is equivalent to a random effect ANOVA, and is defined as follows (Peugh, 2010):

$$ICC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (3)$$

where σ_{u0}^2 represents the between-group variance and σ_e^2 the within-group variance. In our case, the ICC can be interpreted as the degree of similarity of households within the same geographic cluster (DEP or REG).

Step 3. Building the level-2 model: Following the results of the previous step, the next entails adding the level-2 explanatory factors to explain the grouping effect; five explanatory factors were investigated in this study. To consider the regional effects, we used two geographic factors: department (DEP) and region (REG), both of which are categorical variables that indicate the identification number of the French territorial subdivision. The level-1 explanatory factors are reported in Table 1.

Table 1
List and description of level 2 explanatory variables

Department variables				
Variable	**Mean**	**Min**	**SD**	**Max**
Energy price	0.14	0.07	0.06	0.03
Regional income	39,918	22,686	6,821	73,428
Poverty rate	14.01	7.98	3.22	24.11
Heating degree days	2,490	1,286	386.24	3,153
Unemployment rate	9.50	5.70	1.74	13.90

Figs. 2 and **3** show the energy consumption across French regions and departments, which fluctuate across both. However, it is clear that the variation in energy consumption is more significant across departments. Those in the North and East largely are cold (e.g., Nord Pas-de-Calais and Strasbourg), while those in the South typically are warm (e.g., Nice and Marseille). We expected that households in colder climates need more energy for heating and less for cooling, and a recent publication of the French Department of Observations and Statistics (SOeS) stated that heating accounts for approximately 70% of French household energy demand.

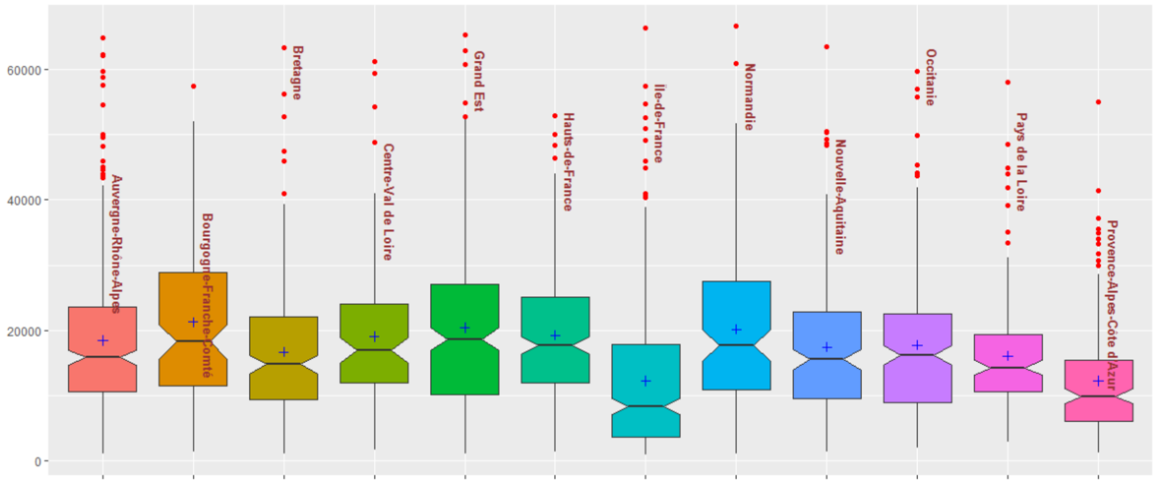

Fig. 2. Energy consumption in kWh by region

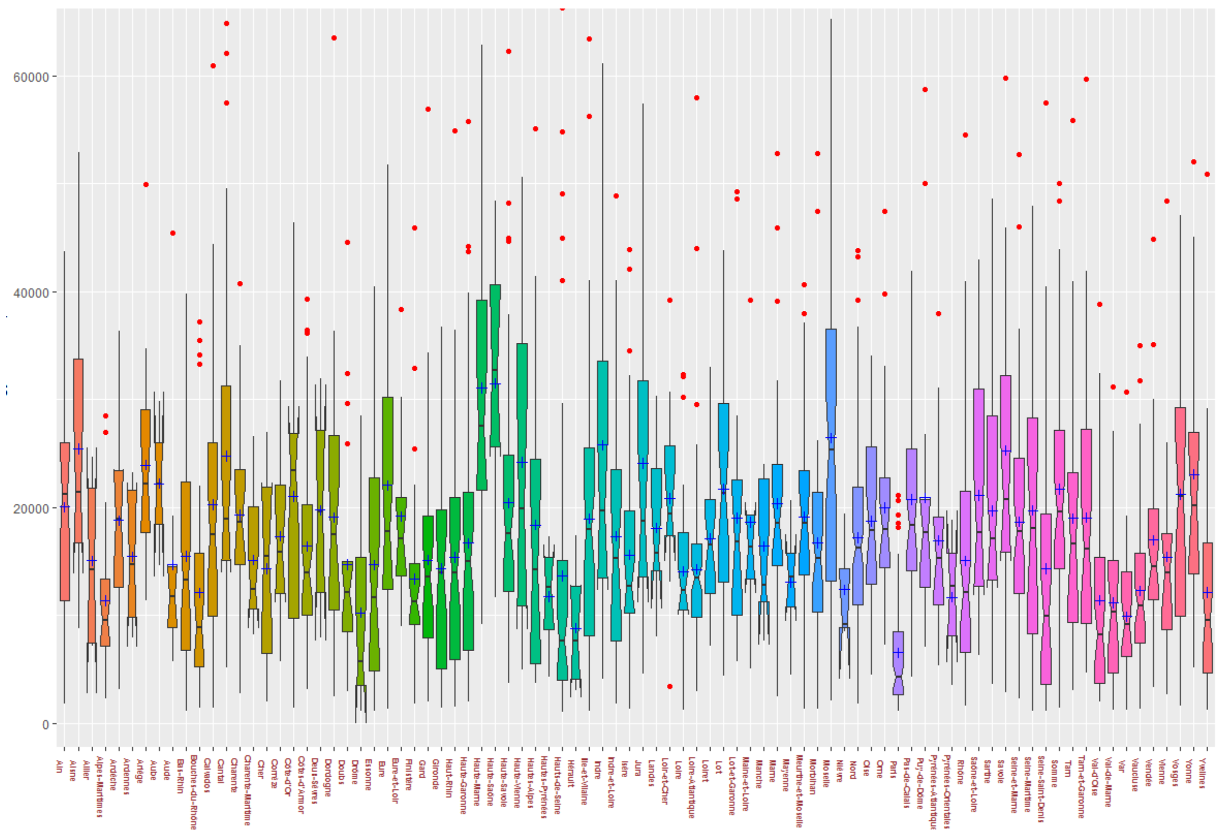

Fig. 3. Energy consumption in kWh by department

Step 4. Building the level-1 model: This step introduces the explanatory factors and examines their effects on the variation in household energy consumption. Based on the variables available in PHEBUS and the findings from previous literature, all possible household, behavior, and housing attributes were selected as candidates in the model. However, to enhance the accuracy and stability of the predictors and avoid dimensionality, we used

Adaptive Elastic Net, which is an innovative variable selection approach that combines the Adaptive Lasso and the advantages of quadratic regularization shrinkage (Zou and Zhang, 2009). Table 2 reports the level-1 explanatory variables.

Table 2
List and description of level-1 explanatory factors

Variables	**Categories**	**Mean/freq. (%)**
Energy consumption (dependent variable)		**17205.00**
Energy Price		**0.14**
Household attributes		
Income available (€)		39918.00
Seniority in the house		17.80
Number of household members (NHM)	1	22.40
	2	36.20
	3-4	33.20
	More than 5	8.20
HPR age	Less than 44	0.30
	From 45 to 66	0.50
	More than 67	0.20
Employment status (ES)	Top managerial profession	18.20
	Intermediate profession	22.80
	Employee	22.70
	Worker, routine and manual Occupations	26.10
	Other	10.20
Tenure type (TT)	Own	76.30
	Rent	23.70
Current situation	Working	56.90
	Unemployed	43.10
Housing characteristics		
Home size (m^2)		105.70
Housing temperature (^{0}C)		20.80
Appliance use ratio		1.80
Energy class (EPC)	A-B-C	15.90
	D	27.20
	E	29.80
	F-G	27.10

Housing type	Shared building	27.40
	Individual house	72.60
Climate zone	H1	59.90
	H2	33.90
	H3	6.20
Urban structure (US)	Rural commune	25.30
	From 2,000 to 20,000 inhabitants	18.50
	From 20,000 to 200,000 inhabitants	17.90
	More than 200,000 inhabitants	24.70
	Paris conurbation	13.60
Heating energy (HE)*	Electricity	34.60
	Gas	36.70
	Wood	6.00
	Fuel	15.80
	GPL	1.90
	Urban Heating	1.40
	Other energies	3.70
Heating system (HS)	Shared central	9.20
	Individual central	77.10
	Other	13.70
Renovation Works	Yes	50.20
	No	49.80
Behavioral features		
	No	18.10
Never reduce or turn off heating in winter	Yes	63.10
	No	36.90
Never reduce heating in the bedrooms except during holidays	No	68.20
	Yes	31.80
Heating restriction	Yes	21.90
	No	78.10
Duration windows open per day	More than 15 min	55.30
	Less than 15 min	30.90
	Occasionally	13.70
Heating management when house is unoccupied	Turn off the heater	27.50
	Reduce heating temperature	54.90
	Leave the heater at the same temperature	17.60

Household preference between economy and comfort	Comfort	58.40
	Economy	41.60
Hours house unoccupied per week	Less than 4H	58.10
	4-8 H	23.10
	More than 8H	18.80

Adaptive Elastic Net Selection

To assess the quality of an econometric model and improve its interpretability and performance, two aspects are crucial: (1) parsimony, particularly when analysing multi-dimensional data, and (2) prediction accuracy. Therefore, in the analysis of multi-dimensional data, it is important to decrease the number of factors and reduce the computation load and model complexity. To model energy demand and capture the relevant determinants of domestic energy use, it is necessary to use an approach that optimizes the model's accuracy and incorporates the complexity of the phenomenon. An important and still challenging issue known well in the field of modern statistical modelling lies in the variable selection process. It known well in today's statistical literature that, particularly in the presence of large numbers of predictors and multicollinearity, a simplistic and standard variable selection approach, such as step-wise, performs poorly both with respect to estimating coefficients and standard errors and selecting repressors, and leads to unstable subset selection that results in a model with poor prediction accuracy (Zou and Zhang, 2009). To overcome these shortcomings, several promising techniques have been proposed recently in the literature, including least absolute shrinkage and selection operator (LASSO), ridge regression, and Elastic Net (Liu and Li, 2017).

Because of its statistical and computational properties that enjoy oracle property, Lasso has become one of the most popular approaches in recent years (Tibshirani, 1996). However, Adaptive Elastic Net Regularisation (AdaEnetR), which is a new and innovative method, has several advantages that distinguish it from the Lasso approach (Liu and Li, 2017), as AdaEnetR is a convex combination of the ridge regression and the adaptive Lasso penalties (Zou and Zhang, 2009). These authors demonstrated that the Elastic Net can enhance the prediction accuracy of Lasso notably, particularly when the correlations among the independent variables are high. Algamal and Lee (2015) argued that AdaEnetR is more robust, also enjoys the so-called oracle property, and outperforms the other penalization procedures with respect to grouping effect, prediction accuracy, and factors selection. More recently, by using real data and numerical simulation, Hu et al. (2018) demonstrated that AdaEnetR is a competitive alternative method for high-dimensional model selection problems.

The regularized regression using the AdaEnetR of β is defined as follows (Li et al., 2011):

$$\hat{\beta}_{AdaEnetR} = \left(1 + \frac{\lambda_2}{n}\right) \arg\min_{\beta} \left\{ L_n(\beta) + \lambda_2 \sum_{j=1}^{Pn} \beta_j^2 + \lambda_1 \sum_{j=1}^{Pn} \widehat{\omega}_j \left|\beta_j\right| \right\} \qquad (4)$$

where λ_1 and λ_2 denote the regularization parameters, in which $\lambda_1 = 0$ leads the ADEN estimate back to the Ridge regularization, and $\lambda_2 = 0$ leads to the Adaptive Lasso estimate. $X_j = (X_{1j}, \dots, X_{nj})^T, j = 1, \dots, p_n$ are the linearly independent response variables.

The variables selection method proposed not only selects the relevant factors to make the model simple to interpret, but also enhances the predictors' accuracy and stability and avoids dimensionality.

Multilevel Regression Model

To improve our understanding of residential energy consumption determinants, we suggested a multilevel (or hierarchical) regression model, often known as a mixed-effects or random-effects model. By considering the clustered nature of the PHEBUS survey, the hierarchical model suggested offers a new approach to improve understanding of the household and environmental influences on residential energy consumption.

Multilevel models are particularly appropriate when analysing data with complex structures that involve stratified characteristic levels that form a combination of micro- and macro-unity, for example, households and their contextual environments or geographic location, which were defined as DEP and REG in this study.

Multilevel regression models have become a standard econometric tool used to examine unobserved heterogeneity in relations between factors that are measured for individuals clustered within higher-order units (Muthén and Asparouhov, 2009). This heterogeneity is depicted by continuous latent variables that vary between clusters, i.e., are expressed as random slopes and intercepts. Compared to a more traditional regression model, the principal advantage of hierarchical regression models is that regional effects are extracted from the variance of residential energy consumption, and thus allow the remaining variance to be explained with usual explanatory variables. In statistical models such as multiple regression models, there always is an unobserved component, in which the model does not explain an aspect of reality. In a multilevel model, dissociating different levels of observation allows this unobserved heterogeneity to be detected precisely and provides a measure of variance per level.

Assuming that we have K predictor factors X at the lowest level, and L predictors Z at the highest level, following Hox et al.'s (2010) specification, our model may be expressed contextually using the following equation:

$$Y_{ij} = \gamma_{00} + \sum_{k} \beta_{k0} . X_{kij} + \sum_{l} \beta_{0q} . Z_{lj} + \sum_{k} u_{kj} . X_{kij} + u_{0j} + e_{ij} \quad (3)$$

In Eq. (3), Y_{ij} is the annual energy consumption of household i in geographic division j (DEP). X_{ij} is the matrix of the level-1 explanatory variables and Z_j is the matrix of the level-2 explanatory variables. Other parameters in the equation need to be estimated. β_{00} is the intercept for the fixed effects,

3. Results and Discussion

This research sought to address some of the many claims that have been made about the role of various factors that affect residential energy consumption, and many conclusions can be drawn from the empirical results displayed in Table 5.

Null model assessment results

The first stage of analysis estimated the appropriate sample weights following Carle's approach (2009). Then, we calculated the ICCs to determine whether hierarchical regression was needed to quantify the grouping effects. The fit statistics of the null models are shown in Table 3. The ICCs for the null model with department as the location variable were greater than 10% (ICC=15% for the null model with scaled weight A). In the model with scaled weight B, ICC=20%, i.e., regional effects, explained approximately 20% of the variation in French residential energy consumption. The results confirmed the need to use multilevel regression to capture the regional effects. These findings were consistent with the results of previous research on the influence of geographic attributes on household energy use (Tso and Guan, 2014).

Table 3
Comparison of fit statistics of the null model results.

	Null model DEP		**Null model REG**	
	Weight A	**Weight B**	**Weight A**	**Weight B**
AIC	5864.20	5838.10	5968.70	5940.00
BIC	5881.40	5855.30	5986.00	5957.20
σ_{u0}^2	0.11	0.11	0.05	0.05
σ_e^2	0.62	0.48	0.68	0.53
ICC $\sigma_{u0}^2/(\sigma_{u0}^2+\sigma_e^2)$	**15%**	**20%**	**7%**	**11%**

Model fit statistics

The goodness-of-fit in multilevel regression models is assessed by the proportions of variance explained within the corresponding levels and expressed by each level's R^2. Table 4 reports the results, according to which we concluded that the MRM proposed outperforms the standard OLS regression.

Table 4
Comparison of fit statistics of OLS and multilevel regression.

Fit Statistics	OLS	REML			
	Scaled Weight A	Scaled Weight A		Scaled Weight B	
	Full model	Null model	Full model	Null model	Full model
AIC	3062	5864.20	2773	5855.30	2783
R^2	0.70	0.05	0.70	0.27	0.76

The variance proportions explained are estimated by the following equations:

Variance explained by level 1:

$$R_{L1}^2 = [1 - \hat{\sigma}_e(full\ model)/\hat{\sigma}_e(full\ model)] = 0.67$$

Variance explained by level 2:

$$R_{L1}^2 = [1 - \hat{\sigma}_{\mu 0}(full\ model)/\hat{\sigma}_{\mu 0}(full\ model)] = 0.31$$

These findings substantiate MRM's ability and effectiveness in quantifying approximately 0.67% of the geographic effects (aggregated level) and 0.31% of the household and dwelling effects (individual level).

Final Model Results

This study developed a combined bottom-up and top-down model to inform researchers and policymakers about the importance of considering regional effects on residential energy consumption. The empirical results of our model are summarized in Table 5.

The final model included all of the variables selected by the AdaEnetR procedure. Figs. 4 and 5 present the coefficient shrinkage and distribution of the sum of squared errors of the tuning parameter λ. AdaEnetR not only can assess the direction of index and important factors selected simultaneously, but also may avoid assessing the unknown link function through the nonparametric method.

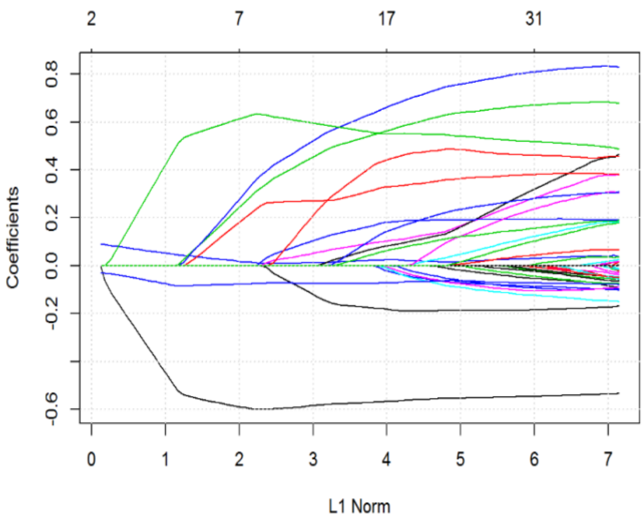

Fig.4. Adaptive Elastic Net shrinkage

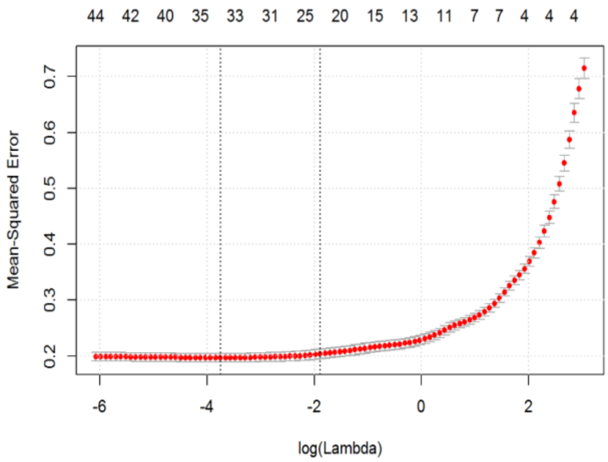

Fig.5. distribution of the sum of squared errors of the tuning parameter λ

The results of the final model are displayed in Table 5. The model fit was reasonable overall and suggested that the final model explained 76% of the variation in residential energy consumption.

Table 5
Parameter estimates of the final model

Variable	Categories	Estimate	*p* (Sig.)	SE
Log of department energy price		0.01	N.S.	0.05
Log of department days degrees		0.47	***	0.07
Department poverty rate		-0.01	N.S.	0.01
	Level 1 variables			
Log of energy price		-0.54	***	0.02

Log of surface		0.48	***	0.03
Log of housing temperature		0.47	***	0.12
Appliance ratio		0.17	***	0.03
Number of household members (NHM)	2 persons vs. 1 person	0.17	***	0.03
	3-4 persons vs. 1 person	0.38	***	0.05
	More than 5 persons vs. 1 person	0.44	***	0.07
HRP age	From 45 to 66 vs. < 44	0.053	**	0.03
	More than 67 vs. < 44	0.08	**	0.03
Heating system (HS)	Shared central vs. Other	0.87	***	0.06
	Individual central vs. Other	0.72	***	0.04
Energy class (EPC)	D vs. A-B-C	0.00		0.03
	E vs. A-B-C	0.10	***	0.03
	F-G vs. A-B-C	0.17	***	0.03
Housing type	Individual house vs. shared building	0.17	***	0.03
Heating energy (HE)	Electricity vs. other	-0.18	***	0.05
	Gas vs. other	0.20	***	0.05
	Fuel vs. other	0.38	***	0.06
	GPL vs. other	0.32	***	0.08
	Urban heating vs. other	0.24	**	0.09
Household preference between economy and comfort	Economy vs. Comfort	-0.08	***	0.02
Lower the heating in bedrooms during the night	No vs. Yes	0.06	***	0.02
Duration housing unoccupied per week	More than 8H vs. 4-8 H	-0.10	***	0.03

Note: The dependent variable is the natural logarithm of household energy consumption. *** p<0.01; ** p<0.05; * p<0.10

Level-2 variables

The results of the regression model shown in Table 5 indicated that the heating degree days (HDD) was the primary factor that explains individual energy consumption demand at the department level. It is worth mentioning that HDD is a meteorological measure used to estimate the severity of the climate with respect to heating use over a period and in a location. The higher the HDD, the more heating is needed. In our paper, HDD was assessed at the department scale and reflected the heating needs of housing located in the zone during 2012. The effect was strong (+0.46) and positive: the higher the HDD, the more energy the household consumes. This result is consistent with the descriptive statistics presented in Figs. 2 and Fig 3, showing that colder departments or regions (such as Hauts-de-France, Normandie, or Grand-Est) are more likely to consume more energy. As energy consumed for heating accounts for approximately 65% of the total energy consumption in France, and because of the high heterogeneity in climate conditions between departments, HDD's significant effect on energy consumption is well-justified.

Thus, the results of the model suggest that HDD may offer a rich source of information to improve the understanding of the residential energy consumption spectrum, particularly as previous research has argued that climatic conditions play an important role in household energy consumption (Belaïd, 2017; Lin et al., 2014).

The energy price was not significant at the department level. This result seems reasonable because there are no local disparities in France relative to the energy price. This finding also is consistent with that of Tso and Guan (2014), who suggested that division-average energy prices' effects on US residential energy consumption are not significant. Finally, the poverty rate did not discriminate housing energy consumption, although a negative effect was expected.

Level-1 variables

Most of the level-1 variables in our model were significant at the 5% level, including age of HRP, household size, housing type, and energy performance certificate. As expected, the price of energy has a negative effect on energy consumption. The energy price elasticity is approximately -0.54, which is consistent with previous findings that found that estimates varied from 0.20 to 1.60 in absolute values (Belaïd, 2016; Dubin and McFadden, 1984; Labandeira et al., 2006; Nesbakken, 1999; Nesbakken Runa, 2002).

Households' socioeconomic characteristics, such as the number of occupants and the age of the reference person, were found to have a significant positive effect on energy consumption. Previous research has demonstrated that, indeed, the life-cycle stage influences energy consumption patterns (Brounen et al., 2012; Leahy and Lyons, 2010; Lévy and Belaïd, 2018). Our findings supported the normal distribution of energy consumption as a function of the HRP's age. Belaïd and Garcia (2016) suggested that individual aspirations are a key feature in explaining the right tail of the normal distribution, while family size explains the left (children's birth). More precisely, household energy demand increases with succeeding stages of the life cycle between the child-rearing years and when the children leave the family. Thereafter, energy demand decreases throughout the remaining stages of the cycle. In fact, the normal distribution of domestic energy demand was demonstrated first by Fritzsche (1981). Using 1972-

73 data from the Bureau of Labor Statistics survey, he substantiated the hypothesis above when he found a significant difference in home energy use with respect to succeeding stages of the family life cycle.

Dwelling attributes also affect French household energy consumption significantly, in which dwelling size was found to have a significant positive effect on energy consumption. A 10% increase in the surface area of the housing increases energy consumption by 5%. Moreover, energy-efficient dwellings (energy classes A-B-C) indeed consume less than do non-energy efficient dwellings: -10% compared to housing units in energy class E and -17% compared to those in energy classes F-G. In addition, individual houses consume more energy than do shared buildings.

Finally, our results suggested that household behavioral factors play an important role in residential energy consumption. Not surprisingly, the temperature in the home influences energy consumption strongly: a 10% increase in the indoor temperature during the winter leads to a 5% increase in energy consumption. Preferring economy over comfort in the use of heating at the individual level implies an 8% decrease in global energy consumption.

Finally, dwellings that are occupied less often during the week, or where individuals indicated they lower the bedroom temperature during the night, consume less energy ceteris paribus. In recent years, many researchers have become increasingly aware of the effect that consumers' energy behaviors and attitudes have on natural global energy demand and environmental quality in general (Belaïd et al., 2018; Lévy et al., 2014). Influencing and changing energy consumption behavior and attitudes within households has a huge potential both to reduce energy demand and address the climate change issue. Therefore, it is essential to preserve environmental resources.

5. Conclusions and policy implications

This research addressed some of the many questions that have been posed to acquire a more comprehensive understanding of the factors that affect residential energy demand. Energy policymakers are concerned increasingly about understanding the main factors that influence household energy use, which is seen as a key strategy to reduce energy demand in the residential sector. Focusing on the French case, this study examined the role of various factors in energy renovation decisions. We explored further the role that geographic variations may play in aggregate household energy demand. Recognition of this distinction can help reveal the complex relations in variables of household energy consumption.

The methodological innovation used in this paper was a mixed bottom-up and top-down statistical approach based on the multilevel regression model (MRM) and an innovative variable selection approach to examine both geographic and household effects on energy use. Using an MRM and the AdaEnetR technique to analyze micro-level data from the 2013 French PHEBUS survey, this paper provides valuable information about the roles geographic, socioeconomic, and dwelling attributes play in residential energy demand.

MRM offers energy policymakers who must analyze large, complex survey datasets a relatively new approach to examine individual and contextual influences on residential energy consumption, and provides a valuable solution to address the clustered nature of the data to examine sources of variation within and across clusters.

The specific objective of this research was to obtain a clearer understanding of residential energy demand in the context of energy efficiency improvements, including its various causal factors, and to assess the literature that examines these variations critically. In particular, we recommend a multilevel regression approach to allocate total fluctuations of residential energy demand among geographic variations and household and housing variations. Further, we explored the role that variations in geographic conditions might play in aggregate household energy demand in more detail.

Fit statistics for the null model with department as the location variable were greater than 10%, and ICC=20% for the null model with scaled weight B; thus, regional effects explain approximately 20% of the variation in French residential energy consumption. This confirms the importance of using multilevel regression to capture these effects. Our empirical findings showed MRM's ability to quantify effectively approximately 0.67% of geographic effects (aggregate level) and approximately 0.31% of household and dwelling effects (individual level). Further, the analysis suggested that household attributes are an important factor in residential energy consumption patterns, and the proportion of the total variance explained was approximately 76% compared to <70% using the OLS regression model.

The MRM model suggests that, among the level-2 explanatory variables, family HDD is one of the most influential factors in household energy consumption, and accordingly, we concluded that regional effects are crucial in examining household energy consumption patterns. Further, most of the level-1 explanatory factors were found to be significant at the 5% level, including household socioeconomic factors and behavior attributes, and housing characteristics.

The findings of this research contribute to recent research designed to increase our knowledge of the salient influences on household energy consumption by incorporating several theoretical approaches and empirical methods, including economics and sociology. These findings can be used to support subsequent decision-making in energy policies and the formulation of efficiency programs.

In addition to contributing to the energy policy debate, this study contributes to ongoing research on residential energy consumption by providing a more elaborate overview of its various facets. However, there is a further challenge for researchers and policymakers in developing multiple strategies and intelligent policies to foster improved energy efficiency in the residential sector, including: (1) regulation reforms (e.g., cost-effective energy pricing, energy efficiency targets by sector, codes/standards with enforcement mechanisms development, etc.); (2) data and information collection (e.g., databases on energy consumption, information, and case studies, etc.); (3) incentive and financial measures (e.g., public sector energy efficiency financing, residential and home appliance credit, etc.); (4) technical capacity improvement (e.g., certification programs and energy audit/manager training, development of energy management systems, etc.), and (5) institutional reforms (e.g., dedicated entities with energy efficiency mandates, clear institutional roles/accountability, and authority to formulate, implement, evaluate, and report on programs, etc.).

References

Algamal, Z.Y., Lee, M.H., 2015. Regularized logistic regression with adjusted adaptive elastic net for gene selection in high dimensional cancer classification. *Comput. Biol. Med. 67*, 136–145. https://doi.org/10.1016/j.compbiomed.2015.10.008

Belaïd, F., 2017. Untangling the complexity of the direct and indirect determinants of the residential energy consumption in France: Quantitative analysis using a structural equation modeling approach. *Energy Policy 110*, 246–256. https://doi.org/10.1016/j.enpol.2017.08.027

Belaïd, F., 2016. Understanding the spectrum of domestic energy consumption: Empirical evidence from France. *Energy Policy 92*, 220–233. https://doi.org/10.1016/j.enpol.2016.02.015

Belaïd, F., Bakaloglou, S., Roubaud, D., 2018. Direct rebound effect of residential gas demand: Empirical evidence from France. *Energy Policy 115*, 23–31. https://doi.org/10.1016/j.enpol.2017.12.040

Brounen, D., Kok, N., Quigley, J.M., 2012. Residential energy use and conservation: Economics and demographics. *Eur. Econ. Rev., Green Building, the Economy, and Public Policy 56*, 931–945. https://doi.org/10.1016/j.euroecorev.2012.02.007

Carle, A.C., 2009. Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Med. Res. Methodol. 9*, 49. https://doi.org/10.1186/1471-2288-9-49

Dubin, J.A., McFadden, D.L., 1984. An econometric analysis of residential electric appliance holdings and consumption. *Econometrica 52*, 345–362. https://doi.org/10.2307/1911493

Fritzsche, D.J., 1981. An analysis of energy consumption patterns by stage of family life cycle. *J. Mark. Res. 18*, 227–232. https://doi.org/10.2307/3150956

Geng, Y., Chen, W., Liu, Z., Chiu, A.S.F., Han, W….Cui, X., 2017. A bibliometric review: Energy consumption and greenhouse gas emissions in the residential sector. *J. Clean. Prod. 159*, 301–316. https://doi.org/10.1016/j.jclepro.2017.05.091

Hox, J.J., Moerbeek, M., Schoot, R. van de, 2010. *Multilevel analysis: Techniques and applications*, 2nd ed. New York, NY: Routledge.

Hu, J., Huang, J., Qiu, F., 2018. A group adaptive elastic-net approach for variable selection in high-dimensional linear regression. *Sci. China Math. 61*, 173–188. https://doi.org/10.1007/s11425-016-0071-x

Labandeira, X., Labeaga, J.M., Rodríguez, M., 2006. A residential energy demand system for Spain. *Energy J. 27*, 87–111.

Leahy, E., Lyons, S., 2010. Energy use and appliance ownership in Ireland. *Energy Policy 38*, 4265–4279. https://doi.org/10.1016/j.enpol.2010.03.056

Lévy, J.-P., Belaïd, F., 2018. The determinants of domestic energy consumption in France: Energy modes, habitat, households and life cycles. *Renew. Sustain. Energy Rev. 81*, 2104–2114. https://doi.org/10.1016/j.rser.2017.06.022

Lévy, J.-P., Roudil, N., Flamand, A., Belaïd, F., 2014. Les déterminants de la consommation énergétique domestique (The determinants of domestic energy consumption). *Flux vol*, 40–54. https://doi.org/10.3917/flux.096.0040

Li, X., Li, G., Yang, S., 2011. Adaptive elastic-net for general single-index regression models, In X.Y. (Ed.): *Nonlinear mathematics for uncertainty and its applications, advances in intelligent and soft computing*. Berlin, Heidelberg, Germany: Springer, pp. 509–516. https://doi.org/10.1007/978-3-642-22833-9_62

Lin, W., Chen, B., Luo, S., Liang, L., 2014. Factor analysis of residential energy consumption at the provincial level in China. *Sustainability 6*, 7710–7724. https://doi.org/10.3390/su6117710

Liu, W., Li, Q., 2017. An efficient elastic net with regression coefficients method for variable selection of spectrum data. *PLOS ONE 12*, e0171122. https://doi.org/10.1371/journal.pone.0171122

Mazur, A., Rosa, E., 1974. Energy and life-style. *Science 186*, 607–610.

Muthén, B., Asparouhov, T., 2009. Multilevel regression mixture analysis. *J. R. Stat. Soc. Ser. A Stat. Soc. 172*, 639–657.

Nesbakken, R., 1999. Price sensitivity of residential energy consumption in Norway. *Energy Econ. 21*, 493–515. https://doi.org/10.1016/S0140-9883(99)00022-5

Nesbakken Runa, 2002. Energy consumption for space heating: A discrete–continuous approach. *Scand. J. Econ. 103*, 165–184. https://doi.org/10.1111/1467-9442.00236

Robinson, W.S., 2009. Ecological correlations and the behavior of individuals. *Int. J. Epidemiol. 38*, 337–341. https://doi.org/10.1093/ije/dyn357

Schulte, I., Heindl, P., 2017. Price and income elasticities of residential energy demand in Germany. *Energy Policy 102*, 512–528. https://doi.org/10.1016/j.enpol.2016.12.055

Sonderegger, R.C., 1978. Movers and stayers: The resident's contribution to variation across houses in energy consumption for space heating. *Energy Build. 1*, 313–324. https://doi.org/10.1016/0378-7788(78)90011-7

Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol. 58*, 267–288.

Tso, G.K.F., Guan, J., 2014. A multilevel regression approach to understand effects of environment indicators and household features on residential energy consumption. *Energy 66*, 722–731. https://doi.org/10.1016/j.energy.2014.01.056

Zou, H., Zhang, H.H., 2009. On the adaptive elastic-net with a diverging number of parameters. *Ann. Stat. 37*, 1733–1751. https://doi.org/10.1214/08-AOS625