



HAL
open science

Nudging electricity consumption in households: A case study of french residential sector

Muhammad Salman Shahid

► **To cite this version:**

Muhammad Salman Shahid. Nudging electricity consumption in households: A case study of french residential sector. Electric power. Université Grenoble Alpes [2020-..], 2022. English. NNT: 2022GRALT031 . tel-03766410

HAL Id: tel-03766410

<https://theses.hal.science/tel-03766410v1>

Submitted on 1 Sep 2022

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.



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : GENIE ELECTRIQUE

Arrêté ministériel : 25 mai 2016

Présentée par

Muhammad Salman SHAHID

Thèse dirigée par **Benoit DELINCHANT**, Université Grenoble Alpes

et codirigée par **Béatrice ROUSSILLON**, Université Grenoble Alpes

et co-encadrée par **Daniel LLERENA**, Laboratoire d'Economie Appliquée de Grenoble (GAEL) et **Frédéric WURTZ**, CNRS

préparée au sein du **Laboratoire Laboratoire de Génie Electrique**

dans l'**École Doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)**

Nudging la consommation d'électricité dans les ménages : Un cas d'étude du secteur résidentiel français

Nudging electricity consumption in households: A case study of french residential sector

Thèse soutenue publiquement le **7 avril 2022**,
devant le jury composé de :

Monsieur Benoît DELINCHANT

PROFESSEUR DES UNIVERSITES, GRENOBLE INP, Directeur de thèse

Monsieur Romain BOURDAIS

PROFESSEUR DES UNIVERSITES, CentraleSupélec , Rapporteur

Madame Sandrine SELOSSE

Chargé de recherche HDR, MINES PARIS TECH, Rapporteur

Madame Béatrice ROUSSILLON

Maître de conférences HDR, UNIVERSITE GRENOBLE ALPES, Co-directrice de thèse

Monsieur Jérôme LEDREAU

MAITRE DE CONFERENCE, UNIVERSITE DE LA ROCHELLE,
Examineur

Monsieur Stephan MARETTE

DIRECTEUR DE RECHERCHE, INRAE ILE-DE-FRANCE-VERSAILLES-GRIGNON, Examineur

Monsieur Yvon BESANGER

PROFESSEUR DES UNIVERSITES, GRENOBLE INP, Président

Table of Contents

Table of Contents	i
List of Figures	vii
List of Tables.....	xi
Acknowledgement.....	xiii
Abstract.....	xiv
Chapter I	1
I.1 Electricity Supply, Associated CO ₂ Emission and Necessity of Energy Flexibility .	3
I.1.a Energy transition as a tool to mitigate climate change	3
I.1.b The quest for sustainable energy mix in European Union and in France.....	5
I.1.c The impact of renewable energy mix on electricity price.....	7
I.1.d Intermittency of renewable energy and need of energy flexibility.....	8
I.2 Energy Flexibility and its usual modes of implementation	10
I.2.a Incentive based direct Flexibility and its limitations	11
I.2.b Price based indirect flexibility and its limitations	13
I.3 The challenge of energy flexibility in residential sector.....	15
I.3.a Potentiality of residential sector	15
I.3.b Identification of potential appliances for energy flexibility	16
I.3.c Non-monetary indirect energy flexibility: A new mechanism for residential sector	18
I.4 Conclusion.....	21
Chapter II.....	23
II.1 Price based energy flexibility in the light of classical economics.....	25
II.2 Behavioral insights for indirect energy flexibility.....	26
II.2.a Nudge.....	28
II.2.b Green Nudge	29
II.2.c Review of nudges used in the field of energy worldwide.....	30
II.2.d The Set of Green Nudges Used In the Experiments	33

II.3	An experiment of indirect energy flexibility in residential sector using green nudges	33
II.3.a	Methodological aspect of the experiment	34
II.3.b	Protocol of interaction between subjects and experiment organizers.....	35
II.3.c	Design of set of green nudges	41
II.3.d	Architecture of Experiment	44
II.4	Prediction of modelled national load curve of residential sector	50
II.4.a	Random Forest Regression.....	51
II.4.b	Regression Model Creation.....	52
II.4.c	Third Party APIs used for Forecast.....	53
II.4.d	Regression Model Parameters	56
II.4.e	Model Evaluation and relative importance of features	57
II.4.f	Performance Metric	60
II.4.g	Conclusion on the prediction of French residential load curve.....	61
II.5	Conclusion.....	61
Chapter III	63
III.1	Introduction	65
III.2	General Statistics of the alerts.....	65
III.2.a	Green Alerts	67
III.2.b	Orange Alerts	68
III.3	Comparison of statistical variables of control group and treated group	69
III.4	Comparison of normal energy consumption between two groups.....	72
III.4.a	Visualization of statistical distribution on normal days	73
III.4.b	Assessment of statistical difference in the normal energy consumption	75
III.5	Analysis of the impact of nudge signals	76
III.5.a	Visualization of statistical distribution on alert days	76
III.5.b	Assessment of statistical difference in energy consumption on alert days ...	78
III.5.c	Quantification of change in the energy consumption during the period of intervention	81
III.6	Conclusion	87

Chapter IV.....	89
IV.1 Introduction	91
IV.2 Literature review of Indirect Feedback.....	91
IV.3 Formulation of Reference load curve.....	92
IV.3.a Key Performance Indicators	94
IV.3.b Techniques studied to formulate reference load curve	96
IV.3.c Analysis for the suitability of a technique to formulate the reference load curve 102	
IV.4 Reference load curve in the form of feedback.....	106
IV.5 Impact of error ϵ non – alert in the calculation of energy flexibility.....	108
IV.6 Analysis regarding reference load curve as an alternative of control group.....	109
IV.7 Measure of response of individual subjects towards orange alerts with respect to the reference load curve	112
IV.8 Conclusion.....	116
Chapter V	117
V.1 Introduction	119
V.2 Non-Intrusive Load Monitoring	119
V.2.a Energy disaggregation datasets	120
V.2.b NILM Techniques	121
V.2.c Remarks on the state of the art.....	122
V.3 Methodology	123
V.4 Committed appliances by the subjects.....	123
V.5 The characteristics of IRISE dataset.....	125
V.5.a Committed appliances of treated group in the IRISE dataset.....	125
V.5.b The statistics of appliances in IRISE dataset.....	126
V.6 Reproducibility of an open-source energy disaggregation algorithm	128
V.6.a Training and testing on same dataset (REFIT)	130
V.6.b Training on REFIT and testing on REDD dataset: transfer learning.....	131
V.7 NILM on IRISE dataset.....	133
V.7.a Removing outliers from IRISE dataset.....	133

V.7.b	Predictive model creation based on IRISE dataset.....	134
V.7.c	The results of energy disaggregation in terms of performance indicators ..	136
V.8	Transferability on the EXPESIGNO dataset	138
V.9	Conclusion.....	139
Chapter VI.....		141
VI.1	General Conclusion	143
VI.2	Future Prospects	146
VI.2.a	Use of other nudges of behavioral science for energy flexibility	146
VI.2.b	Architecture development for experiment.....	147
VI.2.c	Reference Load curve.....	147
VI.2.d	In-depth data analysis of potential appliances for energy flexibility	148
ANNEX A.....		149
A.1	Web interface of Etude Elec Site	149
A.2	Homepage of personal account.....	149
A.3	Default commitments for green alert.....	150
A.4	Default Commitment of Orange Alert.....	150
A.5	Feedback of the alert	151
ANNEX B		153
B.1	Household Characteristics	153
B.2	Presence of appliances	155
B.3	Number of appliances in household	157
B.4	Energy Consumption during weekdays	158
B.5	Energy Consumption during the weekend.....	161
ANNEX C		165
C.1	The results of OLS regression for all green alerts	165
C.2	The results of OLS regression for green alerts of session 1.....	166
C.3	The results of OLS regression for green alerts of session 2	167
C.4	The results of OLS regression for all orange alerts	168
C.5	The results of OLS regression for orange alerts of session 1.....	169
C.6	The results of OLS regression for orange alerts of session 2	170

ANNEX D.....	171
D.1 The Mann-Whitney U Test for green alert days of both sessions using random forest (sequence to sequence) reference load curve in place of control group	171
D.2 The Mann-Whitney U Test for orange alert days of both sessions random forest (sequence to sequence) reference load curve in place of control group	171
ANNEX E.....	173
E.1 The split of IRISE households data with respect to each appliance	173
E.2 The length of data subsets for training the predictive model of each appliance	174
ANNEX F	175
F.1 Test result for hotplate on IRISE house 993 at sampling rate of 30 minutes	175
F.2 Test result for electric heating on IRISE house 990 at sampling rate of 30 minutes	175
F.3 Test result for electric oven on IRISE house 997 at sampling rate of 30 minutes	176
F.4 Test result for electric cooker on IRISE house 902 at sampling rate of 30 minutes	176
F.5 Test result for clothes drier on IRISE house 996 at sampling rate of 30 minutes	177
F.6 Test result for dishwasher on IRISE house 997 at sampling rate of 30 minutes	177
F.7 Test result for microwave oven on IRISE house 997 at sampling rate of 30 minutes	178
F.8 Test result for tv on IRISE house 997 at sampling rate of 30 minutes.....	178
F.9 Test result for washing machine on IRISE house 997 at sampling rate of 30 minutes.....	179
BIBLIOGRAPHY.....	181
RÉSUMÉ.....	199

List of Figures

Figure I.1	Global Warming Projection w.r.t different scenarios until 2100 (Climate Action Tracker, 2021)	4
Figure I.2	Annual Power Production (TW) in France	6
Figure I.3	Annual CO ₂ Emission (MTCO ₂ eq.) in France	7
Figure I.4	(a) Conventional merit order (b) New merit order including renewable energy sources (Bahar & Sauvage, 2013)	8
Figure I.5	Maturity of Energy Storage Technologies until 2017 (Nguyen et al., 2017)....	9
Figure I.6	Flexibility solutions relevant to intermittent renewable production (International Energy Agency, 2018)	10
Figure I.7	Hierarchy of demand response (Alasseri et al., 2021).....	11
Figure I.8	Modelled average daily load curve of residential sector (ENEDIS, 2021)	15
Figure I.9	Annual Energy Consumption (kWh) of household appliances (Energie Douce, 2021; Energie Mobile, 2021; Réseau de Transport d'Électricité France, 2017, 2018, 2019, 2020a)	17
Figure I.10	Nominal power of household appliances (Energie Développement, 2013; Energie Douce, 2021; Energie Mobile, 2021)	17
Figure I.11	Average household energy saving with respect to type of feedback(Ehrhardt-Martinez & Donnelly, 2010)	19
Figure I.12	Behavioral Insight based signal to implement indirect energy flexibility.....	20
Figure I.13	Indirect feedback in graphical form of load curve.....	21
Figure II.1	Factors pertaining to decision making	26
Figure II.2	Green nudge to reduce the excessive towel paper use (Lars Biesewig & Annie Krautkraemer, 2018)	29
Figure II.3	An example of randomized controlled field trial (Haynes et al., 2012).....	35
Figure II.4	Load shifting from evening peak hours to afternoon	37
Figure II.5	Load curtailment during peak hours in the evening	38
Figure II.6	Curious questions and corresponding designed nudge treatment.....	41
Figure II.7	Feature salience through comparison of load curves	44
Figure II.8	Schematic Diagram of project “EXPESIGNO”	45
Figure II.9	Timeline illustrating the signaling on day ‘D-1’	45
Figure II.10	Timeline illustrating the reminder SMS and period of action on day ‘D’ ..	46
Figure II.11	Timeline illustrating the indirect feedback on day ‘D+3’	46
Figure II.12	Calendar of the months for orange alerts based on EJP.....	49
Figure II.13	Calendar of the months for orange alert based on TEMPO for 1st session	49
Figure II.14	Random Forest Regression	51

Figure II.15	Principal of regression model creation, validation and use in real time ...	52
Figure II.16	Load curve of residential profile coefficient by ENEDIS	53
Figure II.17	The procedure of smoothening the measured temperature of 32 cities	55
Figure II.18	The prediction model for predicting profile coefficient of day ‘D’	56
Figure II.19	The correlation of profile coefficient with lagged profile coefficients.....	56
Figure II.20	Prediction of French residential load curve by random forest regression.	58
Figure II.21	Relative Importance of temporal features	58
Figure II.22	Relative Importance of temporal features and lagged profile coefficient..	59
Figure II.23	Relative Importance of all features in model	59
Figure II.24	The comparison of performance metric.....	60
Figure III.1	The impact of lockdown on French energy consumption during 2020 (Réseau de Transport d’Électricité France, 2020a).....	67
Figure III.2	The distribution of non-alert days of green sessions	74
Figure III.3	The distribution of non-alert days of orange sessions	75
Figure III.4	The distribution of non-alert days and alert days of green session 1.....	77
Figure III.5	The distribution of non-alert days and alert days of green session 2	77
Figure III.6	The distribution of non-alert days and alert days of session 1 of type ‘orange’ 78	
Figure III.7	The distribution of non-alert days and alert days of session 2 of type ‘orange’ 78	
Figure III.8	The statistical results of OLS regression for green sessions	83
Figure III.9	The percentage change in the energy consumption on alert day as compared to non-alert day.....	84
Figure III.10	The statistical results of OLS regression for orange sessions	85
Figure III.11	The percentage change in the energy consumption on alert day as compared to non-alert day for 1st session of orange alerts	86
Figure III.12	The percentage change in the energy consumption on alert day as compared to non-alert day for 2nd session of orange alerts	87
Figure IV.1	An ideal reference load curve.....	93
Figure IV.2	A practical reference load curve.....	93
Figure IV.3	Formulation of reference load curve using Kernel Density Estimation	97
Figure IV.4	Selecting the minimum energy measured load curve as reference load curve 98	
Figure IV.5	Selecting the maximum energy measured load curve as reference load curve 99	
Figure IV.6	Green reference load curve formulation by taking average of measured load curve of day ‘D-1’ and day having minimum consumption between day ‘D-2’ and day ‘D-5’ 99	

Figure IV.7	Orange reference load curve formulation by taking average of measured load curve of day ‘D-1’ and day having maximum consumption between day ‘D-2’ and day ‘D-5’	99
Figure IV.8	Random Forest Point to Point technique for creation of green reference load curve	100
Figure IV.9	Random Forest Point to Point technique for creation of orange reference load curve	100
Figure IV.10	Training the model and predicting the instance 6:30 PM for a day ‘D’...	101
Figure IV.11	Random Forest Sequence to Sequence for creation of green reference load curve	101
Figure IV.12	Random Forest Sequence to Sequence for creation of orange reference load curve	101
Figure IV.13	Distribution of energy difference indicator for green reference load curves	103
Figure IV.14	Distribution of R ² score for green reference load curves	104
Figure IV.15	Distribution of energy difference indicator for orange reference load curves	105
Figure IV.16	Distribution of R ² score for orange reference load curves	106
Figure IV.17	Indirect feedback using reference load curve for (a) Green Alert and (b) Orange Alert	107
Figure IV.18	The energy flexibility of subjects who curtailed their load for most of the orange alerts	114
Figure IV.19	The energy flexibility of subjects who did not remain positive for most of the orange alerts	115
Figure IV.20	The energy flexibility of subjects whose effort of energy flexibility cannot be determined	115
Figure V.1	Non-intrusive load monitoring (Parson, 2017)	120
Figure V.2	Schematic diagram of energy disaggregation of measured load curves on alert days	123
Figure V.3	Frequency of appliances committed for green alert	124
Figure V.4	Frequency of matched appliances in IRISE dataset with committed appliances	125
Figure V.5	Frequency of closely matched appliances in IRISE dataset with committed appliances of treated group	126
Figure V.6	Energy disaggregation of washing machine via model trained on REFIT dataset and tested on REFIT house no. 8. Grey curve represents aggregated consumption, red curve represents the actual consumption of washing machine and blue curve represents the predic	131

Figure V.7 The comparison of peak energy consumption of washing machine in REFIT dataset with REDD dataset 132

Figure V.8 Energy disaggregation of washing machine via model trained on REFIT dataset and tested on RED house no. 3. Grey curve represents aggregated consumption, red curve represents the actual consumption of washing machine and blue curve represents the predicted curve of washing machine 132

Figure V.9 Detection of outliers in washing machine of a household in IRISE dataset 134

Figure V.10 Test result for dishwasher on IRISE house 997 at sampling rate of 30 minutes 138

List of Tables

Table I.1	Short Term and Medium Term Targets set by EU to mitigate climate change (Amanatidis, 2019)	5
Table I.2	French multi-year energy programs (Ministry of ecological transition and solidarity France, 2019)	6
Table I.3	Merits and Limitations of traditional energy flexibility	14
Table II.1	Attributes of dual system of thinking (J. S. B. T. Evans, 2008)	27
Table II.2	The nudge based indirect energy flexibility experiments in households	31
Table II.3	The nudges used in the experiment	33
Table II.4	Statistics of performance metric for each type of model	61
Table III.1	Number of households in each group with respect to each year	65
Table III.2	Number of alert days with respect to each session and alert type	66
Table III.3	Basis of triggered green alerts	68
Table III.4	Basis of triggered orange alerts	68
Table III.5	Chi-square test of independence of various variables	70
Table III.6	The Mann-Whitney U Test for non-alert days of both types of sessions	76
Table III.7	The Mann-Whitney U Test for green alert days of both sessions	79
Table III.8	The Mann-Whitney U Test for orange alert days of both sessions	80
Table III.9	The variables of ordinary least square test	81
Table III.10	The assessment of impact of green alert on alert day as compared to non alert days	84
Table III.11	The assessment of impact of orange alert on alert day as compared to non-alert days	86
Table IV.1	Ideal and practical case of energy difference indicator on non-alert days	95
Table IV.2	The choice of techniques to formulate reference load curve	96
Table IV.3	Green and orange load curves to measure energy flexibility Eflexibility, alert	109
Table IV.4	The Mann-Whitney U Test for green alert days of both sessions using $MeanPD - 1, i, MinPD - 2, i, PD - 3, i, PD - 4, i, PD - 5, i$ reference load curve in place of control group	110
Table IV.5	The Mann-Whitney U Test for orange alert days of both sessions using $MeanPD - 1, i, MaxPD - 2, i, PD - 3, i, PD - 4, i, PD - 5, i$ reference load curve in place of control group	111
Table IV.6	Sensitivity analysis of thresholds of identifying responses	113
Table V.1	Some open source NILM datasets	120
Table V.2	Statistics for IRISE appliances at 10 minutes sampling	127

Table V.3	Statistics for IRISE appliances at 30 minutes sampling.....	127
Table V.4	The length of data subsets in REFIT and REDD dataset.....	129
Table V.5	The comparison of model creation with and without GPU	130
Table V.6	Trained and tested on REFIT dataset at sample period of 8 seconds	130
Table V.7	The statistical characteristics of appliances in REFIT and REDD dataset...	131
Table V.8	Trained on REFIT and tested on REDD dataset at sample period of 8 seconds 132	
Table V.9	The length of data subsets for training the predictive model of each appliance 135	
Table V.10	The performance indicators for testing the predicted models	135
Table V.11	The results with respect to performance indicators for data sampled at 10 minutes	136
Table V.12	The results with respect to performance indicators for data sampled at 30 minutes	137

Acknowledgement

First of all, I would like to thank my thesis director, **Benoît Delinchant**, for giving me the opportunity to work on this project and for always being available to discuss different problems related to project and thesis. His precious suggestions and guidance helped me to learn a lot over the course of project development and thesis redaction. I would also like to thank my thesis co-director, **Béatrice Roussillion**, whose guidance assisted me a lot for the work done regarding experimental economics. I would also like to thank my other co-supervisors **Frédéric Wurtz** and **Daniel Llerena**, for their unfailing motivation. Working with them on this project is both enriching and stimulating. Under the supervision of these supervisors, I am able to present a thesis based on the cross-disciplinary impact of experimental economics and electrical engineering.

The rest of the list is long, and I am thankful to all those who supported me through the development of experimental work and thesis redaction. I would specifically like to thanks:

- **Tiansi Laranjeira**: For her help and support regarding the operation and debugging of experiment architecture.
- **Jérôme Ferrari**: For a) indicating that we can use EJP/TEMPO as a protocol if there are no more anticipated PP1/PP2 days and b) his assistance regarding creating and deploying a SMS sending platform.
- **Nils Artiges**: For his guidance to overcome different obstacles in the project and thesis work. His remarks on reports and presentations help me a lot to improve the presentation skills.
- **Adélaïde Fadhuile** and **Penelope Buckley**: For their guidance regarding econometric analysis of the energy consumption.
- **Cédric Lanu** and **Erick Raberanto**: For working on the Etude-Elec sub-domain and keeping track of subjects activity on this platform.
- **Manon Boudet**, **Manon Bacchetta** and **Emma Papet**: For working on the analysis on socio-technical characteristics of the households as part of their course project.

Finally, I would also like to acknowledge and thanks the funding agency to work on this project. This work has been supported by the Auvergne Rhône-Alpes region, with "Pack Ambition Recherche 2017 - Expesigno projet" (17 01110301). This work has also been partially supported by the CDP Eco-SESA receiving fund from the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02).

Abstract

The energy transition towards renewable energy sources and distributed generation is essential to mitigate climate change. However, the renewable energy sources are intermittent in nature, and therefore it may cause network congestion during peak consumption in future. Therefore, demand side energy flexibility is inevitable in complement to the supply side management and energy storage. The residential sector is the largest energy consumer and is therefore essential to implement demand-side energy flexibility. In addition to the existing approaches of energy flexibility, we present in this thesis a non-monetary “indirect” flexibility through the implementation of a behavioral science experiment with 175 French households. The set of nudges used in this experiment does not oblige the participating households to implement energy flexibility for either monetary gain or loss aversion.

Nudge alerts are generated by predictive algorithms and sent by SMS to a group of households with the aim of carrying out either load shifting or load shedding. The nudge alerts are complemented with the suggestions about using (or not using) committed appliances to implement energy flexibility. After each alert, the group receives graphical feedback in which the load curve measured by LINKY smart meter is superposed by a theoretical reference curve specific to each household. In comparison to a control group, we can quantify the impact of these nudge signals on the energy consumption of the treated group. Finally, an energy disaggregation algorithm is presented that will allow us to detect the equipment used during the flexibility and thus will do an analysis of the commitment of households to look for potential appliances for energy flexibility.

Chapter I

Electricity grid and social challenge: Inevitability of residential consumer participation in energy flexibility

Energy flexibility has been in practice since many decades. However, in a future energy mix predominant with distributed generation of renewable sources, the need of demand side energy flexibility will be inevitable to complement supply side management and energy storage. The residential sector remains the focus of demand response programs, however the traditional demand response technique brings certain constraints towards the decision making of residential consumer.

This chapter emphasizes the need on energy flexibility in the future. In addition to existing demand response technique, a new technique is briefly introduced. The purpose of this technique is to use behavioral insights to aid the decision making of residential consumer without causing any monetary loss.

I.1 Electricity Supply, Associated CO₂ Emission and Necessity of Energy Flexibility

I.1.a Energy transition as a tool to mitigate climate change

The supply of sustainable energy is a necessity for humankind. Solar energy remains the most used form of energy as it provides heat and light throughout the day. The discovery of fossil energy sources and technologies introduced industrial eras in the world. The first industrial era in 19th century served the humanity with coal for mass scale energy consumption. The second industrial era in 20th century centralized electricity as bulk energy source for mass scale consumption. Electrical energy is traditionally generated by base load power plants and peak load power plants. The base load power plants (such as hydroelectric, coal-fired and nuclear power plants) are operational all the time. The energy provided by these power plants ensures fulfillment of minimum level of demand on an electrical grid. Conventionally, the peak load power plants (such as oil-fired and gas-fired power plants) are connected to the electric grid for balancing peak energy consumption.

The reliance on fossil fuel power plants brings CO₂ emission, which is a main contributor in the phenomenon of global warming. This leads to the climate change. The issue of climate change has been in discussion worldwide in a series of conferences since 1992. As a result, 191 head of states ratified ‘Paris Agreement’ in COP21¹. This agreement binds the signatories to hold “the increase in the global average temperature to well below 2 °C the pre-industrial levels” by the end of 21st century (United Nations, 2015).

The global warming beyond 1.5 °C is an existential threat to the biosphere and it has a considerable impact on human life. It is estimated by the pledges and targets of various countries that the global warming may reach 2.4 °C by the end of this century; however, with the current policies in place, the global warming may reach 3.1 °C by the end of this century (Climate Action Tracker, 2021). The projections are alarming for the sustainability of biosphere and the impact of climate change will become harsher over the course of time. Some impacts of climate change that the humanity is currently facing are as follows:

- Increased frequency of direct impacts i.e., drought, flood, heatwave and wildfire;
- Increase in indirect impacts i.e., water quality, air quality and ecological change;
- Variation in environmental factors i.e., geography and vegetation;
- Social impacts i.e., loss of habitation, poverty and hunger;

¹ COP21 is the 21st Conference of Parties (COP21) organized by UNFCCC (United Nations Framework Convention on Climate Change) in Paris in November 2015.

- Health impacts like mental stress, undernutrition, respiratory and cardiovascular disease (World Health Organisation, 2018)

The scientists argue that the atmospheric CO₂ concentration must not exceed 350 ppm² to maintain the global warming significantly less than 1.5 °C, however this threshold is surpassed in 1988 whereas it is exceeded to 410 ppm in 2020 (Desing & Widmer, 2021). Figure I.1 represents the global warming projection until 2100. It is therefore required to decrease the atmospheric CO₂ concentration by decreasing the CO₂ emissions.

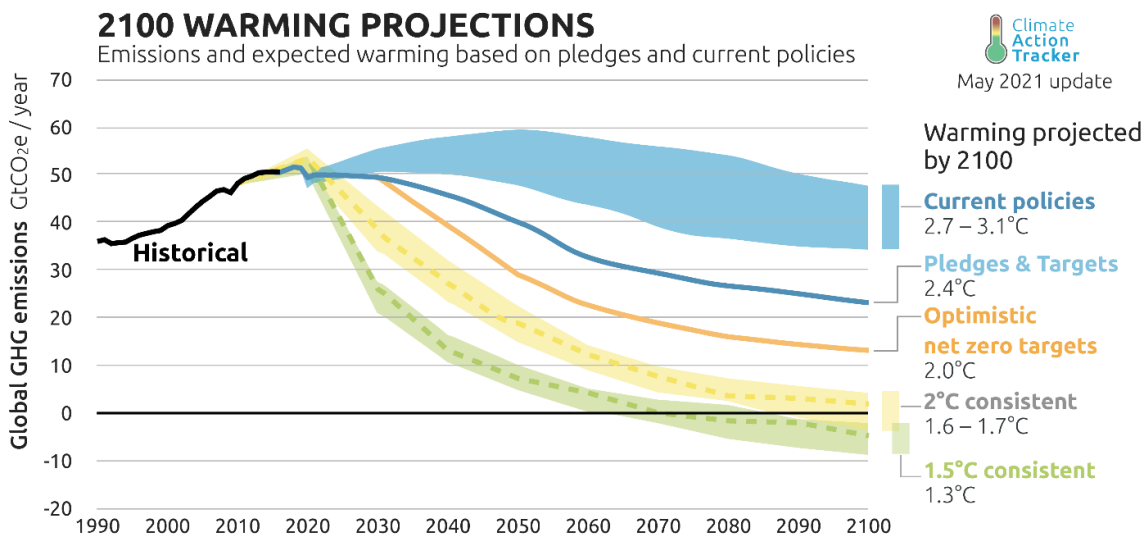


Figure I.1 Global Warming Projection w.r.t different scenarios until 2100 (Climate Action Tracker, 2021)

During 2018, an increase of 1 °C in the global temperature of the earth is observed with respect to the average global temperature of pre-industrial period (1850 to 1900), whereas an increase of 2.1 °C is observed in metropolitan France with respect to the average temperature of the period between 1961 and 1990 (Manuel Baude et al., 2021). The power generation from fossil fuel is a major contributor of CO₂ emissions in the world. During 2018, the electricity production accounted for 41% of CO₂ emission in the world, whereas 6.67% of total CO₂ emission in France is caused by electricity production (Manuel Baude et al., 2021; Réseau de Transport d'Électricité France, 2018).

Keeping in account the CO₂ emissions pertained to the electricity production, the exigency of mitigating climate change brought an evolution in the energy sector during 21st century. This evolution is not only caused by the continuous depletion of primary fossil fuel, but it also imparts the quest for consuming sustainable energy. In the third industrial era, the renewable energy (e.g., solar and wind energy) is preferred over the fossil fuel plants at

² Parts per million

grid level as well as at micro-grid level. The transition of energy landscape is under process and is essential to bring sustainable energy to the doorstep of humankind.

The application of renewable sources also bring innovation in the domain of micro-grid. Prosumerism (also known as auto-consumption) is one of these innovations, which tend individuals to generate, self-consume and sell electrical energy from renewable source(s) at their local site. Generally, the prosumers install photovoltaic solar panels or biogas plant along with electricity storage at their local site. Local energy communities also came in to being as an extension of individual prosumers. These energy communities produce and self-consume energy from renewable sources. The prosumers sell excess energy to either national grid or make peer-to-peer energy exchange on low voltage micro-grid. The decentralization of renewable energy production (with its merits and limitations) makes individuals (and local communities) more self-sufficient in their need of energy consumption via micro-grids, which may also solve the problem of price fluctuation.

I.1.b The quest for sustainable energy mix in European Union and in France

In pursuit of having a sustainable energy mix, European Union is actively working in coordination with its member countries by issuing climate change directives. For this purpose, the European Union set short term, medium term and long-term goals for 2020, 2030 and 2050 respectively. As a long term goal, the European Union aims “to achieve net-zero greenhouse gas emissions by 2050 through a socially fair transition in a cost-efficient manner” (European Commission, 2018).

Table I.1 Short Term and Medium Term Targets set by EU to mitigate climate change (Amanatidis, 2019)

Targets	Directives	Climate and Energy Package 2020	Climate and Energy Framework 2030
GHG emission reduction target (below 1990 levels)		20%	40%
Increase in share of energy consumption from renewable sources		20%	27%
Improving energy efficiency to reduce the use of primary energy compared to forecasted levels		20%	27%

During 2019, the GHG emissions was reduced by 24% (as compared to 1990 level) in 27 EU member states; in which the power sector remained a major contributor with a reduction of 15% GHG emissions (European Commission, 2020). The share of final electricity from renewable energy production is 34% in 2020 (Potrč et al., 2021). These achievements set a baseline for EU to move forward for fulfilling the medium-term ambition until 2030.

Table I.2 French multi-year energy programs (Ministry of ecological transition and solidarity France, 2019)

Program	Energy Program 2019-2023	Energy Program 2024-2028
GHG emission from energy consumption (as compared to 2016)	14%	30%
Increase in share of energy consumption from renewable sources (as compared to 2017)	50%	100%
Decrease in the primary energy consumption from fossil fuels (with respect to 2012)	20%	35%

In line with the United Nations and European Union directives, France has also set targets for reducing GHG emissions and increasing the share of renewable energy for electricity production. A study from ADEME³ suggests 47% and 82% share of renewable energy in electricity mix until 2030 and 2050 respectively (ADEME, 2014). In this context, the installed capacity of wind and solar energy has been increasing in France, while the percentage share of fossil fuel power plants has been decreasing since 2017. It can be observed by Figure I.2 that the renewable power production (from hydro-electric, wind and solar) increased by 17%, whereas the non-renewable power production (i.e. fossil fuel) decreased by 23% in 2018 as compared to previous year. Consequently, the CO₂ emission decreased by 27% in 2018 as compared to previous year (as illustrated by Figure I.3).

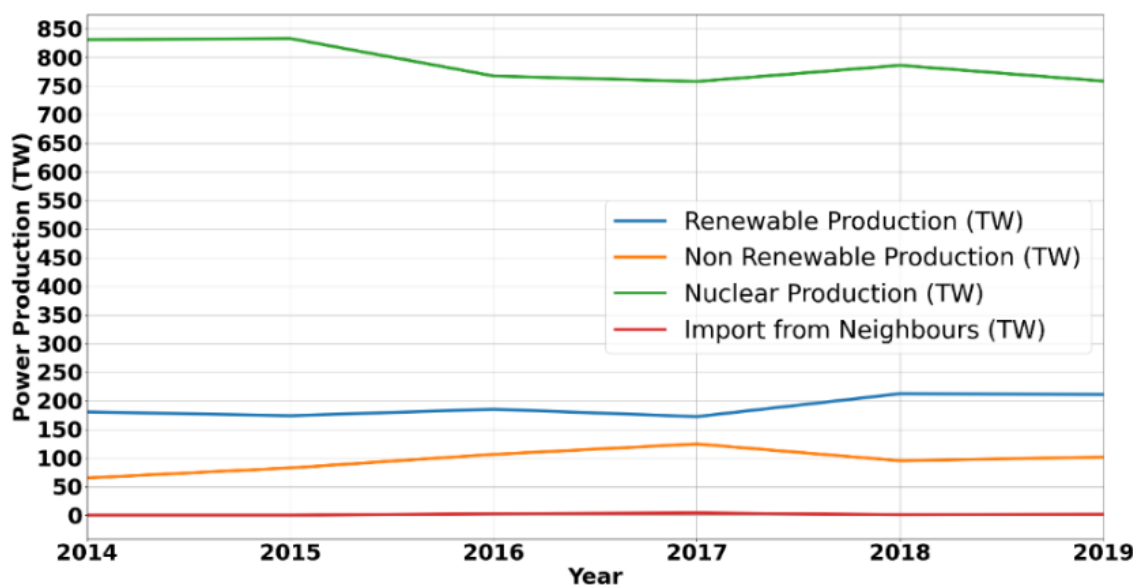


Figure I.2 Annual Power Production (TW) in France⁴

³ ADEME is the French Environment and Energy management Agency.

⁴ eCO2mix power data: <https://www.rte-france.com/eco2mix/telecharger-les-indicateurs>

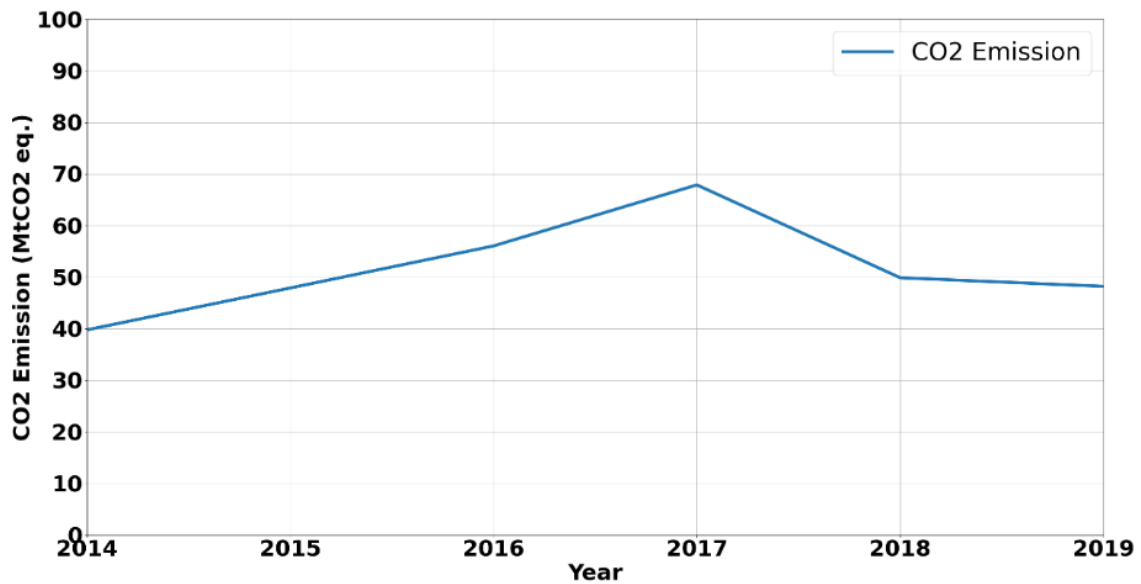


Figure I.3 Annual CO2 Emission (MtCO2 eq.) in France⁵

I.1.c The impact of renewable energy mix on electricity price

The diversity in the marginal cost of generating one unit of electricity establishes a merit order among the electricity generation sources. Despite having high capital cost of installation, hydroelectric energy and nuclear energy have low marginal cost of electricity production due to low operational/fuel cost. Conventionally, these energy sources enter the merit order at the lowest price level. Contrary to this, the fossil fuel power plants have low capital cost of installation, however they have high operational cost. Higher operational cost elevates the marginal cost of one unit of electricity production; therefore these plants follow hydroelectric and nuclear plants in the merit order. The marginal cost has effect on the pricing when the electricity demand increases. As mentioned above, the fossil fuel power plants are connected to keep the energy balance in the grid during peak demand. Therefore, if real time pricing mechanism is followed, the price of electricity significantly increases for electricity production by fossil fuel power plants (as illustrated in Figure I.4 (a)).

Contrary to this, the marginal cost of one unit of electricity production by renewable sources is near zero, except for biomass-fired power plants (Bahar & Sauvage, 2013). Therefore, the renewable energy sources enter the merit order at the lowest level along with hydroelectric energy. Therefore, the price per unit of electricity remains nearly unchanged for same demand. Consequently, the CO₂ emission from fossil fuel power plants is avoided. However, this is not true all the time. The intermittency of renewable sources may push the electricity generation towards fossil fuel power plants to meet the peak demand. Yet, the generated energy quantity from fossil fuel power plants will be low compared to the

⁵ éCO2mix power data: <https://www.rte-france.com/eco2mix/telecharger-les-indicateurs>

conventional merit order. Hence, the CO₂ emissions and cost of electricity will also be lower.

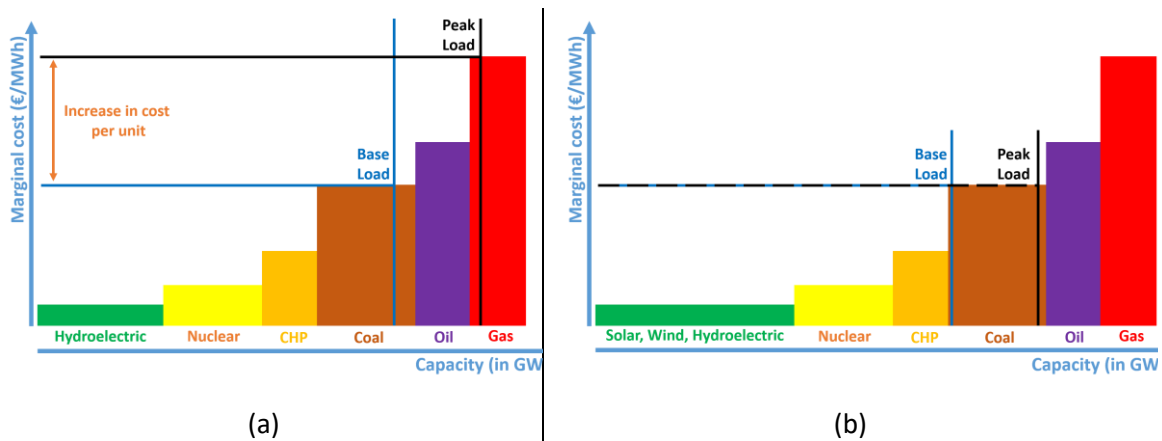


Figure I.4 (a) Conventional merit order (b) New merit order including renewable energy sources (Bahar & Sauvage, 2013)

I.1.d Intermittency of renewable energy and need of energy flexibility

Where the renewable energy mitigates the effects of climate change, it poses constraint to electricity grid. Unlike the conventional power plants, the energy produced by renewable energy is intermittent in nature. e.g. the solar plants can only produce energy in daylight, the wind turbines rotates with wind speed and run of river plants produces electricity with respect to the volume of flowing water.

This intermittency of renewable energy production induces network congestion on the electric grid during peak consumption hours. The imbalance in the grid is harmful to the electricity infrastructure and hence to the reliable supply of electricity. The supply side management has a limited role in mitigating the effects of network congestion. It corresponds to the actions taken to ensure the efficient supply of energy; which includes production from efficient power plants and reduction of line losses by up-gradation of transmission and distribution networks (Karunanithi et al., 2017). It incurs recurring cost; therefore, it increases the price of electricity and is not sustainable for affordable energy consumption.

Energy storage can be deemed as another solution to reduce the impact of intermittency of renewable energy. The solution is very effective for micro-grid e.g., the excess energy produced by PV panels during the day can be stored to use at night. Many energy storage technologies have matured over the years (as illustrated in Figure I.5), however, like supply side management, the energy storage technologies also incur capital cost and recurring operation and maintenance costs. In addition, the energy storage technologies wastes around 10-30% of energy for its own consumption (AL Shaqsi et al.,

2020). The energy transition towards intermittent renewable energy is introducing the world to a new paradigm, in which the demand must be adjusted according to the intermittent supply. In this case, “it is unlikely that storage capacity will fully compensate for supply volatility, at least for short term” (Giet, 2019).

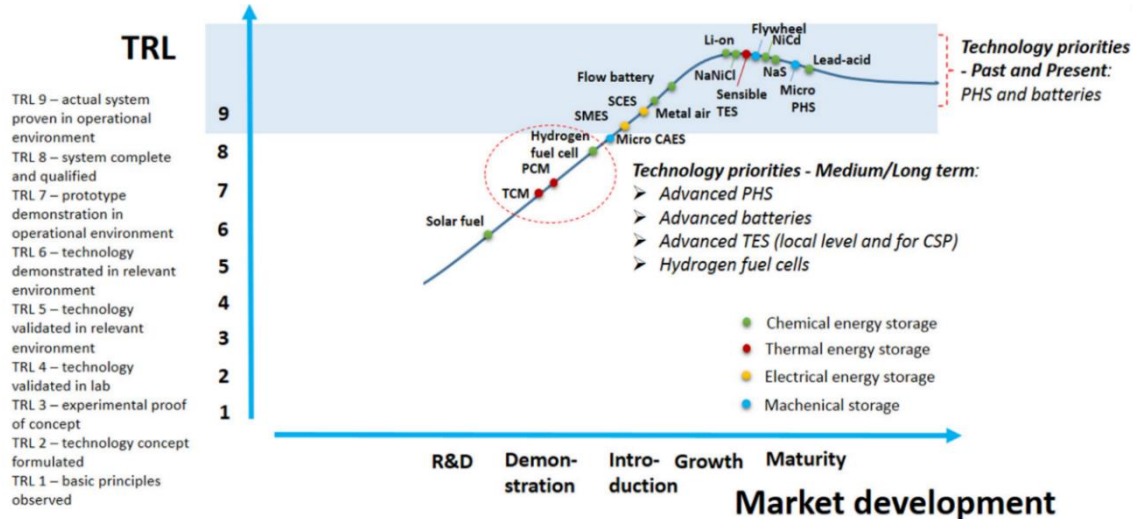


Figure I.5 Maturity of Energy Storage Technologies until 2017 (Nguyen et al., 2017)

In the context of energy transition, the future energy grid will be massively energized with renewable energy. The energy demand is also increasing with time. Therefore, it can be anticipated that the intermittency of renewable energy and the increased demand will frequently pose network congestion on the electrical grid. Where it will bring stress to the electrical network and will be harmful for the sustainable power dispatch to the consumers, it will also bring loss of comfort to the consumer. In addition to this, it will also incur maintenance cost to the electrical grid. With an optimistic outlook of 100% energy transition in France in 2050, ADEME suggests complementing this energy transition with the inclusion of energy storage and demand side energy flexibility. Energy storage is expected to address the intermittency problem of renewable energy sources at supply side. ADEME suggests three types of energy storage solutions namely intra-day, weekly and seasonal. It is estimated that in a scenario of 100% renewable energy in 2050, the energy storage by weekly, intra-day and seasonal solution would be 7 GW, 12 GW and 17 GW respectively (ADEME, 2015).

The demand side energy flexibility tends the consumer to consume energy within the available capacity. In the case of residential consumer, the energy flexibility is significant to curtail the peak by either load curtailment or load shifting. Valley filling can also be practiced in response to a DSO or aggregator signal and is more significant for industrial consumers or aggregated load. In its study of 100% energy transition, ADEME identifies 4 potential loads that can actively participate in the challenge of energy flexibility. These potential loads are water heating, white appliances (i.e., washing machine, cloth dryer and

dishwasher), electric vehicle and space heating. With an estimated annual peak demand of 134 GW, the demand side energy flexibility potential of aforementioned loads would be 3 GW, 3 GW, 7 GW and 25 GW per annum respectively (ADEME, 2015). This brings a potential demand side flexibility of 28.36%. It is cost-effective and brings in consumer participation to meet the challenge of sustainable energy use. The solution is also cheaper than supply side management and energy storage. It can traditionally lessen the impact of intermittency in the same time scale of 10 seconds and 30 minutes.

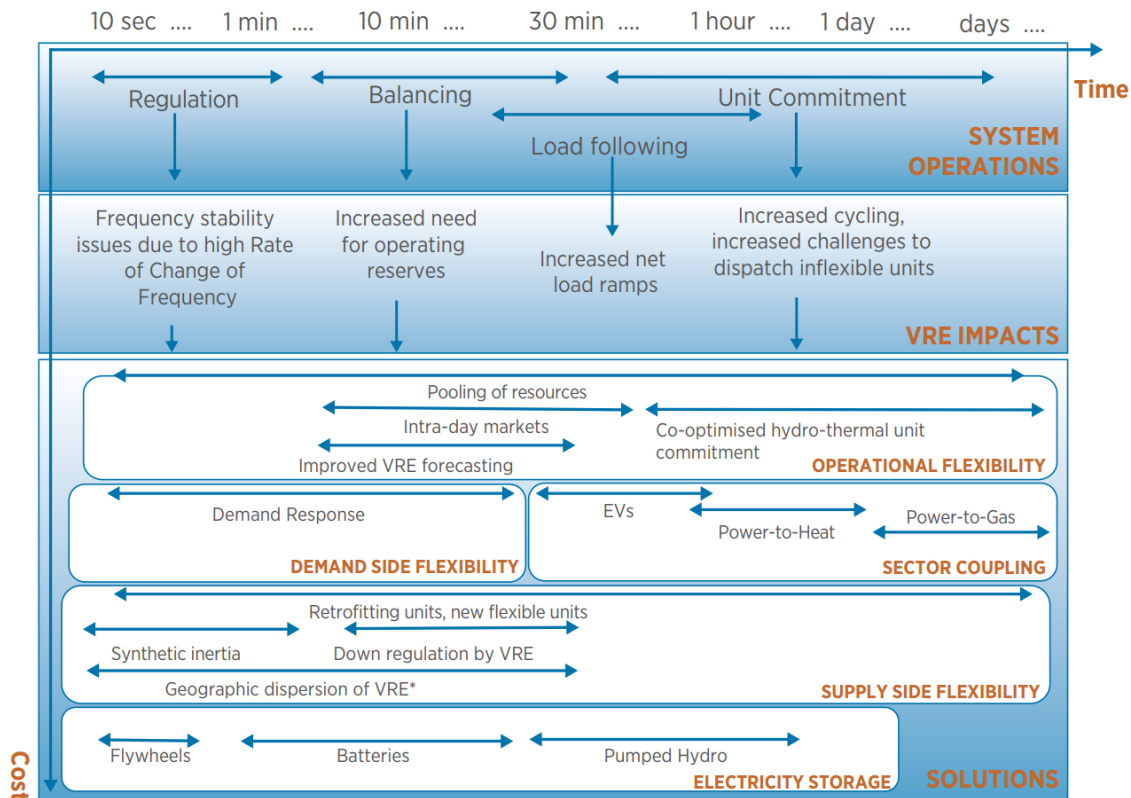


Figure I.6 Flexibility solutions relevant to intermittent renewable production (International Energy Agency, 2018)

I.2 Energy Flexibility and its usual modes of implementation

The definition of the term *energy flexibility* trickles down from a consumption sector (i.e. Industry, Agriculture ...) to the individual consumer. It is also widely presented by the terms of demand side management and demand response. As the buildings are massive energy consumers, therefore the definition of energy flexibility is generally attributed to buildings. “The energy flexibility of a building is the ability to manage its demand and generation according to local climate conditions, user needs, and energy network requirements” (Jensen et al., 2017).

The history of energy flexibility in the residential sector starts in 1915, when an entrepreneur in Switzerland invented an electric water heater called CUMULUS⁶. It is one of the first DHW (Domestic Hot Water) storage system enabling production during low electricity price hours. This invention was commercialized in 1923 for private housing. The necessity of energy flexibility was thoroughly established in the decade of 1970. The electricity generation mix of USA⁷ was dominated by fossil fuel power plants in 1970s. To reduce the impact of energy crisis in 1970s, the distribution system operators in USA introduced energy flexibility programs. The need of energy flexibility is further catalyzed by the phenomenon of climate change. Hence, the energy flexibility programs are categorized into energy crisis era programs and climate change era programs (Ehrhardt-Martinez & Donnelly, 2010).

The energy flexibility can be implemented directly or indirectly. The term direct or indirect is used with respect to DSO (Distribution System Operator) or aggregator. The direct energy flexibility is achieved due to incentivizing the consumer, whereas conventionally, the indirect flexibility is achieved by offering the consumer a dynamic tariff. The classification of energy flexibility is briefly given in Figure I.7.

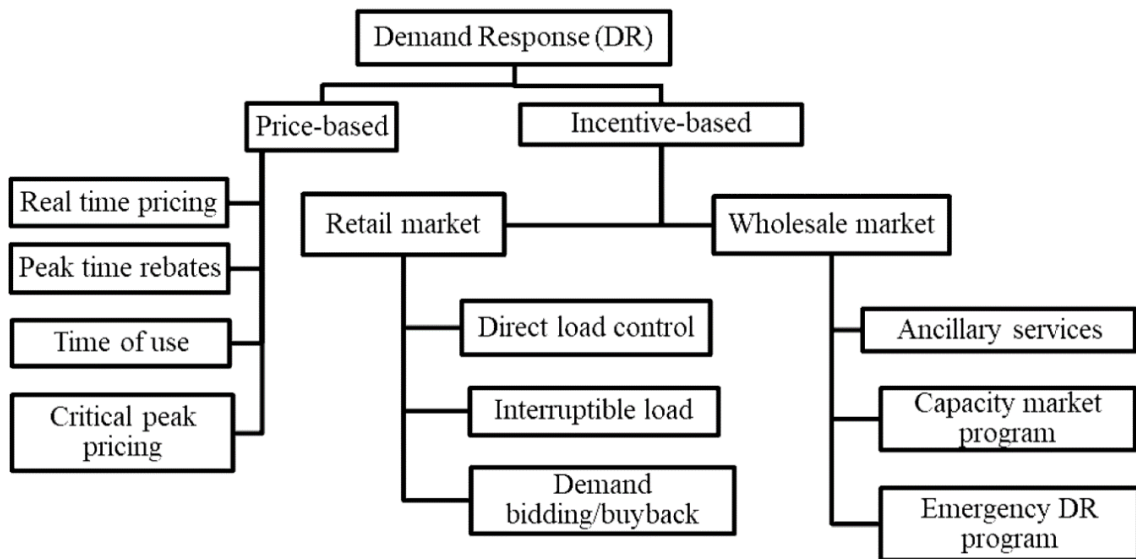


Figure I.7 Hierarchy of demand response (Allasseri et al., 2021)

I.2.a Incentive based direct Flexibility and its limitations

Incentive based programs allow the DSO⁸ or aggregator to implement direct flexibility. It is also known as direct load control. Keeping in consideration the network

⁶ History of CUMULUS : <https://cumulus-lamarque.com/historique/>

⁷ United States of America

⁸ Distribution system operator

condition, the DSO or aggregator may achieve direct flexibility by switching the end-user load in accordance with a predefined contractual agreement. It is intrusive and requires installation of switching devices (e.g., relays and circuit breakers) at the end-user location. In addition to this, it does not consider the participation of energy consumer to implement energy flexibility in real time.

Direct flexibility is significant for industrial sector to curtail big loads during peak hours. For instance, *Energy Pool*⁹ is an aggregator of industrial loads, data centers and hospitals, which is practicing the business of direct flexibility in France. Around 1500 MW of flexible capacity is available to *Energy Pool*; whereby through an optimal decision-making for its clients, *Energy Pool* identifies flexibility potential of the clients, integrates the demand response and offers load adjustments in different markets (Eid et al., 2015). The clients of *Energy Pool* receive specific payments for their participation in either energy based or capacity based trading (Pool, 2015).

Contrary to the *Energy Pool*, *Voltalis*¹⁰ is another aggregator which targets the residential users for load curtailment of electrical heating. The electrical heating accounted for 36% of the annual electricity consumption in French residential buildings during 2017 (ADEME, 2018). When *Voltalis* receives a signal from the transmission system operator, it curtails the electrical heating of its clients in short time intervals through a device installed at client's premises. The clients have a choice to opt out at any time by pushing a button on the device and use their electrical heating normally. *Voltalis* trades the aggregated flexibility in balancing markets and demand response mechanisms of the transmission system operator (Eid et al., 2015). The clients do not receive any financial benefit for reducing their electric heating load, however, they receive a reduction in their electricity bill as reward of participating in demand response.

Despite its robustness, the direct flexibility has certain limitations. The capacity curtailed through direct load control varies from client to client, therefore, certain concerns exist among the clients regarding the privacy and balanced use of direct load control (Haque et al., 2019). Additionally, it normally targets big loads like heating and cooling in households, which brings in loss of comfort for the residents. Where it incurs cost to the distribution system operator, it is also intrusive to the customer's energy consumption. Therefore, the consumers may be reluctant to give the control of their domestic appliances.

⁹ Energy-Pool : <https://www.energy-pool.eu/en/>

¹⁰ Voltalis : <https://www.voltalis.com/>

I.2.b Price based indirect flexibility and its limitations

Contrary to the direct load control, price-based programs are introduced for encouraging the energy consumers to implement energy flexibility themselves. These programs are grouped as indirect load control, therefore provide indirect energy flexibility. They bring the cognitive burden of flexibility to the consumer and impel the consumer through extrinsic motivation to manage their energy consumption for monetary reasons. Therefore, in response to any of the price-based instrument and in accordance with the client's degree of flexibility, the client practices load curtailment or load shifting.

Time-of-use (TOU) is one of the price-based energy flexibility instruments. As indicated by the name, electricity price is set in accordance with the time of use. In TOU tariff, the electricity price is higher during peak hours and is lower during off-peak hours. It encourages the clients to consume less energy during peak hours for economic saving. *HC/HP (Heures Pleines/Heures Creuses)* is a time of use tariff offered by the energy providers in France. More than 40% of the subscribers of *EDF*¹¹ are subscribed to this tariff. The price of electricity is relatively cheaper for 8 hours of a day than the rest of the day. The time slots for these off-peak hours are defined by *ENEDIS*¹² and vary from city to city.

Real time pricing (RTP) is a price-based energy flexibility instrument, which is related to electricity price in real time on hourly basis. Real time pricing is in practice by some DSOs in Nordic countries. Critical peak pricing (CPP) is another price-based energy flexibility instrument. *EDF* has introduced two color coded critical peak pricing mechanisms and sends day ahead notifications to its residential subscribers.

- *TEMPO* is a mechanism of critical peak pricing introduced by *EDF*, in which the electricity price for a day 'D' varies with respect to the color assigned on day 'D-1'. In this regard, the days of normal price of electricity is characterized by white color. The electricity price on blue colored day is cheaper while it is more expensive for red colored day as compared to the white color days (Albadi & El-Saadany, 2007).
- *EJP* is another instrument of critical peak pricing, in which the peak consumption day is color-coded in red and the subscribers pay a low tariff all year round, except for the 21 red days with higher tariff (Bivas, 2011).

Despite that the dynamic pricing is specifically more pertinent for the residential sector, it poses certain difficulties for its effective implementation. It requires installation of a smart meter at consumer site. The smart meter collects energy consumption data of the

¹¹ Electricity of France : <https://www.edf.fr/>

¹² French distribution system operator : <https://www.enedis.fr/>

consumers and log it on the DSO server. This collection can be quite intrusive to a consumer’s personal life as one can learn substantially about the habits and endeavors of a person by looking at energy consumption data. Therefore, the skeptic consumer is reluctant to allow the DSO to install smart meter owing to data privacy and security reason (Goulden et al., 2014).

Another factor that hinders the decision making of a residential consumer is lack of motivation. This lack of motivation occurs owing to limited information and the uncertainty about the potential consequences of the flexibility (Dütschke & Paetz, 2013; Goulden et al., 2014; Krishnamurti et al., 2012). In a longer run, this lack of motivation does not cause any formation of habit in response to dynamic signal pricing. The third constraint is that the consumer often deems a low, invisible and uncertain return from their effort, especially with dynamic tariff (Hargreaves et al., 2010).

Another constraint of consumer regarding dynamic pricing is the everyday life of a household (Dütschke & Paetz, 2013; Goulden et al., 2014). To maximize the benefit from the subscribed energy flexibility tariff, the consumer has to take action with respect to the price signal within the defined time slot of the DSO/aggregator. The consumer cannot postpone these actions if the consumer might not be able to follow the signal subjecting to externalities (absence from home, unfavorable ambient temperature etc.). In consequence, neither the peak consumption on the grid is considerably reduced, nor the electricity bill of the consumer. Table I.1 illustrates the merits and limitations of traditional energy flexibility.

Table I.3 Merits and Limitations of traditional energy flexibility

Characteristics	Direct Flexibility	Indirect Flexibility
Merits		
Nature of signal	Incentive based	Price based
Significant sectors	Industrial	Residential
Type of response	Automated	Cognitive burden for the occupant
Response time	Fast Response	Delayed Response
Guarantee of load control	Yes	Possibly
Anticipated type(s) of flexibility	Load curtailment	Load curtailment, load shifting, Possibly valley filling
Limitations		
Is it intrusive?	Yes	Possibly, but normally no
Installation of switching devices needed?	Yes	No
Target number of loads	A few loads, mainly electric heating for domestic users	A number of loads, depending upon the degree of flexibility of consumer
Intervention may cause loss of comfort?	May be	It depends upon the degree of flexibility of consumer

I.3 The challenge of energy flexibility in residential sector

I.3.a Potentiality of residential sector

According to RTE¹³, the residential sector is the biggest electricity consumption sector in France. During the period between 2017 and 2019, The share of total annual energy consumption of residential sector increased from 36% to 38% (Réseau de Transport d'Électricité France, 2017, 2018, 2019, 2020a). The residential sector remained focal energy consumption sector in energy flexibility programs worldwide (Ehrhardt-Martinez & Donnelly, 2010; Hatton & Charpentier, 2014). Based on the measured power consumption of participating entities throughout France, ENEDIS models a load curve of each consumption sector. These modelled load curves are called profile coefficients¹⁴, whereas “a profile reflects a pattern of consumption or generation, i.e. how an average customer consumes or generates electricity over time” (Réseau de Transport d'Électricité France, 2020b).

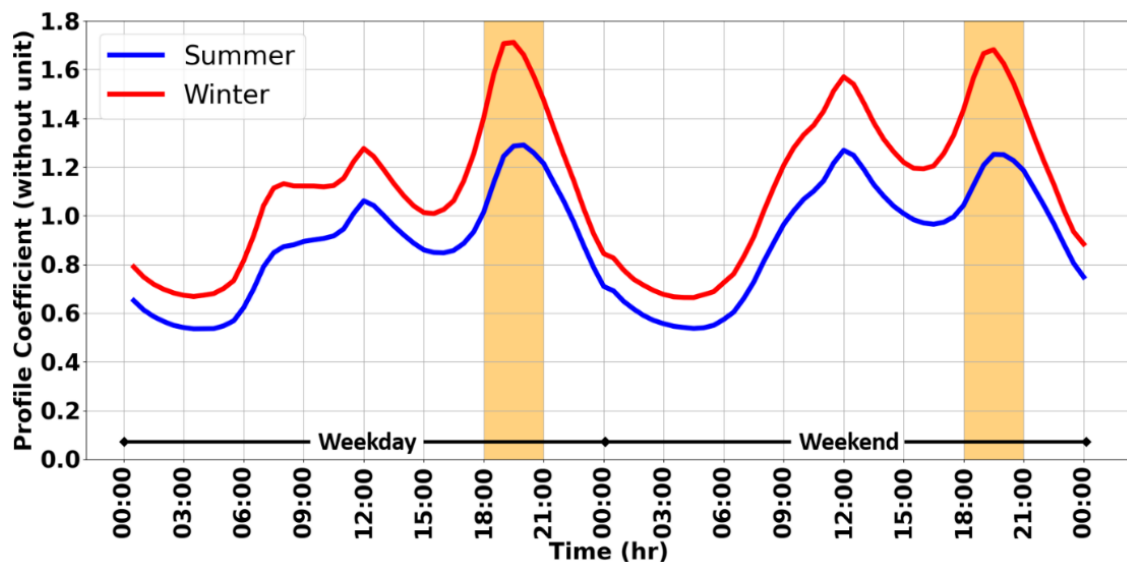


Figure I.8 Modelled average daily load curve of residential sector (ENEDIS, 2021)

The modelled average daily load curve of residential sector is illustrated in Figure I.8. It is based on the historical data of profile coefficient of residential sector. Independent of the season, three major consumption peaks are observed: one for weekday evenings, one for weekend mornings and a last one for weekend evenings. The energy consumption in winter is more than in summer due to the use of electric heating. In terms of energy flexibility, it is estimated that the residential sector can contribute up to 1 GW of (direct and indirect) energy flexibility annually by 2030, which is equal to the combined flexibility

¹³ French national grid operator: <https://www.rte-france.com/>

¹⁴ Profile coefficients: <https://data.enedis.fr/explore/dataset/coefficients-des-profils/information>

expected by batteries and pumped hydro energy storage by the same year (Le Dréau et al., 2019). Which makes it a significant sector to achieve energy flexibility.

As mentioned above, *Voltalis* targets the electric load of residential sector as an aggregator, while giving a choice to its clients to opt out. Implementing energy flexibility on electric heating is not often exercisable. In case of direct flexibility, it may cause loss of thermal comfort to the residential customer. Furthermore, in the case of indirect flexibility, lack of motivation from dynamic pricing and fear of comfort loss may prevent the residential consumers to dial down their radiators during the hours of low ambient temperature. In such a case, it is more interesting to find out other loads that can actively take part in the process of energy flexibility.

I.3.b Identification of potential appliances for energy flexibility

The household appliances are categorized as follows:

- **Continuously consuming appliances:** There are some appliances which consumes energy continuously round the year e.g., refrigerator, freezer etc.
- **Binary appliances:** The binary appliances (e.g. kettle and coffee maker) are ON/OFF appliances and are becomes operational for a few minutes.
- **Multi-stage appliances:** The appliances such as washing machine and dishwasher are termed as multi-stage appliances as they have multiple cycles of operation for each usage.

Here, the first question arises that which appliance has more potential in achieving the goal of energy flexibility. Usually, the intended duration of energy flexibility is in the range of few minutes to few hours. Therefore, an appliance has more potential of flexibility if it offers large amount of flexible power in a short period. It can be observed from Figure I.9 that the maximum annual energy consumption of each household appliances is 500 kWh (except for water boiler). The average duration of operation of each appliance is different. E.g., the fridge is a continuous load, yet it consumes energy in the same fashion as cloth dryer (which is a multicycle load). However, they consume energy in the same fashion. Therefore, it is found difficult to identify potential appliances in a household by looking only at their annual energy consumption.

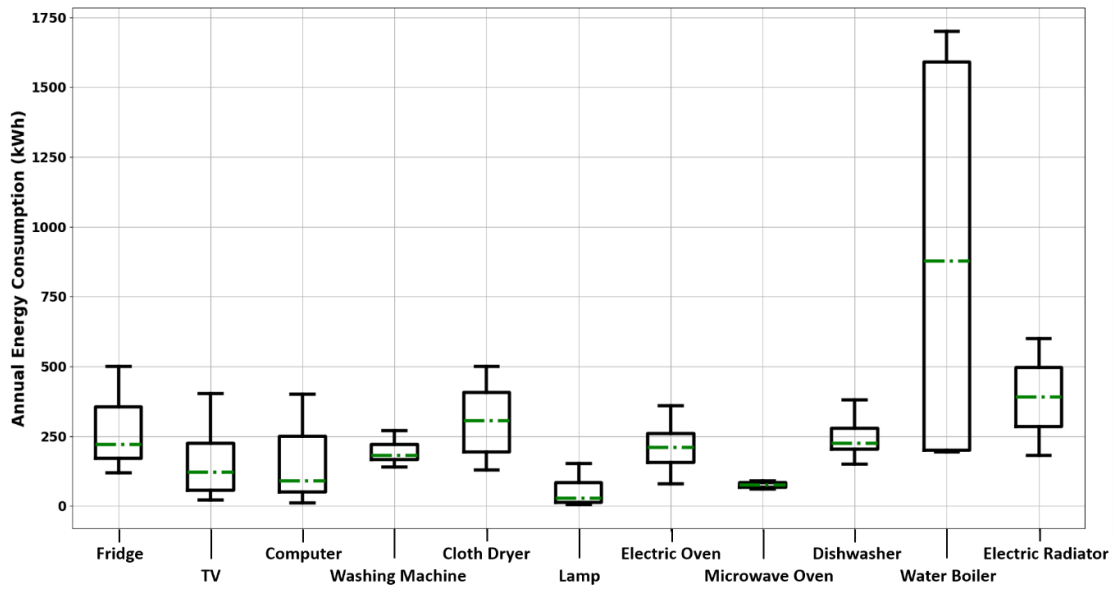


Figure I.9 Annual Energy Consumption (kWh) of household appliances (Energie Douce, 2021; Energie Mobile, 2021; Réseau de Transport d'Électricité France, 2017, 2018, 2019, 2020a)

Figure I.10 suggest that the nominal power of fridge is around 10 times less than the cloth dryer. Therefore, for a short intended period of interventions, cloth dryer offers more potential than fridge. If we recall the study of ADEME regarding the potential appliances for energy, it reflects that around 3 GW of energy flexibility can be obtained from these appliances in year 2050 (ADEME, 2015).

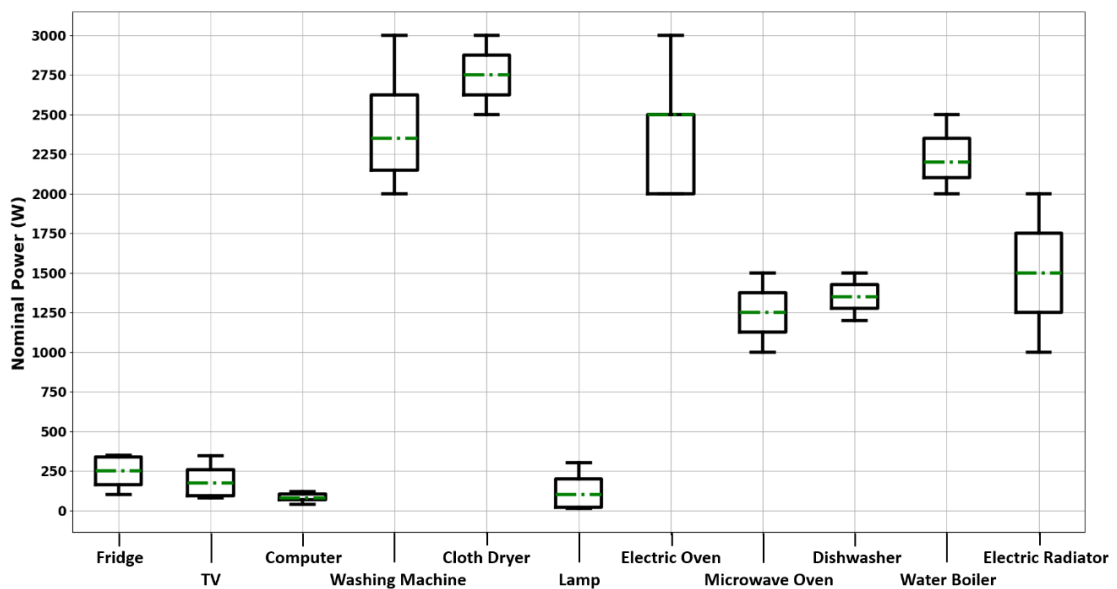


Figure I.10 Nominal power of household appliances (Energie Developpement, 2013; Energie Douce, 2021; Energie Mobile, 2021)

I.3.c Non-monetary indirect energy flexibility: A new mechanism for residential sector

The limitation of direct flexibility and indirect flexibility for residential sector are briefly discussed in section I.2. In the case of dynamic pricing, where the consumer is given a choice to implement energy flexibility, it also brings in the notion that not implementing energy flexibility cause monetary loss. Generally, the residential consumer does not know the purpose of dynamic pricing i.e., reducing peak consumption to meet either intermittency of renewable energy or to avoid fossil fuel generation.

The classical economics suggests that a residential consumer behaves rationally towards the dynamic pricing. However, despite available information on pricing, there exists certain behavioral obstacles and externalities that hinders consumer decision to implement energy flexibility (Sirin & Gonul, 2016). The deviation from rational decision making is not an anomaly, rather it indicates that the decision making structure is influenced by factors like culture, environment and ambient temperature (Gowdy, 2008). In some cases, the DSOs offers dynamic pricing on an opt-in basis, yet the overall consumer participation is low, signaling the weakness of the opt-in design (Schneider & Sunstein, 2017). In this case, it is required to influence the preference of residential consumer in such a way that it offers zero risk to the residential consumer for implementing energy flexibility.

Unlike classical economics based policy of incentivizing the consumer, the behavioral economics has demonstrated that the decision making can be influenced by non-monetary and pro-social impulses (Pratt & Erickson, 2020). A way of doing this is to introduce a competition among the consumers. It has been found that a comparison of home energy usage with a focus on peer consumption decreased energy consumption by 1% to 2% (Ayres et al., 2013). The competition-based programs tend the consumer to win only and therefore are effective during the duration of competition. In case of unlimited duration with no considerable monetary gain, the consumer might abandon its efforts after a period. Thus, these type of programs does not remain sustainable in a long run.

Another important aspect of energy flexibility that is usually neglected in indirect energy flexibility is to bring the human in the loop through providing feedback. In case of direct feedback, the demand response operator is obliged to give feedback to its customers in real time. Contrary to this, indirect feedback is given to the consumer after the flexibility period is finished. A review of 36 energy flexibility programs carried out between 1995 and 2010 has been performed to measure the impact of feedbacks (Ehrhardt-Martinez & Donnelly, 2010). The category called indirect feedback is characterized by global information such as the monthly invoice, provided with a consequent delay of several months. The energy savings linked to this feedback are relatively low, up to 8% with daily

information. The other category corresponds to real-time feedback directly related to the occupants' action. This immediacy is much more valuable in improving occupant behavior, which can reach average savings of more than 10%. New French buildings are subject to regulations related to efficiency (i.e., RT2012), which requires to have an energy consumption monitoring system that informs occupants, at least monthly, of their energy use, by energy type. However, sub-metering by usage or by dwelling is not required if a mathematical disaggregation method is implemented.

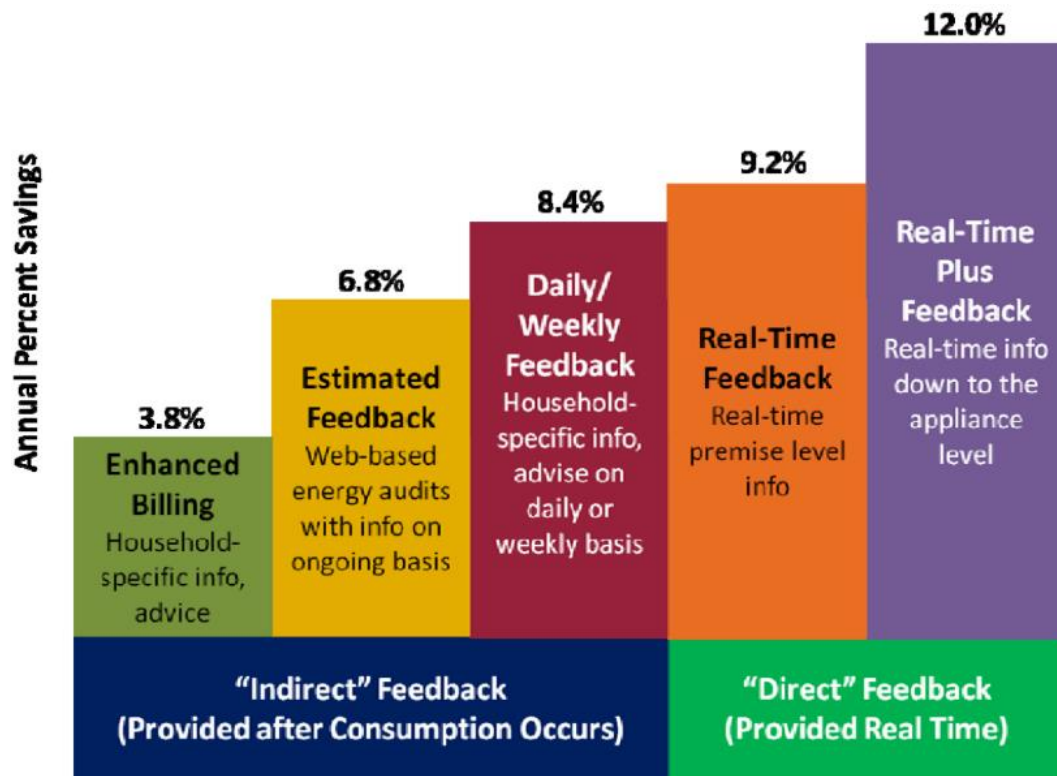


Figure I.11 Average household energy saving with respect to type of feedback(Ehrhardt-Martinez & Donnelly, 2010)

As an alternative to price based indirect energy flexibility, the behavioral economics provides a key to influence the preference of a residential consumer while offering zero risk at the same time. For this purpose, non-monetary and pro-social signals should be sent to the consumer. The ecological signals are found most effective in relatively liberal communities (Costa & Kahn, 2013). This thesis presents the design and results of a conceived indirect energy flexibility mechanism based on behavioral insights. In this regard, a new non-monetary, pro-environmental indirect energy flexibility mechanism is designed and tested with several households. A similar instrument is introduced in November 2020

by RTE in collaboration with ADEME in France. This instrument is called *mon-ecowatt*¹⁵ and it notifies the subscribers to curtail their load for an anticipated peak consumption.

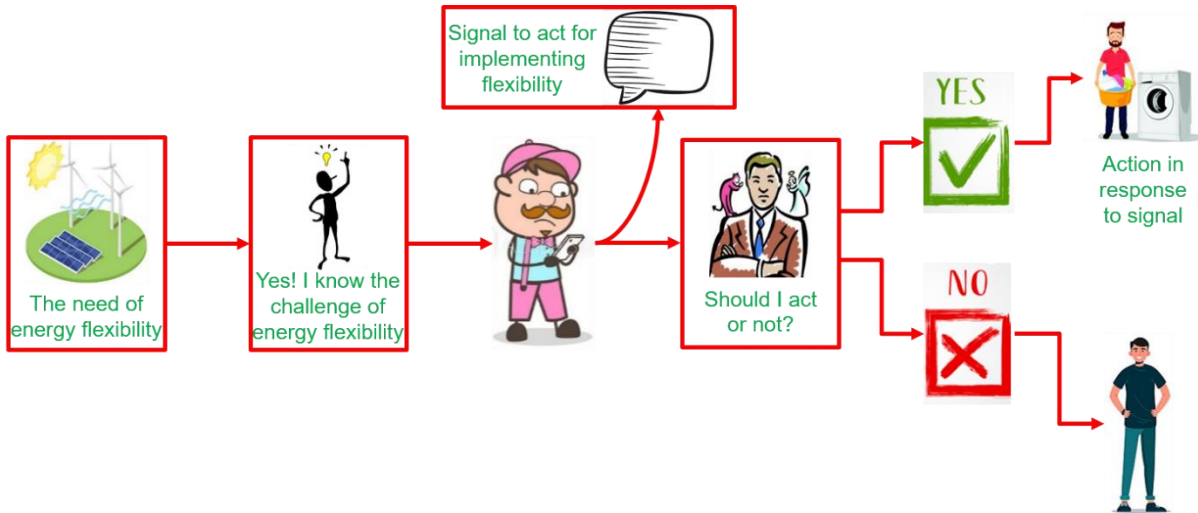


Figure I.12 Behavioral Insight based signal to implement indirect energy flexibility

In this thesis work, pro-environmental nudge signals are introduced to implement indirect energy flexibility with the following characteristics.

- Like price based indirect flexibility, it is non-intrusive.
- No price-based information is given to the households. Instead, the following information is given to the residential consumers
 - o Anticipated network congestion or elevated renewable production.
 - o Certain actions are suggested to do during this period of network congestion or elevated renewable production
- Unlike price based indirect flexibility, the households have zero (monetary) risk to act on the notification. No monetary gain or loss is induced.
- The households are given a system of choice to opt out of their pre-defined commitment with no obligation. *Mon-ecowatt* does not offer any such procedure.
- Unlike *mon-ecowatt*, the households are given indirect feedback of their energy flexibility in graphical form of load curve (as illustrated in Figure I.13).

¹⁵ Mon-Eco-Watt : <https://www.monecowatt.fr/>

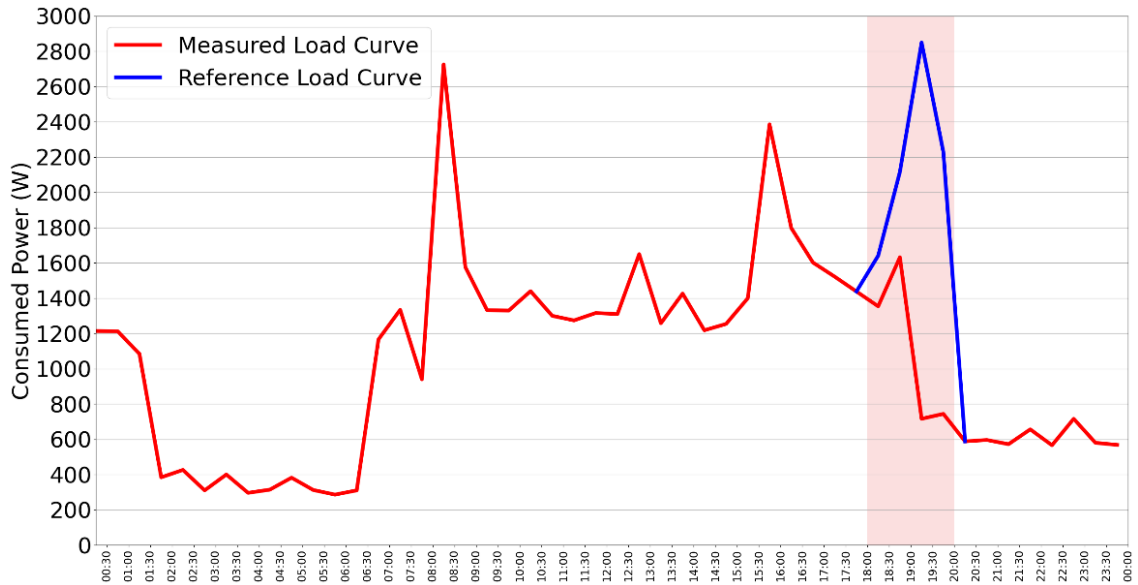


Figure I.13 Indirect feedback in graphical form of load curve

I.4 Conclusion

The power sector is a major contributor of GHG emissions in the world, and therefore is majorly contributing to the phenomenon of climate change. The energy transition from fossil fuel power plants towards renewable power plants is inevitable to mitigate the effects of global warming. However, the renewable energy sources are intermittent in nature. In a future scenario when renewable energy production will be dominant in the electricity mix and demand will be increasing, the intermittency of renewable sources will pose network congestion problems to the electrical grid. The supply side management and energy storage cope with this challenge to a certain extent and therefore requires a performing demand side management.

The residential sector being one of the biggest consumption sectors offers potential for demand side flexibility. Incentive based direct flexibility is significant for controlling big loads; however it is not agreeable for the residential consumer due to unwanted remote switching. The price based indirect flexibility puts the burden of action on the consumer, however, the limited knowledge of dynamic tariff and uncertain return for the effort limits the implementation of indirect flexibility. In this case, it is required to influence the preference of the residential consumer in such way that it presents zero risk to the consumer in implementing energy flexibility. Behavioral economics provide a way to design non-monetary, pro-social and environmental mechanisms of indirect energy flexibility.

This thesis presents an indirect energy flexibility mechanism for residential consumers, based on behavioral insights and indirect feedback. For validation, the mechanism is experimented with a batch of French households using non-monetary and pro-environmental interventions. The following chapters will discuss:

- The significance of behavioral insight (specifically green nudges) in designing a new energy flexibility mechanism
- The design of experiment (including the conceived set of nudges and the architecture for each intervention)
- The day-ahead forecast of modelled national load curve of residential sector
- The measure of effectiveness of nudge signals
- A proof of concept to detect potential appliances for energy flexibility

Chapter II

Case Study: Nudge signals for implementing indirect energy flexibility in French households

The traditional demand response techniques are found to introduce limitations on the decision making of a residential consumer. According to dual process theory, the human brain uses system 2 of thinking in response to dynamic pricing and he poses certain questions to himself, which makes it riskier to act upon the price signal. Behavioral insights is significant in designing a new mode of energy flexibility that tend the energy consumer to use the system 1 of thinking. For this purpose, nudge tool is used that rearranges several choices in a choice architecture and drives individuals towards the desirable behavior.

In this chapter, an overview of the limitation of dynamic pricing are presented. Following this, nudge-based energy flexibility is introduced. The architecture of a nudge based energy flexibility experiment with households is presented in the later part of the chapter.

II.1 Price based energy flexibility in the light of classical economics

The dynamic pricing has been in practice worldwide since many decades as indirect energy flexibility. The non-linear pricing of this tariff is usually based on a number of factors, i.e. the forecasted intermittent production, forecasted peak consumption etc. It is found in the literature that the predominant intention of introducing dynamic pricing is load curtailment. Where it helps the electric grid to avoid network congestion, it also helps the customer to decrease his energy bill in direct proportion to his energy consumption. The impact of dynamic pricing for load shifting is studied by; for instance; (Valenzuela et al., 2012). A dynamic pricing model is also studied “to reduce the overlaps between residential and charging station loads by inspiring the temporal load shifting of electrical vehicles during evening peak load hours” (Moghaddam et al., 2019)

The classical economics presents a theory of rational choice regarding the human behavior towards consumption. This theory states that “all actions are fundamentally rational in character and the people calculate the likely costs and benefits of any action before deciding what to do” (Browning et al., 1999). It constructs a theoretical model of human, which is called as ‘homo economicus’. In a choice paradigm, the ‘homo economicus’ behaves rationally for his self-interest to pursue his defined goal optimally (Urbina & Ruiz-Villaverde, 2019). The classical economics considers that ‘homo economicus’ has characteristics like flawless rationality, unlimited cognitive capacity, narrow self-interest and preference consistency.

The cognitive burden of understanding and acting upon the tariff of dynamic pricing rests on the consumer. It has been observed that the consumers show cognitive difficulty in understanding nonlinear price systems i.e. dynamic pricing (Bartolome, 1995; Ito, 2014). The consumer either lacks information or is unable to process all the information of time varying electricity price in real time. The homo-economicus model suggests that the consumer responds to the actual marginal price; however it has been found that a rational consumer responds to expected marginal price (Ito, 2014; Severin Borenstein, 2009). The expected marginal price is the one that is perceived by the consumer, and it may differ to a larger extent from the actual marginal price. It is also observed that in certain situations despite having updated information, the people do not know what to do in response to a signal. Despite their will to decrease electricity consumption, they do not know which appliances should be effective for either load curtailment during peak hours or load shifting towards off-peak hours. Therefore, they are unable to act upon the signal.

Another challenge comes when the consumer is uncertain about the potential consequences of the flexibility despite having proper information of dynamic tariff. The classical economics deals this situation with the ‘expected utility theory’. It states that the decision maker chooses between risky or uncertain prospects by comparing their expected utility values. Through dynamic pricing, it is expected that the high price during peak hour prompts the consumer to implement energy flexibility (usually in the form of load curtailment). However, any exogenous factor (e.g., extreme ambient temperature) may lead the consumer to think otherwise. If a consumer does not make any effort (i.e., load shifting or load curtailment), the consumer might bring undesirable impact towards the electric grid. If the consumer is uncertain about the consequences of the flexibility, the consumer might not choose (or act upon) the dynamic pricing as it seems too risky.

In conclusion, the dynamic pricing acts as an extrinsic motivator to prompt the consumers for implementing energy flexibility. It is found that the households are likely to exhibit energy flexibility in response to higher prices, yet the dynamic pricing has modest to substantial impact to achieve energy flexibility (Faruqui & Sergici, 2010). The studies show that dynamic pricing is riskier to respond towards the real time need of energy flexibility, mainly due to cognitive limitations of the consumer. Though the homo-economicus model fits in the case of classical economics, yet a transition from homo-economicus towards homo-sapiens is required to drive the residential consumers towards energy flexibility.

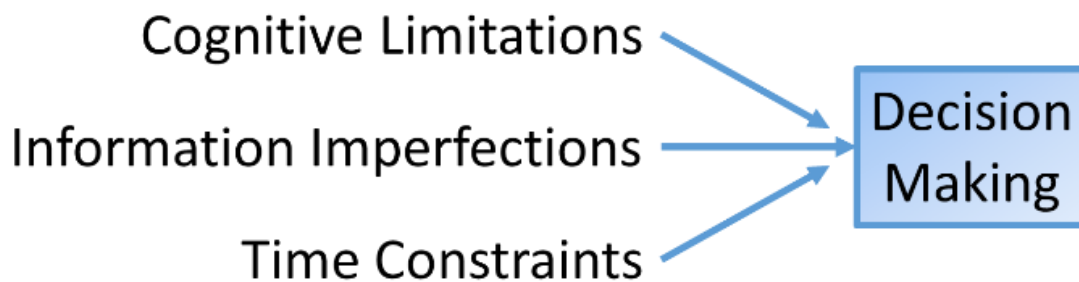


Figure II.1 Factors pertaining to decision making

II.2 Behavioral insights for indirect energy flexibility

According to behavioral economics, the human thinking is represented by ‘dual process theory’. This theory classifies human thinking in two types of systems namely system 1 and system 2. System 1 (of thinking) operates unconsciously, automatically, with little or no effort and with high capacity processes; whereas system 2 (of thinking) operates consciously, in control, with high effort and low capacity process (Kahneman, 2011). Table II.1 illustrates some attributes of dual system of thinking.

Table II.1 Attributes of dual system of thinking (J. S. B. T. Evans, 2008)

Characteristics	System 1	System 2
Consciousness	Unconscious	Conscious
	Implicit	Explicit
	Automatic	Controlled
	Rapid	Slow
	Low Effort	High Effort
	Default Process	Inhibitory
	Holistic, perceptual	Analytic, reflective
Functionality	Associative	Rule based
	Domain specific	Domain general
	Pragmatic	Logical
	Parallel	Sequential

The human behavior is influenced by several factors. It requires to maintain conformity between the actions or feelings with values, goals and public perception of personality (Bénabou & Tirole, 2006). The economics suggests that individuals respond to extrinsic motivators, e.g., monetary incentives, loss aversion etc. However, according to certain studies, the extrinsic motivators undermine the intrinsic motivation of the consumer, has limited impact on treatment and negative impact on the persistence of treatment (Bénabou & Tirole, 2003, 2006).

Dynamic pricing serves as an extrinsic motivator, with the general purpose of driving the consumers to curtail their load. The consumers are susceptible to be influenced by the context, emotions, short-sightedness or any other driver towards irrationality (Shroff et al., 2019). Therefore, the dynamic pricing instigates system 2 of thinking. The cognitive burden of acting on energy flexibility signal is on consumer which enables the consumer to make a controlled and calculated decision after analyzing the situation. The cognitive difficulty of understanding the dynamic pricing makes it riskier to implement following a decision through system 2 of thinking.

Behavioral insights is a toolbox emerged by the interdisciplinary conjunction between economics, psychology and sociology. The consumers are susceptible to be influenced by their immediate environment, emotions and other forms of irrationality. The behavioral insights studies people’s cognitive biases and how they make choices and behave in real life situations. By studying these cognitive biases the behavioral economics serves to develop new methods, mechanisms and other interventions that would help people to achieve what they want (Shroff et al., 2019).

To tend the consumers towards implementing energy flexibility, it is inevitable that certain tools and mechanisms should be developed that instigate system 1 of thinking. In this way, the mechanism of decision making of implementing energy flexibility will be

unconscious and automatic. The consumer does not have to do any mental effort to weigh the possibility of outcome or the opportunity cost of energy flexibility. This mechanism can be used in parallel with the monetary mechanism. Such mechanism can be devised by the utility of behavioral insights, in such a way that the mechanism keeps in account the cognitive ability of a residential energy consumer and external influencing factors.

II.2.a Nudge

Nudge is a technique in behavioral insights which provides indirect suggestions to influence the decision making of an individual or a group. The purpose of nudge is to bring positive reinforcement in the decision making. The most prominent definition of nudge is that “it is any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid.” (Thaler & Sunstein, 2008).

Nudge has been in use by governments around the world to make soft policies. These policies can be regulatory, economic or communication of information. For example, a randomized control trial experiment was performed in UK¹⁶ during 2011. The purpose of this experiment is to see whether and how the repayment rate of overdue taxes could be increased. For this purpose, the control group received usual letter while the treated group received a letter with different text. It has been found that “if the recipient was told that most people in UK paid their taxes on time, repayment rates increased by 5%. However, if they were told about the repayment of the people in their town, the figure increased to 15%.” (van Bavel et al., 2013).

The individuals found it difficult to understand the situation and suffers from an imperfect ability to process new information due to the limitations in cognitive processes (Momsen & Stoerk, 2014). The nudges are distinguished by conventional instruments of policy formation by its virtue of not assuming the economic benefit of rational choice. They are based on dual process theory; where preferences are rearranged in a choice architecture to abridge the thinking process. Nudge instigates the system 1 of thinking for automatic and involuntary decision making. An important element to consider while designing a nudge is the choice architecture offered to an individual or a group for bringing positive reinforcement towards decision making. “The choice architecture refers to the practice of influencing choice by organizing the context in which people make decisions” (Shafir, 2013).

¹⁶ United Kingdom

It is observed through an experiment that nudge is easy to comprehend, resulting in an immediate reduction of energy consumption, whereas it takes longer for an individual to understand the impact of dynamic price on their energy consumption (Buckley & Llerena, 2018). While giving a freedom of choice, the right stimuli are put into focus in choice architecture so that the desired choice is either chosen by default or is easy and automatic to be chosen by the targeted individual or group. The following section demonstrates green nudges and their usability for energy flexibility.

II.2.b Green Nudge

Green nudge is an extension of nudge technique which is attributed to environmental cause and is used for the policy making at national and regional level. The purpose of a green nudge is to use behavioral biases in order to encourage citizens to adopt lifestyles showing a greater respect for the environment (Centre d'analyse stratégique, 2011). An example of green nudge is illustrated in Figure II.2.



Figure II.2 Green nudge to reduce the excessive towel paper use (Lars Biesewig & Annie Krautkraemer, 2018)

The fossil fuel is under the process of phasing out during energy transition while simultaneously integrating renewable energy in the energy mix. The conventional model of energy provision shows that the generation of electricity follows the demand, i.e., higher the forecasted demand, higher will be the need of electricity generation. However, this might not be the case in the future. In a future energy mix predominant in renewable energy, the intermittency of production will pose a constraint on the provision of energy. The continuity of conventional pattern of energy consumption will not only bring network congestion but might also damage the transmission infrastructure.

The distributed generation of renewable energy is considered as a common pool resource. For instance, (Wolsink, 2012, 2014) discusses the characterization and impact of distributed generation as a common pool resource, whereas (Gollwitzer et al., 2018) discussed whether the common pool resource theory “can assist in analyzing and designing sustainable institutions for managing electricity provision in rural mini-grids”. A common pool resource is a certain stock of particular goods from which a flow of resource units can be drawn within its specific boundaries (Ostrom, 1990). The distributed renewable energy generation has the following characteristics which makes it a common pool resource:

- Renewability of stock i.e., energy;
- Scarcity of stock with respect to the optimal location of installation and intermittency of generation;
- Co-production of a common good in the form of energy community (Wolsink, 2012).

The usual pattern of energy consumption depicts that the individuals draw more than the sustainable amount of resource units from the available stock of electricity. This is considered as an overuse of electricity. However, in the future energy mix predominant with renewable energy sources, this usual practice of energy consumption will cause scarcity of energy, either through intermittency of renewable energy or high demand at a certain time. The provision of electricity is excludable in nature i.e., an individual can use the electricity if he/she pays for it. However, the energy scarcity is not excludable as everyone connected to the grid will be affected from this deficiency. In this case, the conventional pattern needs to be altered where the demand should follow the intermittent energy production.

To avoid this socio-technical dilemma in the future, it is inevitable that the individuals should show flexible gesture towards energy consumption. Residential sector being the highest energy consumption sector offers an opportunity that the households should be accustomed with the energy flexibility gestures. In the present, encouraging a responsible and eco-friendly behavior towards energy consumption can contribute to the reduction of GHG emissions (and hence climate change). Though, the benefits of reducing GHG emissions are spread temporally, yet the cost should be borne in the present to be able to enjoy present and future benefits.

II.2.c Review of nudges used in the field of energy worldwide

There are several experiments of indirect energy flexibility based on nudge that has been carried out in the world. Social norms remained effective in the study of non-price-based energy flexibility in most of these experiments. It should be noted that the purpose of these programs was mostly energy conservation. Table II.2 gives an overview of these experiments that were conducted worldwide in residential sector for reduction of energy consumption.

Table II.2 The nudge based indirect energy flexibility experiments in households

Reference	Sample Size (Households) and Location	Type of treatment	Results
(Brandon & Lewis, 1999)	120 UK	Comparative norm, intervention	4.6% decrease in energy usage
(Schultz et al., 2007)	270 USA	Descriptive Norm, Injunctive Norm	Descriptive norm led to decrease of 1.22 to 1.72kWh/day, boomerang effect disappears with inclusion of injunctive norm.
(Costa, 2010)	35000 USA	Descriptive Norm, Injunctive Norm	On average 2% reduction in electricity consumption (1.7% for conservatives, 2.4% for liberals).
(Allcott, 2011)	600000 USA	Descriptive Norm, Injunctive Norm	On average 2% reduction in electricity consumption
(Ayres et al., 2013)	169000 USA	Comparative feedback	On average 1.2 to 2.1% reduction in electricity consumption
(Allcott & Rogers, 2014)	78887 USA	Descriptive Norm, Injunctive Norm	Average 1-1.3% immediate reduction in consumption, discontinuing nudge leads to a decay of effect by 10-20% per year
(Dolan & Metcalfe, 2015)	569 UK	Social descriptive norm	On average 6% reduction in electricity consumption
(Graffeo et al., 2015)	300 Israel	Social descriptive norm	Social norm led to an intention to conserve energy by 12.1 to 34.4% in treated group as compared to control group
(Ito et al., 2018)	691 Japan	Moral Suasion	Moral suasion induced a short-run reduction in peak-hour electricity usage by 8%, economic incentive produced electricity consumption reduction of 14 to 17%.
(Buckley & Llerena, 2018)	240 France	Comparison of nudge and price based	Nudge does not bring welfare loss of individual or group, price based incentive does
(Pratt & Erickson, 2020)	16149 USA	Gamification	Percentage collective change of 4.34 to 11.98 % on treatment days w.r.t control days for households.
(Jorgensen et al., 2021)	143 Australia	Information about peaks in consumption	Reduction in peak consumption by 12% to 20%.

The most prominent non-price nudge field study was performed in USA, in which the energy consumption of a household for a period is compared with the energy consumption of energy efficient neighbors and all neighbors. This comparison is presented in graphical form along with energy conservation tips in the report and is considered as the

descriptive norm. An *injunctive norm* is also added in the report by giving a rating in the form of *efficiency standing*. This efficiency standing categorized the household as *Great*, *Good* or *Below Average*. Both norms are collectively called as Home Energy Report (HER). It is observed from the study that HER causes an average of 2% reduction in electricity consumption (Allcott, 2011), whereas discontinuing the HER leads to decay of energy conservation practice by 10-20% per year (Allcott & Rogers, 2014).

Another significant work in this regard is (Ito et al., 2018), in which the authors studied the effect of *moral suasion* and *economic incentives* for energy conservation using randomized controlled trial in households. The habituation and dishabituation is also studied. It has been found that moral suasion induced a short-run reduction in peak-hour electricity usage by 8%, economic incentive produced electricity consumption reduction of 14 to 17%. In the case of moral suasion, “the treatment effect diminishes after repeated interventions but can be restored to the original level by a sufficient time interval between interventions. Economic incentives induce larger treatment effects, little habituation, and significant habit formation” (Ito et al., 2018).

(Buckley & Llerena, 2018) played demand response as a common pool resource game in a lab experiment to do a comparison between nudge and peak pricing-based interventions. It is found that in the absence of an energy conservation policy, the individuals do not achieve socially optimal level of consumption. “The nudges were quick, easy to understand and resulted in an immediate reduction in consumption in the period following initial feedback. On the other hand, individuals took longer to understand the effect of the increased price on their consumption and so took longer to integrate it into their decision making process” (Buckley & Llerena, 2018).

The impact of load curtailment signal complemented with two types of incentive schemes is studied by (Llerena et al., 2021). An honorary contest is held in one group with an incentive for social cause. The subject that consumes lowest energy during the predefined load curtailment timeslots was considered winner. As incentive, the reward would be given to the NGO (non-governmental organization); chosen by the winner. In addition to this, the name of winner was publicly displayed on the welcome display of the building. The other contest was monetary in nature; in which half of the reward was given to the winner and the rest was distributed among other subjects. The results show that honorary contest bring a significant load curtailment by the subjects. Whereas the monetary contest seems to have no impact for load curtailment, albeit a chance of winning money. This experimental study brings the idea of testing nudge-based energy flexibility in residential sector. The types of nudges used in the experiment and the architecture of nudges is given as follows.

II.2.d The Set of Green Nudges Used In the Experiments

There are several ways that can be used to design a choice architecture for nudging. For this experiment, a set of green nudges are conceived. The characteristics of this set is given below, whereas the information about each nudge in the set is given in Table II.3.

- Unlike most of the past experiments, no social or comparative norm is used as nudge. This is a peculiarity of this experiment.
- Unlike the past experiments which usually relies on extrinsic motivators, the purpose of this experiment is to nudge the intrinsic motivation of residential consumer.
- Feedback is added as a nudge in this experiment. This can be considered as a novel nudge.
- The nudges are designed in such a way that the mechanism for each intervention should be simple, equitable, have low or negligible cost and have no risk for the subjects.

Table II.3 The nudges used in the experiment

Sr. No.	Type of Nudge	Remarks
1	Information ¹⁷	Since a common residential consumer is not aware of the forecasted grid condition, therefore information is given to the subjects of treated group of this experiment. The information pertains either about the forecasted network congestion or about the elevated renewable energy production.
2	Commitment	The commitment nudges tend the individual or group to remain behaviorally consistent by maintaining their commitment.
3	Feedback	Feedback of each intervention is given back to the subjects of treated group, as described in detail in section II.3.c.iii. The purpose of this nudge is to introduce salience ¹⁸ in the intervention. It is expected from this nudge that it serves to keep the subject motivated for the future interventions.

II.3 An experiment of indirect energy flexibility in residential sector using green nudges

As mentioned before, the experiments related to nudge based indirect energy flexibility were conducted for the prime purpose of energy saving, rather than energy flexibility. One important aspect of these experiments is to distribute the subjects into control group and treated group. The control group acts as a standard against which the

¹⁷ Information can also be considered as priming. A primed stimulus is unconsciously held in the associative memory of the decision maker to subsequently influence the decision.

¹⁸ A feature salience makes a choice more noticeable than others by emphasizing certain aspects of choice architecture.

impact of treatment given to the treated group is analyzed. Another important aspect of these experiments is giving direct or indirect feedback to the subjects so that they can also know the impact of their action for energy conservation.

This section demonstrates the design of a behavioral science experiment for the purpose of implementing indirect energy flexibility in the residential sector. Contrary to the earlier experiments available in the literature review, our experiment does not use any social norm among peers or ratings for the participating subjects. Among the diverse nudge tools, non-social nudges are chosen. However, like earlier experiments, it includes a control group and indirect feedback towards the participating subjects. The impact of indirect energy flexibility for load curtailment (during peak hours) and load shifting (from peak hours to off-peak hours) is analyzed in the experiment. The section is divided in to following subsections.

- Methodological aspect of the experiment
- Protocol of interaction between subjects and experiment organizers
- Set of green nudges
- Architecture of experiment

II.3.a Methodological aspect of the experiment

A field experiment is conducted to find out the causal effects of real-world behaviors on energy flexibility. Field experiment is defined as “a data collection strategy that employs manipulation and random assignment to investigate preferences and behaviors in naturally occurring contexts” (Baldassarri & Abascal, 2017). Internal and external validities are related to both lab experiment and field experiment. “Internal validity refers to the ability to draw confident causal conclusions from one's research. External validity refers to the ability to generalize from the research context to the settings that the research is intended to approximate” (Loewenstein, 1999). Contrary to lab experiment, a field experiment has the benefit of having external validity from the real-world environment in a large-scale experiment pool. However, a field experiment has less internal validity as compared to a lab experiment. In contrary to lab experiment, some variables (that affects the treatment group) cannot be controlled in a field experiment.

This field experiment is conducted in a randomized control trial mode. In this type of field experiment, the subjects are assigned to one (or multiple) treatment conditions. The effectiveness of treatment interventions is evaluated by comparing treated subjects with those in a control group (or those who received a different treatment). These interventions affect the normal approach of subjects towards the problem under study (in this case energy consumption). The motive of these interventions is to encourage individuals for taking actions to improve their individual and social well-being (Banerje & Duflo, 2011). Figure

II.3 illustrates an example of randomized field trial. The randomized control trial is used to implement energy conservation for moral suasion of subjects in Japan (Ito et al., 2018). The future electricity grid needs energy flexibility at consumer side; therefore, it is necessary to understand the consumer behavior towards this challenge. For this purpose, this experiment is targeted to know the residential consumer behavior towards the challenge of energy flexibility. It should be noted that though the methodological aspect of this experiment is same as that of (Ito et al., 2018), however the approach i.e. the protocol is different.

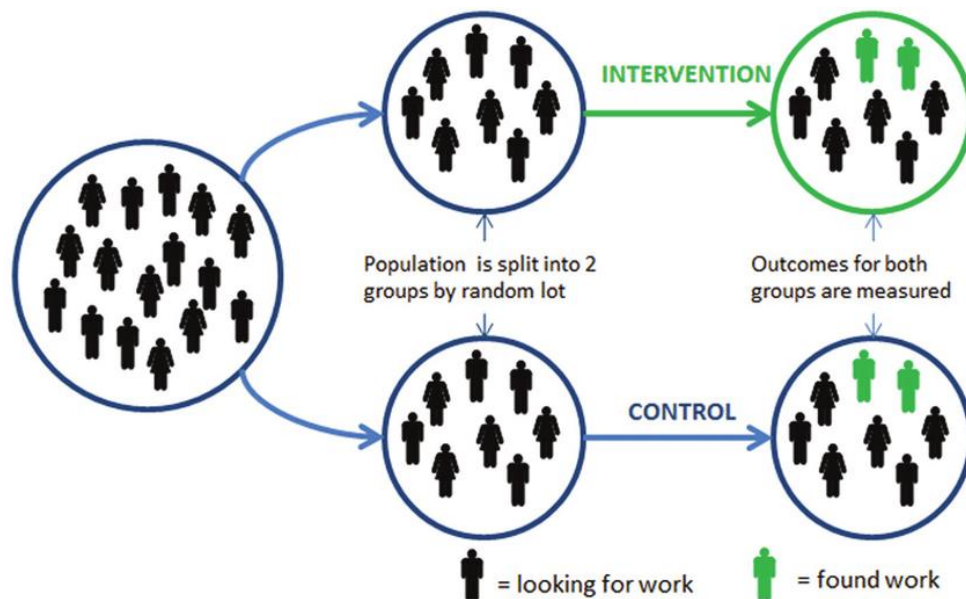


Figure II.3 An example of randomized controlled field trial (Haynes et al., 2012)

II.3.b Protocol of interaction between subjects and experiment organizers

The protocol of the experiment consists of multiple steps that were carried out before and within the duration. It is related to selection and interaction with the subjects of the experiment. It is given as follows.

II.3.b.i Campaign and recruitment of subjects for experiment

The recruitment of subjects was carried out between March and May 2019. A campaign for recruiting subjects was launched by the economists of GAEL¹⁹ lab in collaboration with the social scientists and pro-environment associations. Distribution of flyers, personal interactions and sending emails to the subject already presented in the laboratory panel was the mode of communication of campaigning. The persons listed as

¹⁹ Grenoble applied economics lab : <https://gael.univ-grenoble-alpes.fr/accueil-gael>

student in the laboratory panel were excluded from the recruitment process. The interested candidates were offered to fill in an online form. The gathered information was concerned about the type of housing, the composition of the household, the socio-professional category, the presence of a smart meter, the type of heating, etc.

686 candidates filled the online form to participate in the experiment. At this stage of the recruitment process, the information given to individuals was limited to the launch of a study on energy consumption and the development of renewable energy. A selection criterion is decided in order to select the potential subjects among the candidates. Out of 686 candidates, 175 households are recruited for the purpose of experiment. The selection criteria is as follows:

- The household should be composed of at least 2 people.
- The household should be in the area of metropolitan Grenoble.
- The household should be equipped with *LINKY*²⁰ smart meter. The *LINKY* smart meter permits the experiment organizers to collect the power consumption data from *ENEDIS* as per written consent with the subject.
- The household should not have a specific electricity supply contract (i.e. dynamic pricing).

II.3.b.ii Categorization of recruited subjects into groups

The recruited households are distributed into control group and treated group. The control group comprises of 79 households and treated group comprises of 96 households. The allocation of the selected subjects between the two groups was carried out based on the strictest possible equivalence rule using following criteria.

- The number of people in the household.
- The size of the dwelling.
- The presence of household members in the home at least 3 days per week.
- The number of household appliances in the home.

The treatment group receives treatment in the form of a set of nudges for each intervention. Indirect feedback is also given to the treated group for each intervention. No intervention is made on weekends and French national holidays. The control group does not receive any treatment. It is formed to find out the externalities that are affecting the treatment of treated group during the experiment. It ensures that the energy consumption of treated group is only influenced by the nudge during alert day.

²⁰ LINKY smart meter : <https://www.enedis.fr/le-compteur-linky-un-outil-pour-la-transition-ecologique>

Depending on their treatment assignment, subjects were invited to attend a study presentation session. The organization of the sessions ensured that no communication between subjects belonging to different treatments took place.

II.3.b.iii Remuneration of participation in the experiment

To keep the subjects engaged in the experiment, a small annual remuneration is given at the end of each experiment year. The amount of remuneration is 40€ per subject per year for 3 years of participations. It should be noted that this remuneration is not given as a reward for implementing energy flexibility, rather the remuneration serves as monetary indemnity against a survey and the consent of using subjects' energy consumption data.

II.3.b.iv Types of Alerts

This field experiment is conducted from November 2019 to September 2021. The following types of alerts are sent to the subjects. For the sake of convenience, the day of treatment is hereafter referred as alert day and is symbolized by 'D'.

- Green alert (GA) The purpose of this alert is to nudge the treated group to shift their partial load from evening to the duration between noon and 3:00 PM on alert day 'D'.

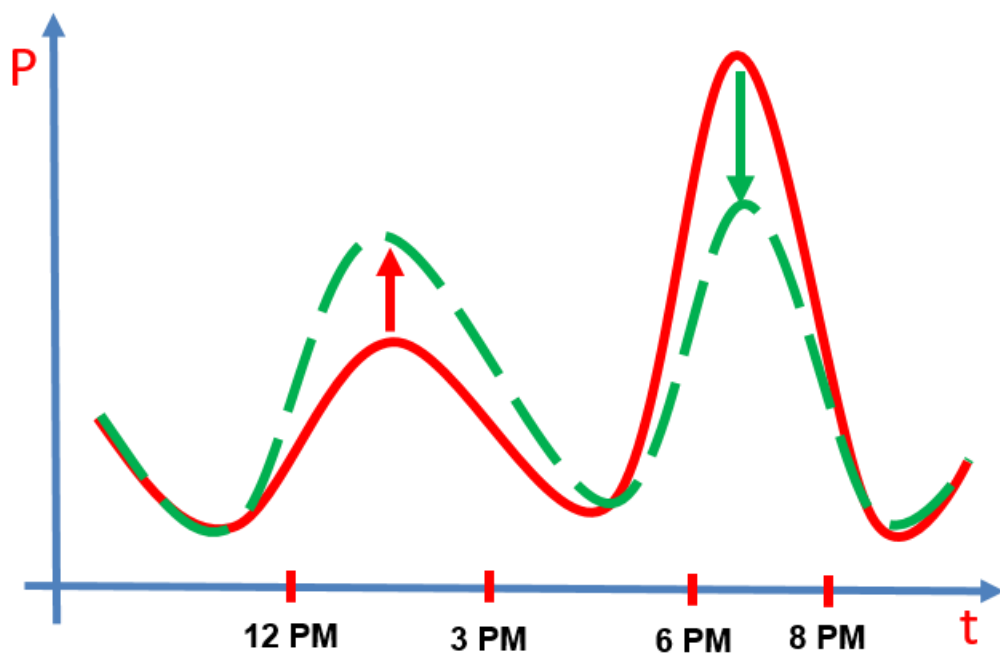


Figure II.4 Load shifting from evening peak hours to afternoon

- Orange alert (OA) The purpose of this alert is to nudge the treated group to curtail their load between 6:00 PM and 8:00 PM on alert day 'D'.

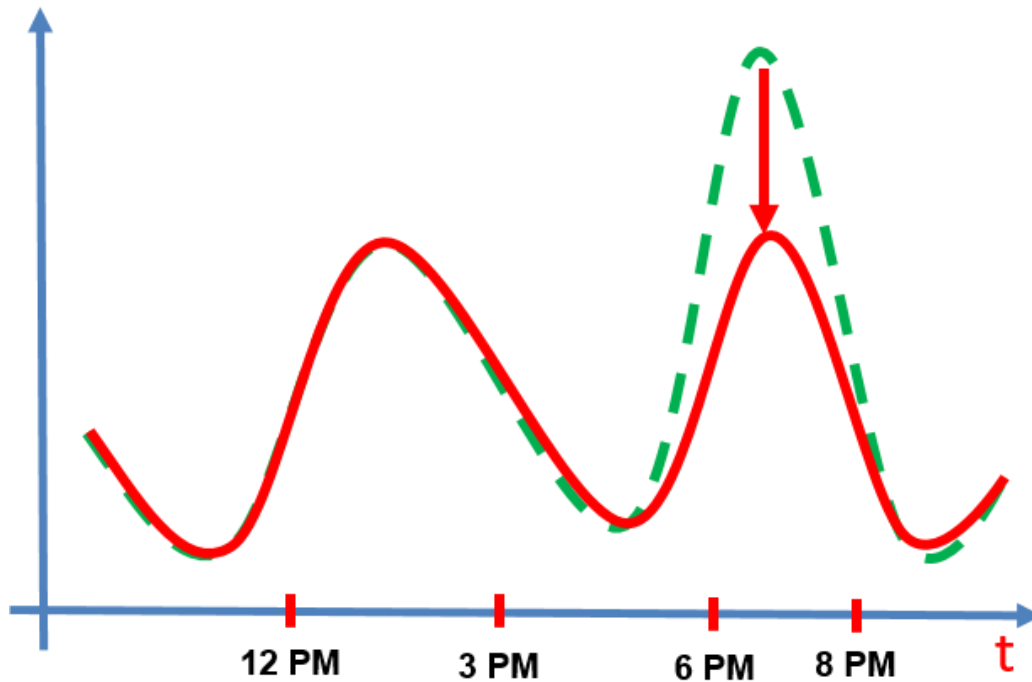


Figure II.5 Load curtailment during peak hours in the evening

II.3.b.v Sessions of Instruction

The presentation of the instructions respected the anonymity of the subjects and no communication between the subjects. To avoid any strategic behavior, especially in terms of the responses concerning the subjects' commitments, the principle of increasing information in several phases was adopted. Each subject was given a personal code to keep them anonymous. The sessions included between 16 and 25 subjects from the same treatment. Each session was organized in a strictly identical manner in 6 phases.

1. The first phase of the instructions consisted of presenting the challenge of energy transition towards renewable energy and consumption by households. Particular attention was paid to the gap between production times e.g., during sun hours in case of solar and consumption times i.e., peak of the household in the evening. To cope with this mismatch, the flexibility of household consumption was highlighted as one of the solutions to avoid the use of polluting thermal installations. Based on this information, the objectives of the study were presented. For the treated group, the aim is to better understand the opportunities for greater flexibility among households by testing recommendations and advice. For the control group, only the considerations of a better understanding of consumption behavior were mentioned.
2. The second phase of the instructions concerned the conduct and organization of the study. The presentation of the study process was identical for both groups. It was indicated that the duration of 2 years of study requires the inclusion of further

phases where subjects would be asked to update their information about their equipment and usage, as well as (only for treated group) a questionnaire for the collection of information about the recommendations and feedback on the periods of intervention. For the control group, the presentation of the organization of the study simply consisted in underlining the need to collect information on the equipment and its uses from the beginning to the end of the study. For the treated group, the presentation of the organization of the study consisted in describing the critical periods for the consumption of electricity.

3. The third phase of the instructions consisted of collecting information on equipment and its use within households. A comprehensive questionnaire was administered before the subjects were informed of the very existence of a commitment principle that they would be asked to carry out in the next phase. This questionnaire, based on a study of the professional literature and institutional reports on household equipment, provides a fairly accurate representation of the dwelling, the set of equipment and their uses for each subject in the study.
4. The fourth phase of the instructions was primarily consisted of the recommendations for the orange alerts and green alerts. In parallel to this presentation, subjects were asked to indicate: i) which equipment was not present in their dwelling and ii) which equipment was not generally used during the intended period of intervention for orange alerts or was generally used during the intended period of intervention for green alerts. This step was particularly important to avoid subjects making commitments to actions that they could not carry out or were already carrying out. It was only based on the collection of this information, in a second stage that subjects were informed of the principle of commitments to be made for each type of period. These commitments were made based on information on the impacts that these could have in terms of energy consumption.
5. For the sake of convenience, the periods of intervention for orange alerts will be called as orange periods, whereas the periods of intervention for green alerts will be called as green periods. The fifth phase focused on the alert system implemented throughout the study, for both the orange and green periods. For each orange or green period on day 'D', this system is based on:
 - a. an alert via SMS on day 'D-1' between 6:45 PM and 7:15 PM
 - b. a second SMS inviting subjects to consult the commitments made at the start of the study and, if necessary, to modify them for the coming period between 7.30 PM and 10.30 PM on day 'D-1'. The consultation of the commitments, and their modifications for the current period, was carried out via a dedicated website with personal and secure access.
 - c. a final reminder SMS on day D, at 7.30 am.

- d. Finally, this fifth phase presented the feedback system on subjects' electricity consumption for the different orange or green periods. Thus, an SMS informs the subjects to view their personal account on the study website where they could find the consumption curves for each day with a period, namely a measured curve, made with the data from the smart meters at a step of ½ hour, and a reference curve which is an estimate of the usual consumption at that time of the day.
6. The sixth and final phase of the instructions concerned the presentation of the operational conditions for data collection, storage and security provisions applied to the data (both relating to the questionnaires and to electricity consumption). This phase ended with the collection of the subjects' consents for sharing of their data from the smart meter according to the regulations in vigor.

II.3.b.vi Survey through questionnaire

A questionnaire is given to each subject to obtain information regarding the presence of various appliances in their respective household. The questions were related to;

- Whether an appliance from given list of appliances is present in the house. The subjects are also asked to mention if any other appliance is present in the house.
- The number of each type of appliances present in the house.
- The time of use of the appliances present in household. For this purpose, a standard day is split into 7 timeslots. In addition to this, the subjects are also asked about the use of appliances during weekdays and weekends.

II.3.b.vii GDPR Consent

A consent is signed with the subjects of the experiment in compliance with GDPR²¹ every year for a period of 3 years. The subjects might opt out of the experiment after a year if they are not interested in signing the consent. Through this consent, the subject gives rights to the experiment organizers to use his/her data. For collecting power consumption data, the subject provides his/her identity number of electricity dispatch point (in French: ID point de livraison) and permits the experiment organizers to obtain the power consumption data of the subject from *ENEDIS* as a third party. This consent binds the experiment organizers to strictly use the data of subject within the limits of experiment. The data includes response to questionnaire, measured load curve, appliance commitments and contact details.

²¹ General data protection regulation (EU) : <https://gdpr.eu/what-is-gdpr/>

II.3.b.viii Website

A dedicated subdomain Etude-Elec²² has been made available to the subjects of the experiment. The sample screenshots of website are given in Annex A. All the subjects have their personal account on the site, which allows them to;

- Respond to the online questionnaire regarding the presence of equipment.
- Change their default commitment for an alert day 'D'.
- View the result of their effort on alert day as feedback in graphical form.

II.3.c Design of set of green nudges

A set of green nudges is introduced in the experiment. Its purpose is to nudge the subjects to implement energy flexibility on alert day 'D'. The timeline of sending these nudge signals gives ample time to the treated subjects for making a decision. This timeline is given in section II.3.d. These nudges are devised to answer the 4 basic questions of a residential consumer i.e.

- Why should energy flexibility be implemented?
- When to act to implement energy flexibility?
- Which appliances should be used to implement energy flexibility?
- Following the implementation of energy flexibility, how did my effort perform?

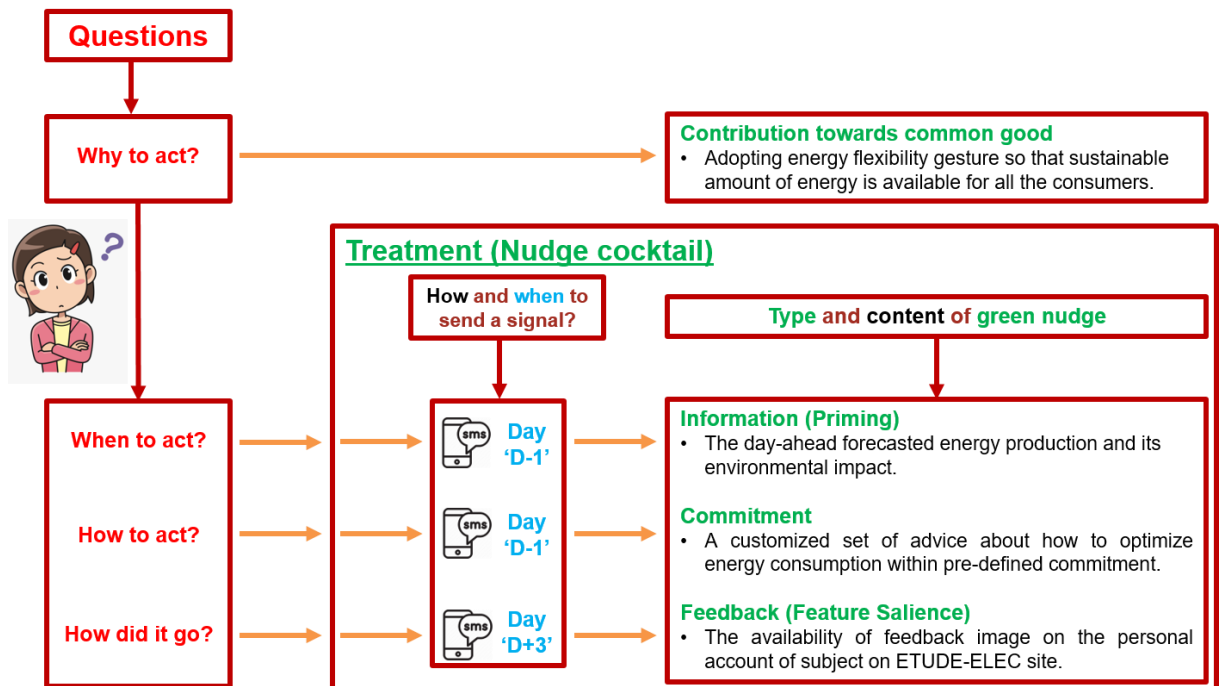


Figure II.6 Curious questions and corresponding designed nudge treatment

²² Etude-Elec : <https://etudelec.univ-grenoble-alpes.fr/>. The website is made for the purpose of study. It may not be available anymore.

II.3.c.i Information

The subjects are primed with an informational message regarding energy production and its environmental impact. “The primed stimulus remains in associative memory and influences behavior in a predictable manner over a certain period of time” (N. Evans et al., 2017). For this purpose, a SMS is sent to the subjects of treated group on day before alert day i.e. day ‘D-1’. This SMS is hereafter called as ‘alert SMS’. The text of alert SMS for each type of alert is as follows:

- Green alert (GA) Tomorrow, from noon to 3 PM, the production of electricity from renewable energy will be maximum.
- Orange alert (OA) Tomorrow, from 6 PM to 8 PM, a consumption peak is expected with maximum use of thermal power plants.

II.3.c.ii Commitment

During the introductory meeting between the experiment organizers and subjects, the subjects filled a questionnaire. This questionnaire contains questions regarding the number of diverse appliances in the household and the use of these appliances during timeslots in a standard day. These timeslots are defined by the experiment organizers e.g., one timeslot is between 6 PM and 8 PM. The purpose of this questionnaire is to make the energy consumption of appliances evident to the subject. Following this, the subjects are given a list of actions that they can possibly implement in the case of receiving an alert. The subjects are asked to give their commitment of using (or not using) one or multiple appliances during treatment period by selecting the actions in the list. In response, the applicants committed appliances in accordance with their prospective degree of comfort during treatment period. In conjunction with alert SMS, a subject specified commitment SMS is also sent to remind each subject of treated group their default commitment on day before alert day i.e. ‘D-1’.

This SMS (a green nudge of type commitment) tends the subject to remain behaviorally consistent in accordance with the commitment. However, the subjects are given a choice to modify their pre-defined commitment via their personal account on the Etude-Elec website within an hour or receiving the commitment SMS. With the help of color codes associated to their default commitment, the subjects are conveyed about which actions should be taken by default, or else they can modify their commitment for corresponding alert. The text of commitment SMS for each type of alert is as follows. For the sake of privacy, the dotted line represents the pre-defined commitment, whereas the asterisks represent the etude-elec website link.

- Green alert (GA) Tomorrow, your commitment is to use To modify, visit *****.
- Orange alert (OA) Tomorrow, your commitment is to NOT use To modify, visit *****.

In addition to the alert SMS and the commitment SMS sent to the treated group on day before alert day ‘D-1’, a reminder SMS is also sent to the treated group on alert day ‘D’. The text of the reminder SMS is as follows.

- Green alert (GA) Reminder: Today, green period from noon to 3 PM
- Orange alert (OA) Reminder: Today, orange period from 6 PM to 8 PM

II.3.c.iii Feedback

This green nudge aims to make ecological choices more noticeable by emphasizing specific aspects of the choice architecture. For each intervention (i.e., alert day ‘D’), the subjects of the treated group are given indirect feedback in graphical form. The subject specific image contains the measured load curve of the subject household on alert day ‘D’, which is superposed by a calculated reference load curve of the same household. The calculation of reference load curve is based on the historical consumption of the household. The formulation of reference load curve will be explained in detail in Section IV.3. Visually, the difference of area between the reference load curve and measured load curve makes the energy flexibility effort noticeable at a glance.

For each subject of treated group, this image (as indirect feedback) is uploaded on his personal account on Etude-Elec website on day ‘D+2’. The subjects of treated group is sent an ‘image visualization SMS’ on the following day ‘D+3’. Irrespective of the alert type, the text for all the subjects of treated group is as follows. For the sake of convenience, the dotted line represents the date of the alert day, whereas asterisk represent the link of Etude-Elec website.

- Green alert (GA) You can view the consumption curve for the period of
- Orange alert (OA), by logging on to *****.

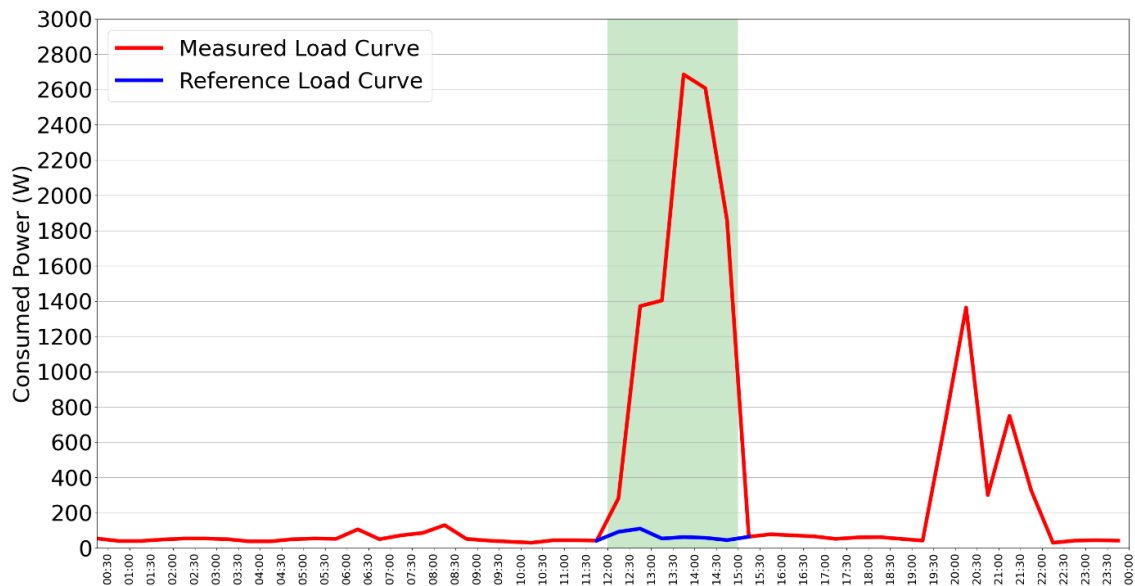


Figure II.7 Feature salience through comparison of load curves

II.3.d Architecture of Experiment

The architecture of conducting this field experiment is set-up by *G2ELab*²³ (Grenoble electrical engineering lab) in close coordination with *GAEL* lab. The architecture is conceived to fulfill the following three prerequisites of experiment.

- How to collect the data of daily measured load curve of the subjects?
- How to identify a day-ahead potential alert?
- What is the protocol of each intervention including giving feedback?

Figure II.8 illustrates the schematic diagram of the experiment. The schematic diagram briefly demonstrates the collection of data, accessing the historic data from ENEDIS as well as accessing the day-ahead forecasted notification via web-services API²⁴. However, it does not demonstrate the protocol of each intervention in the form of timeline. The schematic diagram is described in detail in the sub-sections below, whereas the protocol of each intervention is given from Figure II.9 to Figure II.11.

²³ Grenoble electrical engineering lab : <https://g2elab.grenoble-inp.fr/en>

²⁴ API : Application Programming Interface

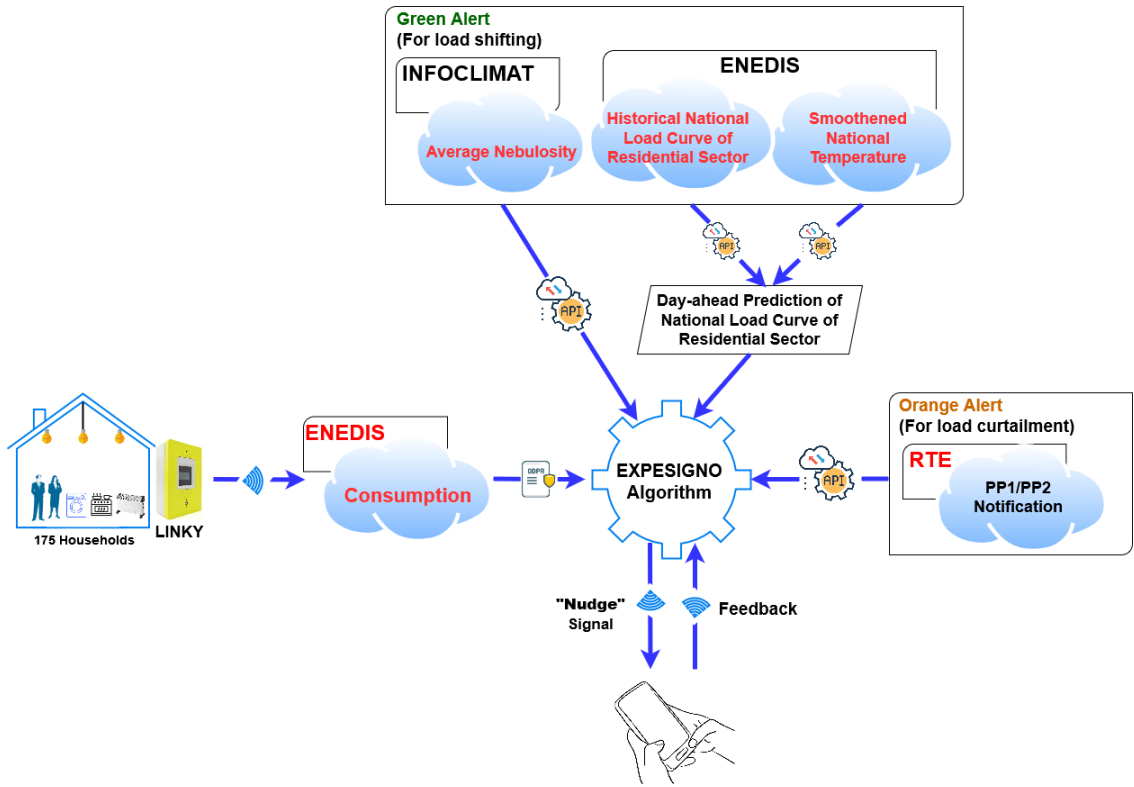


Figure II.8 Schematic Diagram of project "EXPESIGNO"

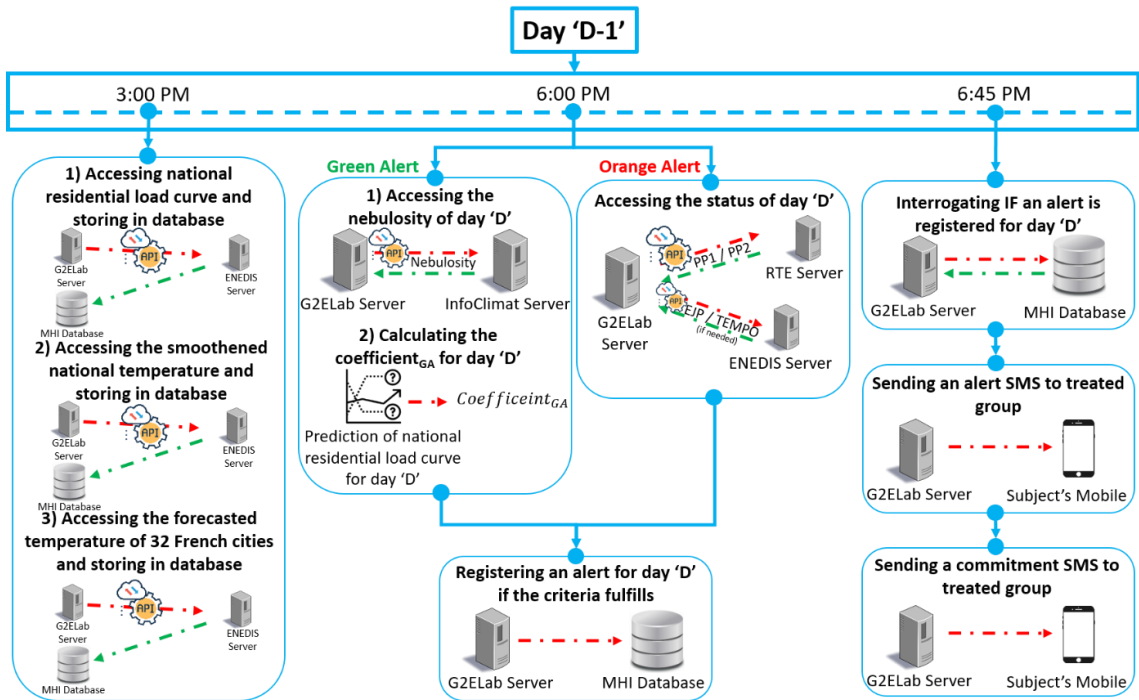


Figure II.9 Timeline illustrating the signaling on day 'D-1'

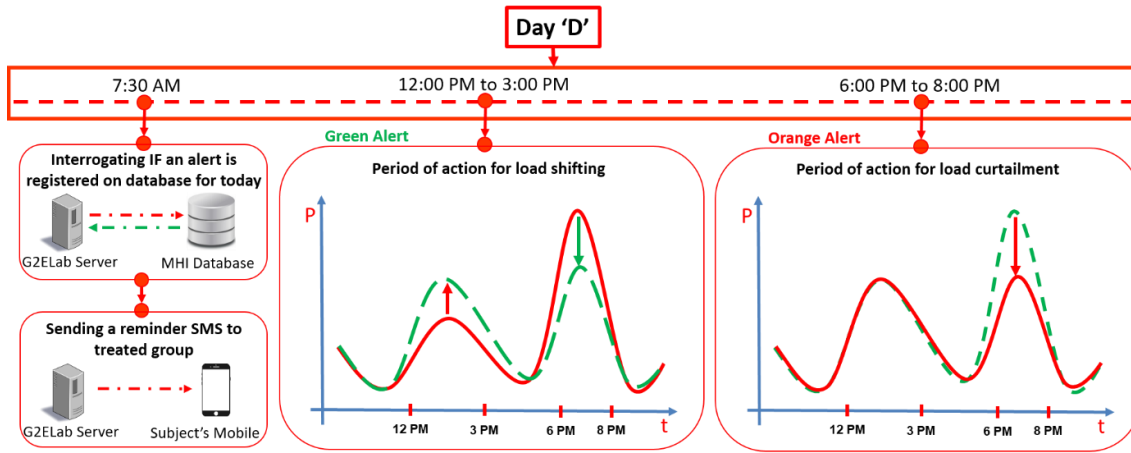


Figure II.10 Timeline illustrating the reminder SMS and period of action on day 'D'

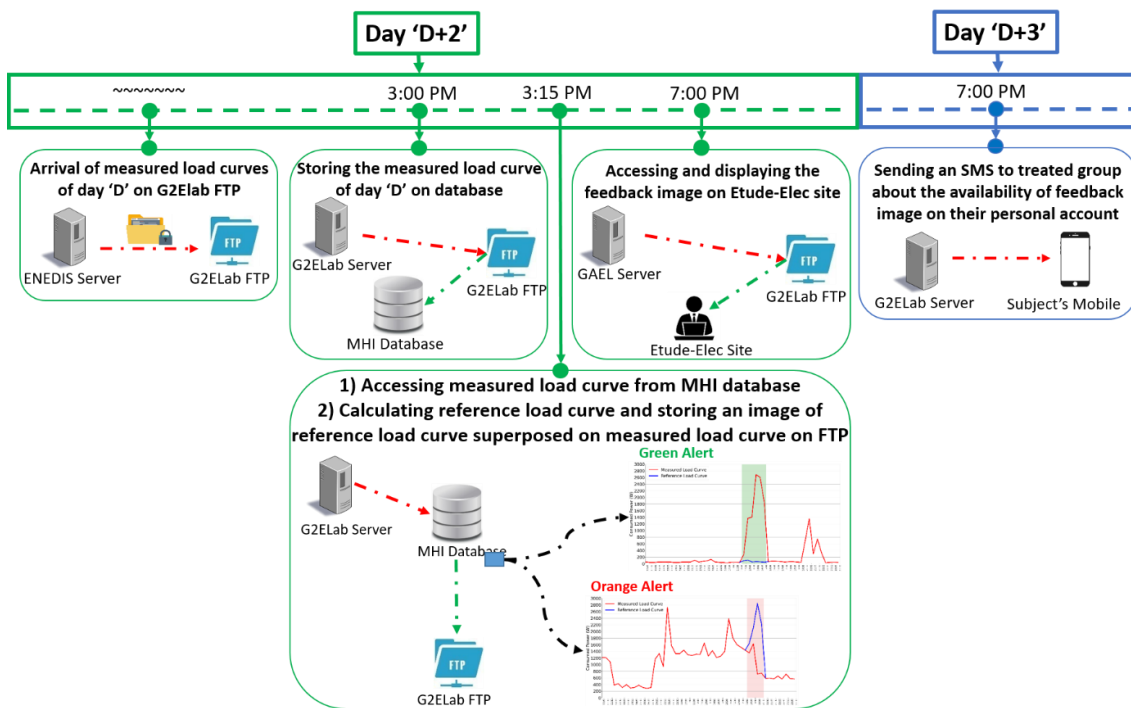


Figure II.11 Timeline illustrating the indirect feedback on day 'D+3'

II.3.d.i Collection of data regarding measured load curve of subjects

It is essential for this field experiment that the experiment organizer must have access to the energy consumption data of all the subjects collected by *LINKY* smart meter. *LINKY* transmits the daily measured load curve of the consumer at a resolution of 30 minutes to *ENEDIS* server (Duplex et al., 2013). In line with a written consent signed by the subjects regarding access to their energy consumption data as a third party, *G2ELab* receives the daily measured load curve of each household from *ENEDIS* in encrypted form. An automatic script decrypts and anonymizes the data as per GDPR consent signed with each household.

II.3.d.ii Identification of a potential day ahead green alert

For the sake of comprehension, it is recalled here that the purpose of this alert is for nudging the treated group to shift their partial load from evening to the duration between noon and 3:00 PM on alert day 'D'. It is pre-defined that the green alert should only be triggered for the day when the production from renewable energy is maximum. There is no such instrument exist on national level that notifies about maximum renewable production, or at least the renewable production more than a defined threshold. Therefore, the following criteria is set for the identification of a day ahead green alert. To register a green alert, both conditions must be fulfilled.

- The average nebulosity in the afternoon of a potential alert day 'D' in Grenoble must be zero. i.e., the sky will be clear and the solar production will be maximum. For this purpose, the forecasted nebulosity by *INFOCLIMAT*²⁵ is accessed via an *API* a day before 'D-1' at 6 PM
- Based on the forecasted (and modelled) French national load curve of residential sector, a coefficient is calculated. This coefficient is a ratio between the day ahead average energy consumption in the evening and day ahead average energy consumption in the afternoon. The forecasting model will be discussed in detail in section II.4. To trigger a green alert, this coefficient should be greater than 1. This implies that the forecasted average evening consumption is greater than forecasted average afternoon consumption. Therefore, the usual energy consumption of the evening can be shifted towards afternoon of the same day (as illustrated in Figure II.4). The coefficient is mathematically given in equation II.1.

$$Coefficient_{GA} = \frac{\sum_{6 PM}^{8 PM} P_{consummed} / t_1}{\sum_{12 PM}^{3 PM} P_{consummed} / t_2} \quad \text{Equation II.1}$$

where t_1 and t_2 are number of hours in the evening and in the afternoon respectively. Their values are 2 and 3 respectively.

II.3.d.iii Identification of a potential day ahead orange alert

For the sake of comprehension, it is recalled here that the purpose of this alert is to nudge the treated group for curtailing their partial load in the evening between 6 PM and 8 PM on alert day 'D'. It is pre-defined that the orange alert should only be triggered for the day when a consumption peak is expected with maximum use of thermal power plants on

²⁵ INFOCLIMAT API: <https://www.infoclimat.fr/api-previsions-meteo.html?id=3014728&cntry=FR>

alert day 'D'. The orange alert is based on the day ahead announcement of *PP1/PP2*²⁶. These days correspond to high electricity consumption determined by *RTE*. The day *PP1* is based on the forecasted day ahead extreme meteorological condition whereas *PP2* is based on forecasted day ahead network congestion.

The announcement of *PP1/PP2* days serves in capacity mechanism to ensure the supply of electricity for forecasted network congestion. The capacity mechanism has been regulated in France since January 2017 to address the growing energy demand while maintaining energy balance in the grid and ensuring secure supply during peak energy consumption. The energy generation companies and load curtailment operators are obliged to commit the availability of their production and load curtailment capacity during these peak periods. By default, *RTE* can attribute at maximum 15 *PP1* days per year and 25 *PP2* days per year (ENOPTEA, 2018) in winter season. For identifying a potential day ahead orange alert 'D', the *PP1/PP2* announcement is accessed via an *API* at 6 PM on day 'D-1'. If a *PP1/PP2* day is declared, the program registers an orange alert for next day 'D'.

It is worth mentioning that during the first wave of orange alert, the meteorological conditions in France were not extreme enough that *RTE* forecasted a congestion on the electrical network for several days. In this case, the rest of the alerts were triggered based on day ahead announcement of *TEMPO*²⁷ and/or *EJP*²⁸. *TEMPO* is a mechanism of critical peak pricing introduced by *EDF*, in which the tariff for color-coded day 'D' varies according to the color assigned on day 'D-1'. The tariff on day 'D' is cheaper than time of use price for blue color, slightly expensive than time of use price for white color and very expensive for red color (Albadi & El-Saadany, 2007). *EJP* is another instrument of critical peak pricing, in which the peak consumption day is color-coded in red and the subscribers pay a low tariff all year round, except for the 21 red days with higher tariff (Bivas, 2011).

²⁶ *PP1/PP2*: <https://www.services-rte.com/fr/visualisez-les-donnees-publiees-par-rte/signaux-pp1-et-pp2.html>

²⁷ *TEMPO*: <https://particulier.edf.fr/fr/accueil/contrat-et-conso/options/tempo.html#/selection-bp>

²⁸ *EJP*: <https://particulier.edf.fr/fr/accueil/contrat-et-conso/options/tempo.html#/selection-bp>

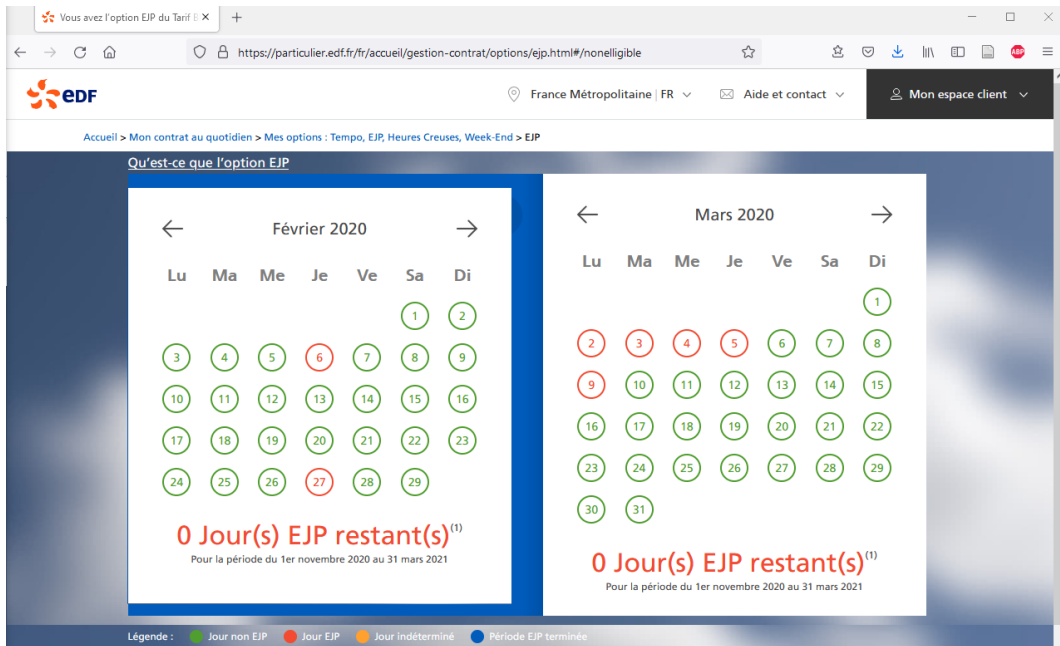


Figure II.12 Calendar of the months for orange alerts based on EJP

EDF notifies subscribers on day ‘D-1’ regarding the day-ahead color of *TEMPO* and *EJP*. Therefore, *EJP* and *TEMPO* are added to the protocol of orange alerts for the months of February and March 2020. Like *PP1/PP2*, the day-ahead notification is obtained via the *APIs* of *EJP* and *TEMPO* at 6:00 PM on day ‘D-1’. The pre-defined condition for a triggered alert based on *EJP* and *TEMPO* is that either of the two instruments notified the day ahead in red color. Figure II.12 and Figure II.13 demonstrate the color-coded days during the months of February and March 2020 for *EJP* and *TEMPO* respectively.

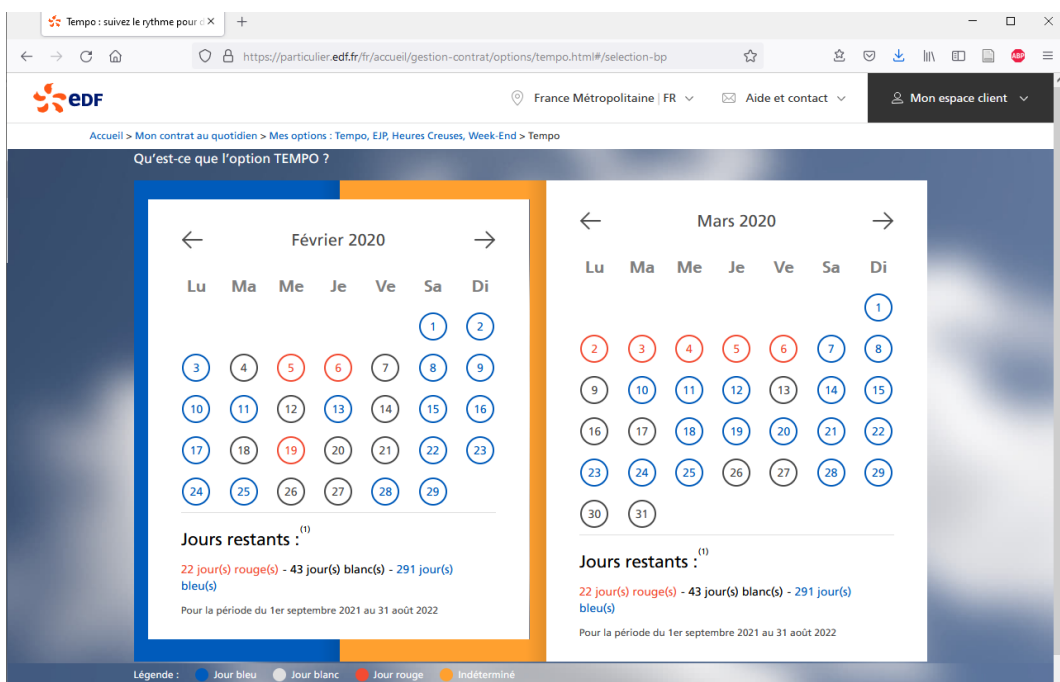


Figure II.13 Calendar of the months for orange alert based on TEMPO for 1st session

The experiment is performed for a period of three years. The alerts are sent during the first two years while the third year is set to find out the habit formation of the households. The timeline is adjusted within the temporal constraints (i.e. the availability of measured load curve by ENEDIS). It is also adjusted in accordance with the observation of (Ito et al., 2018) i.e. to retain the treatment effect by introducing sufficient time interval within consecutive treatments. The triggering of orange alerts is simply based on the *API* notification by *RTE* regarding the forecasted network congestion in the national grid. However, the triggering of green alert is based on two criteria. One of the criteria of green alert is based on the prediction of day ahead residential energy consumption, which is discussed below.

II.4 Prediction of modelled national load curve of residential sector

The criteria for identifying a potential green alert is discussed in the section II.3.d.ii. For the sake of convenience, it is recalled here that one of the criteria is that the value of $\text{coefficient}_{\text{GA}}$ (coefficient of green alert, see equation II.1) should be greater than a threshold value. In this case, the threshold value is 1. The $\text{coefficient}_{\text{GA}}$ is a ratio between the average energy consumption in the evening and the average energy consumption in the afternoon. Since the calculation of $\text{coefficient}_{\text{GA}}$ is based on the day ahead energy consumption (which cannot be measured on day before), therefore it requires forecasting of the day ahead energy consumption.

API for real time prediction of day ahead energy consumption is not available by the French electrical grid operators or a third party. Therefore, a supervised learning model is developed for day ahead prediction of energy consumption. *RTE* provides the historical *gross daily energy consumption*²⁹ data at national level. This data represents the accumulated electricity consumption for all the sectors (residential, industrial, etc.). This might be effective in case of sending a general signal to the consumers of all the energy consumption sectors, however it is more significant if customized signals are sent to the consumers of respective consumption sector, i.e. in our case residential sector.

For this purpose, the historical data of modelled *load curve of French residential sector*³⁰ is used. To refine the prediction, the historical data of *smoothed national*

²⁹ French daily energy consumption : https://opendata.reseaux-energies.fr/explore/dataset/consommation-quotidienne-brute/api/?sort=-date_heure

³⁰ Modelled load curve of French residential sector: https://data.enedis.fr/explore/dataset/coefficients-des-profil/information/?disjunctive.sous_profil

*temperature of France*³¹ is also used. This data is used by *ENEDIS* and is also accessible on their open data platform via an *API*. In addition to this, the forecasted temperature of 32 French cities is also accessed via an *API* of *Climacell* (rebranded as *tomorrow.io*³²). The detail of these *APIs* and their corresponding data is discussed in detail in section II.4.c. Numerous techniques are available in the literature for short term load forecasting. Moghram & Rahman (1989) presented an overview of 5 short term load forecasting techniques namely; multiple linear regression, stochastic time series, general exponential smoothing, state space and Kalman filter. Liu et al. (1996) presented fuzzy logic, neural networks, and autoregressive (*AR*) models in their article. Taylor & McSharry (2007) presented autoregressive approaches such as *ARIMA* and periodic *AR*, or Holt-Winters exponential smoothing and a method based on the principal component analysis (*PCA*). Recently, recurrent neural network (*RNN*) based short term load forecasting has been presented in multiple articles. Bianchi et al. (2017) presents a comparative analysis of *RNN* which includes Long Short Term Memory (*LSTM*) and gated recurrent unit (*GRU*).

The results of different techniques are studied for the purpose of prediction. Normally, *LSTM* prediction gives good results however, it needs a lot of computation, which is not necessarily required for this study. Random Forest regression gives good results as well and needs less computational time than *LSTM*, Therefore, for the purpose of forecasting the French national load curve of residential sector, *random forest regression* model is used. The following section presents the working principal of *random forest*.

II.4.a Random Forest Regression

Random forest is an ensemble supervised learning method for classification and regression. “It operates by constructing a multitude of decision trees at training time and outputting the class that is mean/average prediction (regression) of the individual trees” (Tin Kam Ho, 1995, 1998). Figure II.14 represents the schematic diagram of *random forest regression*.

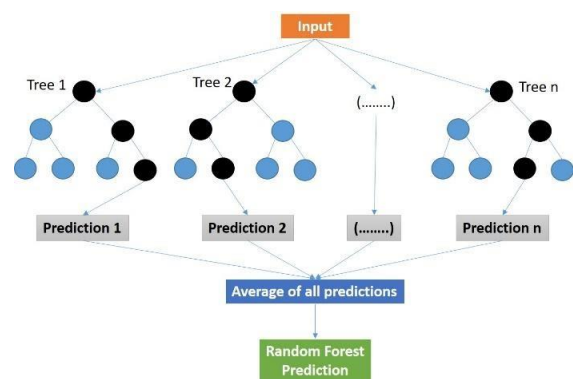


Figure II.14 Random Forest Regression

³¹ Smoothened national temperature : <https://data.enedis.fr/explore/dataset/donnees-de-temperature-et-de-pseudo-rayonnement/table/?sort=-horodate>

³² Climacell (rebranded as tomorrow.io) : <https://www.tomorrow.io/>

II.4.b Regression Model Creation

A supervised regression model has three phases. The first phase is related to the training of a model. For this phase, it is necessary to have sufficient historical data. In case of multiple input variables, the data of all the quantities is needed. Following this, the choice of a suitable regression technique is important. A target variable is also defined in this phase, of whom the prediction will be performed. In the training phase, the historical data of (one or multiple) quantities is used as input (mathematically denoted as X_{train}), whereas the target is defined as output (mathematically denoted as Y_{train}). Using the chosen technique, the model learns that each x input leads to its respective y output (where x is the instantaneous input and y is the instantaneous output).

The second phase is related to validation of model. Usually, a fraction of training data is kept for validation. Since the data is in vectorial form, therefore rather than having a simple arithmetic difference, a validation indicator is chosen in this phase. For the given input data X_{val} , the model performs a prediction \hat{Y} . The difference of prediction \hat{Y} with the target Y_{val} according to the chosen indicator gives us a measure of how well the model is performing. Certain measures can then be taken to improve the model if the model seems to be not working up to the expected standards. The third phase is related to the real time prediction, where a validated model is used to do the prediction \hat{Y} for given input data X .

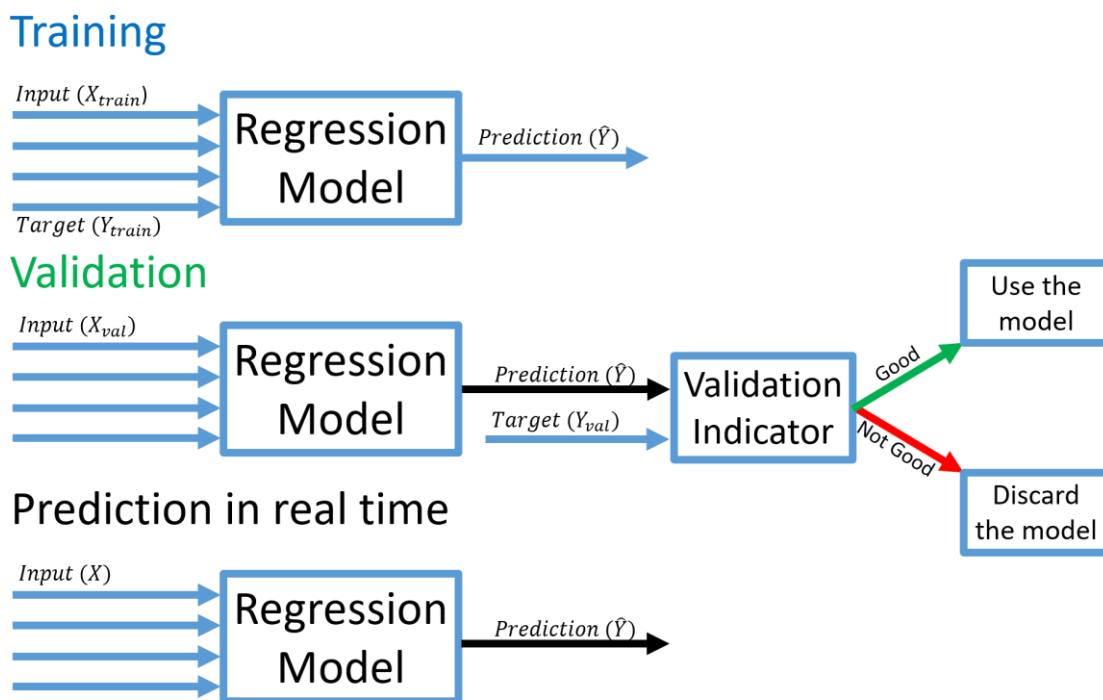


Figure II.15 Principal of regression model creation, validation and use in real time

For the purpose of prediction, *scikit-learn* package (Pedregosa et al., 2011) of *python* language is used. The script access data via *APIs* of *ENEDIS* and *Climacell* (weather

forecasting enterprise) and store it in a database. It should be noted that the *random forest model* created by *scikit-learn* method does not update regularly. Therefore, with the availability of sufficient new data, the python script trains a new model on whole available historical data. For this purpose, the same script runs daily to make a model with ensemble of available data. Another *python* script runs automatically at 6:00 PM every day and make the prediction using the *random forest* model based on the available input data.

II.4.c Third Party APIs used for Forecast

Third party *APIs* are used to obtain updated data from *ENEDIS* and *Climacell* on regular basis. These *APIs* and the data obtained from them are briefly explained below.

II.4.c.i Profile Coefficient of Residential Sector

ENEDIS constructs national profiles of the energy consumption sectors in France. The data of these profiles is published as open source and is validated by energy regulation commission. A profile coefficient depicts a form of consumption or production of electricity over time. Each consumption profile is modelled based on the energy consumption data of a sample of customers (in this case, residential consumers) selected throughout France. A profile represents average behavior of a group of users and reflects how an average individual in the group consumes electricity with the passage of time (Réseau de Transport d'Électricité France, 2020b). Figure II.16 illustrates the profile coefficient of residential sector, displayed on the data visualization platform of *ENEDIS*.

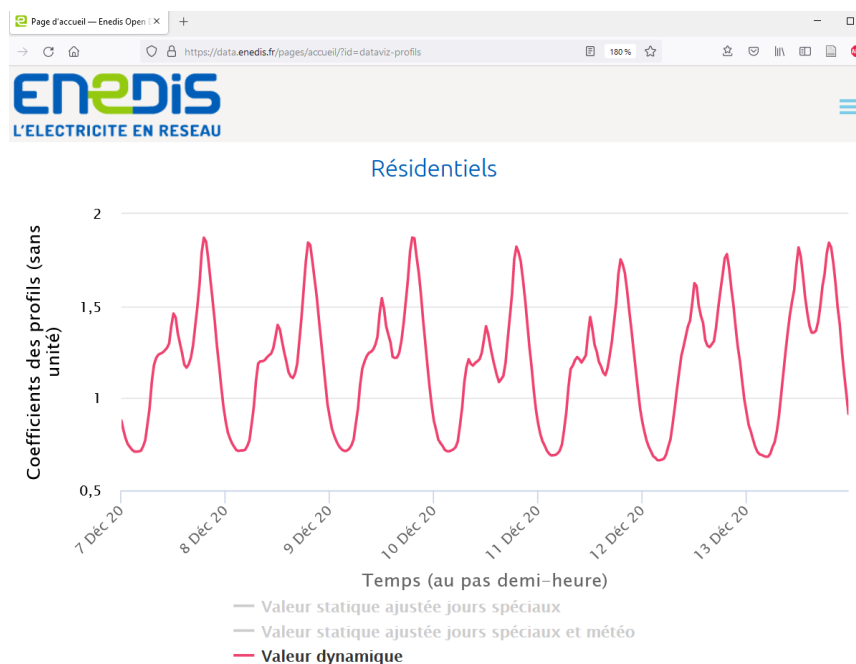


Figure II.16 Load curve of residential profile coefficient³³ by ENEDIS

³³ Enedis Data Visualisation : <https://data.enedis.fr/pages/accueil/?id=dataviz-profil>

For our study, we have selected profile coefficient curves of category *Residential* with sub-category *RES1_BASE*. This sub-category consists of the residential consumers which have contracted power ≤ 6 kVA. This is a classical contracted power which is being subscribed by around 7 out of 10 French households (Pinon, 2018). These households do not have electrical heating in their home. In our experiment, around 8 out of 10 households does not have electrical heating in their home. Therefore, this sub-category is close to the experimental case. Further, there exist 3 types of profile coefficient curves in this sub-category. All these curves are sampled at 30 minutes.

- The *prepared coefficient* depicts a model about how an average customer of this category consumes electricity over time at normal temperature. The normal temperature is further described in section II.4.c.ii. The normal temperature is established by *RTE* for a period of 10 years based on historical measured temperature of previous 30 years. Using the normal temperature, the values are adjusted for special days. The data is available for this coefficient since 26 December 2015;
- The *adjusted coefficient* depicts a model about how an average customer of this category consumes electricity over time at national temperature (instead of normal temperature). The national temperature is further described in section II.4.c.ii. Using national temperature, the values are adjusted for special days and weather. The data is available for this coefficient since 26 December 2015;
- The *dynamic coefficient* is the result of a more precise modelling. This coefficient directly take into account weather and special days (i.e. holidays). The data is available for this coefficient since 1st July 2018 (Réseau de Transport d'Électricité France, 2020b).

Assuming these curves as the modelled load curve of residential sector, the *dynamic coefficient* curve is selected for the forecasting of French national load curve of residential sector.

II.4.c.ii Temperature

This dataset corresponds to the national temperatures of France. Both types of temperature given below is used to calculate a coefficient of proportionality, which is called *temperature gradient*. It refers to the variation in consumption with respect to variation in 1 °C of temperature. However, it needs to consider the inertia of consumption in relation to temperature variations. For this purpose, the temperature is smoothened to take in to account the thermal inertia of buildings. The purpose of this is to make a hypothetical relation between the ambient temperature and the internal temperature of buildings (without knowing the internal temperature). The national temperature is calculated by a

weighted aggregation of measured temperature for 32 French cities where weather stations are installed. The data is sampled at 30 minutes. There are two types of temperatures available in this dataset.

- The *normal temperature* corresponds to a weighted average of smoothed normalized temperature of the 32 weather stations. The normal temperatures per station are established for a period of 10 years based on temperatures observed over a period of 30 years;
- The *national temperature* corresponds to a weighted average of smoothed measured temperatures of 32 weather stations on day ‘D-2’ (Réseau de Transport d’Électricité France, 2020b).

To forecast the French national load curve of residential sector, the *measured national temperature* dataset is taken. The data of this dataset is available until day ‘D-2’ every day.

II.4.c.iii Weather Forecast API

Since we only have data until day ‘D-2’ from *ENEDIS API* of temperature and pseudo-radiation, we need to collect forecasted temperature of the 32 French weather stations. For this purpose, we use the *API* of *Climacell* to gather the forecasted temperature data. For each French weather station, the *API* offers 109 hours of forecast with a time step of 1 hour.

RTE France (2020) presents the method of smoothing the temperature of 32 French weather stations, which is illustrated in Figure II.17. The smoothing is carried out by successive application of the following formulae.

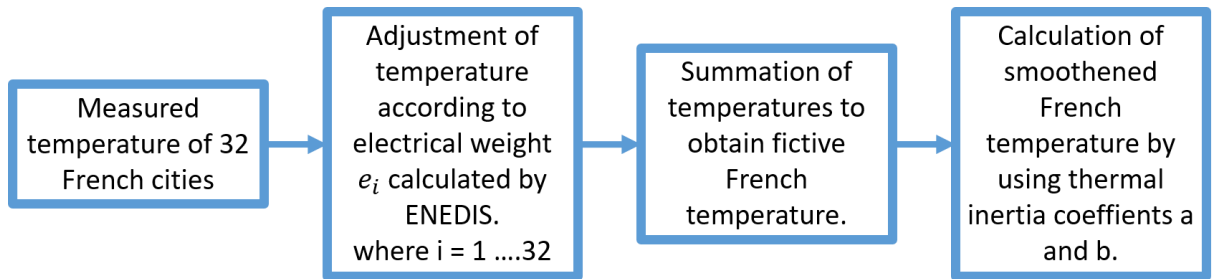


Figure II.17 The procedure of smoothing the measured temperature of 32 cities

$$T_{LT}(h, d, y) = (1 - a_{[h]}) \times T_F(h, d, y) + a_{[h]} \times T_{LT}(h - 1, d, y) \quad \text{Equation II.2}$$

$$T(h, d, y) = (1 - b_h) \times T_F(h, d, y) + b_h \times T_{LT}(h, d, y) \quad \text{Equation II.3}$$

Where h = hour, d = day, y = year, a , b = coefficients adjusting building thermal inertia, $T_F(h, j, n)$ = brute fictive temperature of France, $T_{LT}(h, j, n)$ = Long term temperature and $T(h, j, n)$ = smoothed French temperature.

II.4.d Regression Model Parameters

Figure II.18 gives a schematic diagram of the regression model.

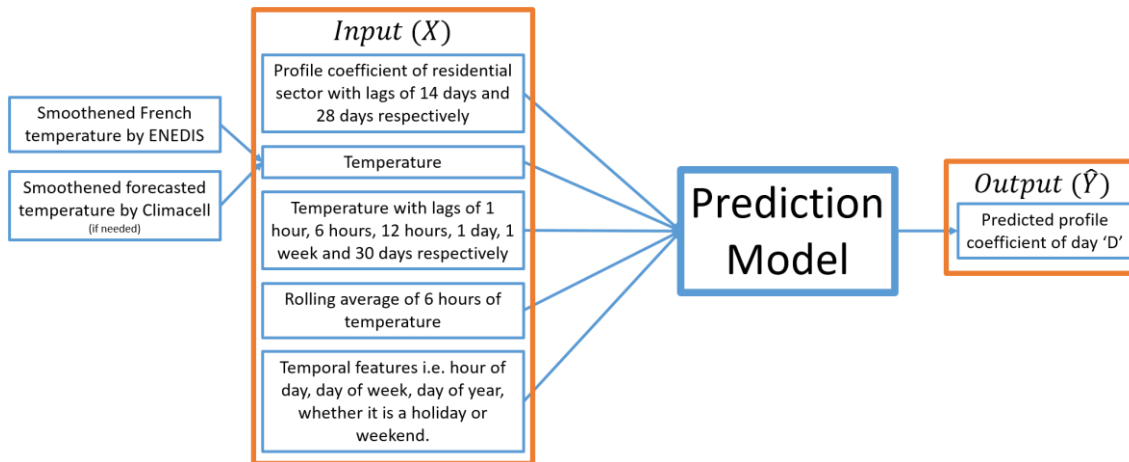


Figure II.18 The prediction model for predicting profile coefficient of day 'D'

II.4.d.i Dynamic Profile Coefficient

The historical data of dynamic profile coefficient obtained by *ENEDIS API French national load curve of residential sector* is used as input 'X' for training the model. In addition to this, 2 parameters of lagged dynamic profile coefficient are introduced in the dataset as part of input 'X'. One parameter of the model is dynamic profile coefficient lagged by 14 days, while other parameter is dynamic profile coefficient lagged by 28 days. The frequency of availability of new data for these curves is variable. The data is generally available until day $D-n$ where $n = 1 \dots 10$. Normally *RTE* updates the data on their open-source platform either weekly or every 10 days. This is the reason that lag of 14 days and 28 days are selected respectively rather than a lag of 1 day or 7 days. Figure II.19 shows a correlation between the actual profile coefficient and the lagged profile coefficient by 14 days and 28 days respectively. It shows the periodicity of the curve of profile coefficient.

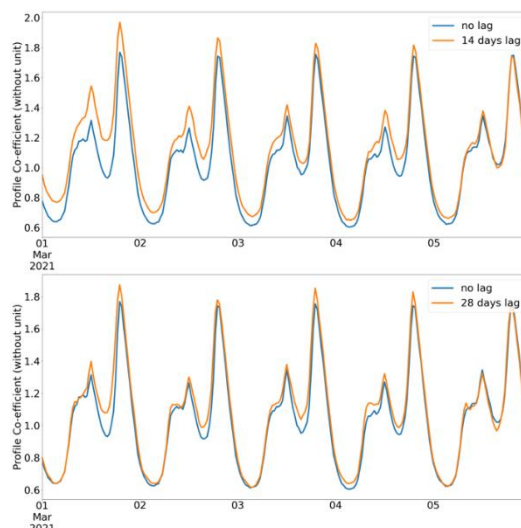


Figure II.19 The correlation of profile coefficient with lagged profile coefficients

II.4.d.ii Temperature

The historical data of smoothed national measured temperature of 32 French weather stations is obtained by *ENEDIS API*. In addition to this, 6 parameters of lagged temperature are added to the dataset. These temperature parameters are lagged by 1 hour, 6 hours, 12 hours, 1 day, 1 week and 30 days respectively. Another parameter of rolling average of 6 hours is also added in the dataset for training of model. Thus, a total of 8 parameters related to temperature are added in the dataset as part of input X for training of regression model. The smoothed national forecasted curve obtained by *ClimaCell API* is used for prediction by the created model.

II.4.d.iii Temporal Parameters

The following temporal parameters are added in the training dataset with one hot encoding. One hot encoding is a process by which categorical variables are converted into binary columns. e.g. a column having 7 day of week can be transformed into 7 binary columns representing each day of week.

- **Hour of day:** 24 one hot encoded columns are introduced to the training dataset for the parameter of hour of the day.
- **Day of week:** 7 one hot encoded columns are introduced to the training dataset for the parameter of day of week.
- **Day of year:** 365 one hot encoded columns are introduced to the training dataset for the parameter of day of year. It should be noted that data of leap day from year 2020 was removed from the model for the sake of convenience.
- **'Is_weekday':** A Boolean column is added as parameter *is_weekday*. The value remains 0 if weekday, else 1 for weekend.
- **'Is_holiday':** A Boolean column is added as parameter *is_holiday*. The value remains 1 if the day is a French holiday, else it remains 0.

II.4.e Model Evaluation and relative importance of features

Figure II.20 illustrates the prediction of load curve for weekdays as well as weekend. Since the model upgrades with the availability of new data, therefore for the purpose of illustration, the results are based on training of a historical data of 3 years and 2 months. It can be seen in the figure that the model is also capable of predicting the weekend, though the aim of prediction for the experiment is to have good prediction for weekdays.

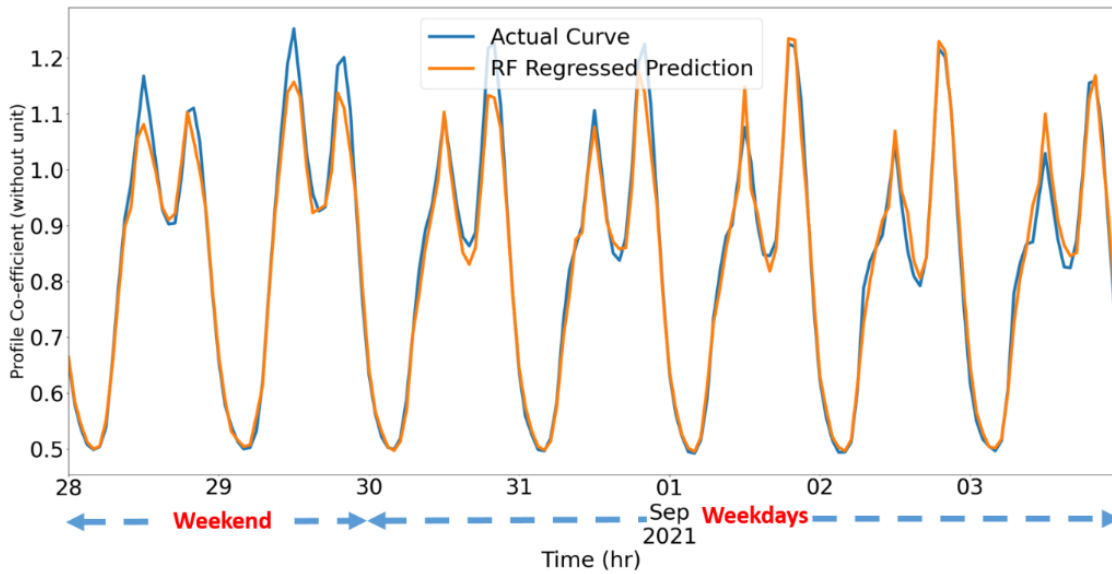


Figure II.20 Prediction of French residential load curve by random forest regression

It is worth mentioning that initially a model consisting of temporal features as input and dynamic profile coefficient as output is analyzed. Following this, the lagged features of dynamic profile coefficient are added in the model. Lastly, the temperature and lagged temperature features are also added in the model. An interesting factor to look at is the relative feature importance of the input parameters. The analysis of base model with temporal features and subsequent introduction of new features is illustrated below with respect to feature importance.

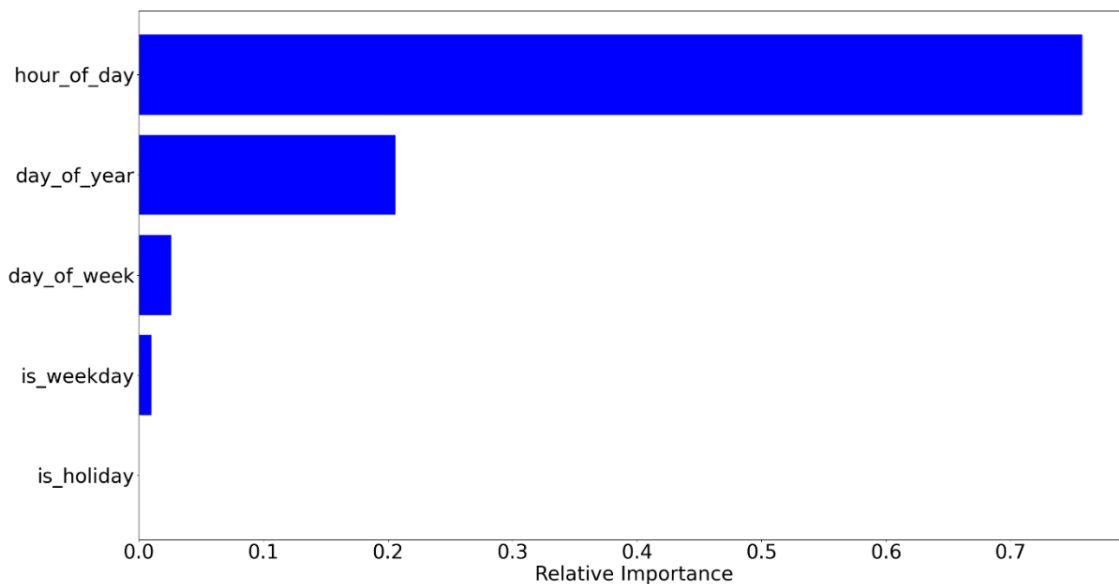


Figure II.21 Relative Importance of temporal features

It can be observed in Figure II.21 that *hour of day* is relatively most important feature among the temporal features. However, with the introduction of lagged features of dynamic coefficient, the temporal feature of *day of year* appears to be relatively most

important feature among temporal features (as demonstrated in Figure II.22). Among all the temporal features and lagged dynamic coefficient features, the dynamic profile coefficients lagged by 14 days remains the most important feature.

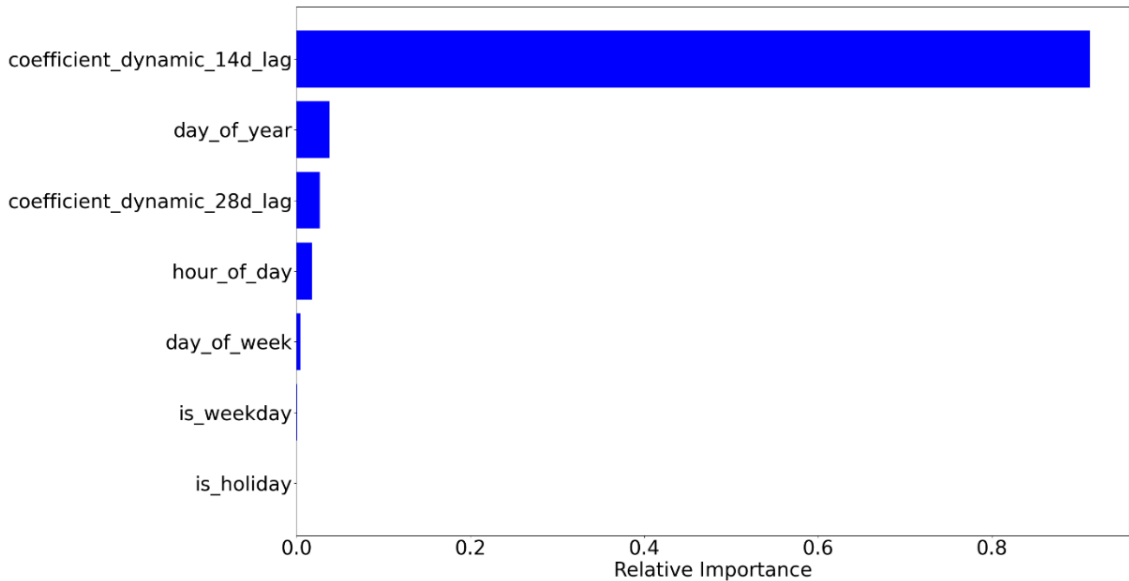


Figure II.22 Relative Importance of temporal features and lagged profile coefficient

Yet it can be observed in Figure II.23, a temperature feature lagged by 12 hours is relatively more important than the temporal features. Despite this, it is found that the temporal features are significant to improve the model performance. It can also be noticed that the feature illustrating whether the target day is a holiday does not seem to have any relative importance among all the features. Likewise, the features illustrating whether the target day is a weekday seems to be least important among all the features relatively.

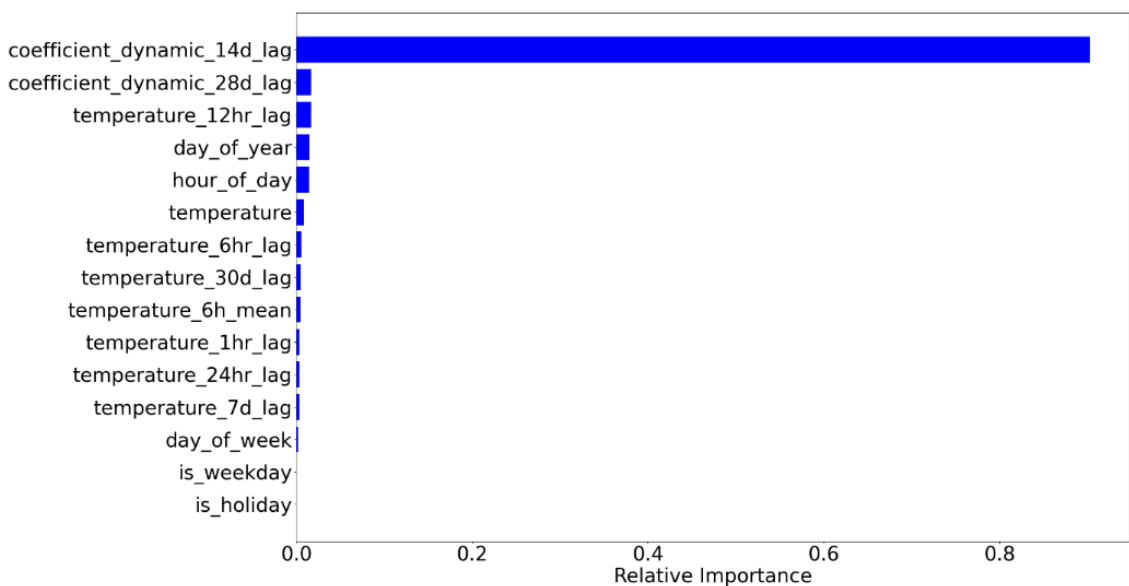


Figure II.23 Relative Importance of all features in model

II.4.f Performance Metric

It is observed in the previous section that the importance of input features can be calculated relative to each other. The relativity changes if new features are added in the model or certain features are removed. However, this does not give an absolute measure of the absolute validation of a model. Therefore, a performance metric is needed to measure the validation of a model. *Mean absolute percentage error (MAPE)* is taken as the performance metric to validate the results. The mathematical formula is given below.

$$MAPE (\%) = \frac{\left| \frac{y_{true} - y_{predicted}}{y_{true}} \right|}{\text{Number of } y_{true}} \times 100 \quad \text{Equation II.4}$$

Figure II.24 demonstrates the distribution of *mean absolute percentage error* recorded over a period of 3 years and 1 month. It should be noted that the outliers are not shown in the figure for the sake of comprehensible presentation. The prediction model having only temporal features gives a large distribution of *mean absolute percentage error*. The average *MAPE* observed in this case is 9.49% with a standard deviation of 6.77%. To improve the performance of model, the lagged features of dynamic coefficient profile and features regarding temperature are subsequently added. The addition of former decreased the average mean absolute percentage error to 6.43% whereas the addition of later decreased the error to 4.11%. The standard deviation also decreased. It concludes that besides temporal features, the prediction has a strong dependence on the historical data and temperature.

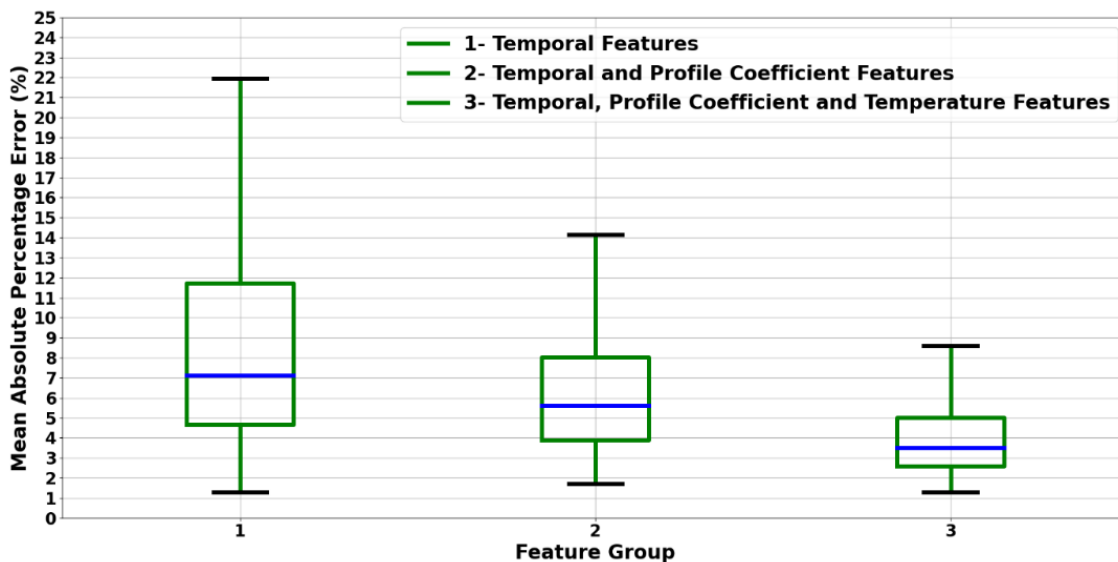


Figure II.24 The comparison of performance metric

Table II.4 Statistics of performance metric for each type of model

Model Input Aggregate Function	Temporal features	Temporal and profile coefficient features	Temporal, profile coefficient and temperature features
Count (Days)	808	808	808
Average (%)	9.49	6.43	4.11
Standard Deviation (%)	6.77	3.48	2.08
Minimum (%)	1.28	1.70	1.29
Maximum (%)	31.45	21.95	14.34

II.4.g Conclusion on the prediction of French residential load curve

A demand side management signal to the residential consumer based on the forecasted grid consumption retains the effect of all energy consumptions sectors. Therefore, it is more significant to send energy flexibility signals to the residential customers solely based on the forecasted condition of residential consumption. One way of doing this is by making a forecast on national load curve of residential sector.

Random forest regression is used to do prediction of modelled French national load curve of residential sector. The analysis of feature importance depicts that the historical value of French national load curve of residential sector is most important, followed by the temperature parameters and temporal parameters. The performance metric suggests that the addition of lagged features related to French national load curve of residential sector and national temperature improves the performance of the model. This prediction serves to calculate the $Coefficient_{GA}$ (green alert coefficient) to identify a day ahead potential alert for load shifting. With a forecasted clear sky and $Coefficient_{GA}$ greater than 1, the treated group is nudged to shift their energy consumption from evening to afternoon on next day.

II.5 Conclusion

Dynamic pricing is a tool to motivate consumer towards moderate energy consumption. The residential consumer faces constraint of limited knowledge and uncertain utility. If a consumer has enough information, the consumer does not know how to implement energy flexibility. The consumer makes decision according to system 2 of thinking by calculating the outcome and opportunity cost of energy flexibility. The residential consumer should be motivated to make automatic decision through system 1 of thinking while keeping in account the social, psychological, and ambient factors. Since the distributed generation serves as common pool resource, the residential consumer should be enlightened about the requirement of energy flexibility.

The general questions of the consumer regarding energy flexibility are identified i.e. why to act, when to act, how to act and how did it go. A green nudge-based energy flexibility mechanism is experimented with a pool of French households. A set of green nudges are used to answer the above questions. It includes information, commitment, and feedback. The purpose of this set of nudges is to intrinsically motivate the treated group.

Two types of interventions are experimented i.e. load curtailment and load shifting. The signal of load curtailment is based on the day ahead forecasted network congestion of the French electric grid. One of the criteria for load shifting signal is based on the forecasted modelled load curve of national residential sector. A prediction model is created using random forest regression. It is observed that by adding the lagging features of historical curve and temperature, the performance of prediction model is improved up to 57% of the baseline model only containing temporal features. It demonstrates the significance of lagged features of historical data for better prediction. Following the presentation of experimental architecture of nudge-based energy flexibility, the next chapter will discuss the efficiency of these nudge signals.

Chapter III

Measuring the impact of nudge signal

This chapter deals with the statistics of alerts i.e. interventions made for energy flexibility. It also presents a comparison of statistical and energy consumption characteristics of control group and treated group. The later section of this chapter discusses the quantification of the energy flexibility implemented by treated group as compared to the control group.

III.1 Introduction

In the previous chapter, the architecture of the experiment representing nudge based indirect energy flexibility in the households is presented. The purpose of this experiment is to study the impact of green alerts for load shifting and the impact of orange alerts for load curtailment. It is recalled here that the purpose of green alerts is to increase the energy consumption during the hours of elevated renewable energy production in the afternoon of summer season, whereas the purpose of orange alerts is to reduce the peak energy consumption during the evening in winter season. This chapter deals with analyzing the characteristics of subjects during alert days and non-alert days. Additionally, it also deals with measuring the impact of nudge signals.

III.2 General Statistics of the alerts

There were 175 households that participated in the study during the first session of orange alert and green alert. These households were categorized into control group and treated group based on strictly possible equivalence. Before the commencement of second session of experiment, the subjects were asked to re-sign a consent for continuing their participation in the experiment. Certain subjects left the experiment before the commencement of second session while some other left the experiment amidst the second session. This is owing to their personal choice, mal communication of smart meter or the fact that the households were relocated. The number of households in each group for each session is given in Table III.1.

Table III.1 Number of households in each group with respect to each year

Sr. No.	Type of group	Control Group	Treated Group	Total
	Year			
01	Year 1	79	96	175
02	Year 2	76	87	163

Two sessions of green alerts as well as two sessions of orange alerts were conducted during a period of 2 years. The experiment was conducted between November 2019 and September 2021. The duration of each session and the number of alerts triggered during each session is given in Table III.2. For the sake of comprehension, the following terminologies will be used hereafter:

- **Green (or Orange):** It represents the type of alerts for load shifting (or load curtailment) respectively.

- **Session X of green alerts (or orange alerts):** The respective session of green alerts or orange alerts. Here the value of X is either 1 or 2.
- **Non-alert days of session X:** The days for which no alert was in place during a session X. These days represent usual behavior of energy consumption of households. The energy consumption during the period of interest is significant to measure the impact of nudges on alert days.
- **Alert Days of session X:** The days for which an alert was in place during a session X. These are the days when the households are expected to implement energy flexibility during period of intervention.
- **Period of interest (or intervention)** The period of a day (in hours) for which an energy flexibility is expected to be implemented by the households of treated group on alert days. This period is between noon and 3 PM for green alerts and between 6 PM and 8 PM for orange alerts. To distinguish it for non-alert days and alert days, this period will be hereafter called as *period of interest* for non-alert days and *period of intervention* for alert days.

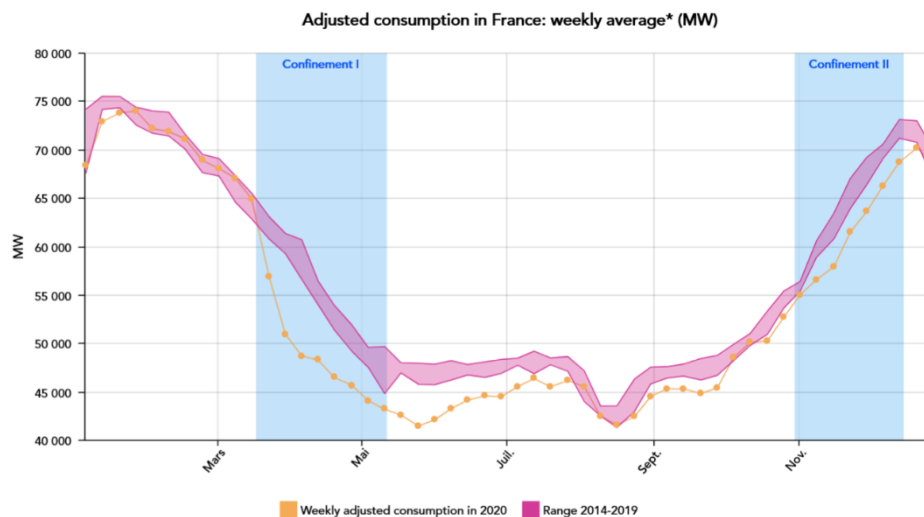
Table III.2 Number of alert days with respect to each session and alert type

Sr. No.	Alert Type	Purpose of Alert	Period of interest (or Intervention)	Duration of session	Number of non-alert days	Number of alert days
01	Green	Load shifting	Noon to 3 PM	First session: 1 st July 2020 – 31 st July 2020 and 24 th August 2020 – 30 th September 2020	46	5
				Second session: 1 st April 2021 – 31 st July 2021 and 24 th August 2021 – 30 th September 2021	108	2
02	Orange	Load curtailment	6 PM to 8 PM	First session: 1 st November 2019 – 31 st March 2020	99	9
				Second session: 1 st November 2020 – 31 st March 2021	96	12

III.2.a Green Alerts

The first session of green alerts was originally planned from April 2020 to September 2020. However, due to the first national lockdown in France caused by COVID-19 pandemic, the first session was re-planned from July 2020 to September 2020. It was done to not bother the subjects during the lockdown (and immediate post-lockdown) period. As it can be also observed in Figure III.1, the weekly average consumption in France during the lockdowns remained different comparing to the range of weekly average consumption of previous 5 years (Réseau de Transport d'Électricité France, 2020a). Therefore, triggering any alert during the lockdown period would not be effective to measure the impact of energy flexibility regarding usual energy consumption behavior. Besides, the behavior of energy consumption during the lockdown would be different than the normal behavior of energy consumption in the households, which would influence the experiment.

It is recalled here from section II.3.d.ii that two criterions were put in place for triggering a green alert. The principal criterion is related to the day-ahead forecasted nebulosity during the afternoon in Grenoble. The nebulosity should be zero. The second criterion is based on calculating day-ahead $Coefficient_{GA}$. This coefficient is a ratio of average evening energy consumption and average afternoon energy consumption. It should be greater than 1. It was expected that there will be a significant number of green alerts during the second session. However, contrary to the expectations, only 2 green alerts were triggered by the automatic system between April 2021 and September 2021. This was since the sky remained cloudy almost throughout the summer 2021 in Grenoble. Table III.3 demonstrates the basis of each triggered green alert for both sessions.



*adjusted consumption, business days only

1st confinement: from Tuesday, 17 March to Monday, 11 May 2020

2nd confinement: from Wednesday, 30 October to Tuesday, 15 December 2020

Figure III.1 The impact of lockdown on French energy consumption during 2020 (Réseau de Transport d'Électricité France, 2020a)

Table III.3 Basis of triggered green alerts

No. of Alert	Date	Average Evening Energy Consumption	Average Afternoon Energy Consumption	$Coefficient_{GA}$
Green Session 1				
01	27 July, 2020	1.15	0.98	1.17
02	25 August, 2020	1.19	1.02	1.17
03	14 September, 2020	1.3	0.98	1.33
04	17 September, 2020	1.26	0.92	1.37
05	22 September, 2020	1.28	0.84	1.52
Green Session 2				
06	14 June, 2021	1.18	0.98	1.20
07	06 September, 2021	1.19	0.92	1.29

III.2.b Orange Alerts

The first orange alert was triggered based on the forecasted negative temperature of Grenoble. This served as a test to find out whether the designed mechanism of the experiment works. Following this, the protocol of sending alerts based on the $PP1/PP2$ notifications was adapted. As the winter was not cold enough to bring network congestion on the national grid, it is deduced that there might be no more $PP1/PP2$ notifications. Therefore, EJP and $TEMPO$ notifications are added as secondary protocol. EDF split the territory of France in 4 zones to notify about the status of day ahead EJP . For triggering orange alerts, the notifications of EJP for south zone 1 is used.

Contrary to the 1st session of orange alert, the winter season during the second session of orange alert was colder and RTE notified 19 $PP1/PP2$ days. It is defined in the protocol that two consecutive alerts should not be sent for two consecutive days. Therefore, only 12 $PP1/PP2$ days were considered during 2nd session of orange alert. Since the number of alert days were enough, therefore no secondary protocol was adapted. Table III.4 demonstrates the basis of each triggered orange alert for both sessions.

Table III.4 Basis of triggered orange alerts

No. of Alert	Date	Basis of alert
Orange Session 1		
01	14 November, 2019	Manual Alert (Temperature < 0)
02	20 December, 2019	$PP1/PP2$
03	20 January, 2020	EJP (South Zone) ³⁴
04	06 February, 2020	EJP (South Zone)

³⁴ The region of Auvergne-Rhône-Alpes, where the experiment is performed is included in the south zone by EDF for EJP notifications : <https://www.fournisseurs-electricite.com/edf/tarifs/ejp/zone>

05	19 February, 2020	TEMPO
06	27 February, 2020	EJP (South Zone)
07	02 March, 2020	EJP (South Zone)
08	04 March, 2020	EJP (South Zone)
09	09 March, 2020	EJP (South Zone)
Orange Session 2		
10	30 November, 2020	PP1/PP2
11	02 December, 2020	PP1/PP2
12	04 December, 2020	PP1/PP2
13	07 December, 2020	PP1/PP2
14	09 December, 2020	PP1/PP2
15	18 December, 2020	PP1/PP2
16	04 January, 2021	PP1/PP2
17	06 January, 2021	PP1/PP2
18	08 January, 2021	PP1/PP2
19	11 January, 2021	PP1/PP2
20	26 January, 2021	PP1/PP2
21	11 February, 2021	PP1/PP2

III.3 Comparison of statistical variables of control group and treated group

Before discussing the impact of nudge signals for energy flexibility, it is important to assess the resemblance between the groups. Though the groups are formed with respect to strictest possible equivalence, it is pertinent to compare various statistical variables of both groups. The information of households regarding the statistical variables was gathered through the questionnaire filled by the subjects. If statistically significant difference is found between the groups for a number of statistical variables, the later results about the impact of nudge signals will be affected and may not be statistically validated.

The purpose of this section is to compare the similarity (or dissimilarity) of certain statistical variables between two groups. The variables that are analyzed in this section is related to the number of occupants (and appliances) in the households, the presence of occupants in the household and the characteristics of the building. For the sake of convenience, the variables are converted into their corresponding fraction in percentage. The variables and their percentage share are given in annex B.

A *Chi-square test*³⁵ (also known as ‘goodness of fit’ test) is performed on the statistical variables between the control group and the treated group. It is used to analyze

³⁵ The analysis was performed in RStudio in the programming language ‘R’. A group of students of econometrics assisted in this evaluation. The evaluation served as a course project for the students.

categorical data and to check the homogeneity of two groups. It measures that if the groups are independent, how well the observed distribution of data fits with the expected distribution. A quality of Chi-square test is that it does not require equality of variances among the study groups or homoscedasticity in the data (McHugh, 2013). Two hypotheses are formulated i.e., the null hypothesis (H₀) and the alternative hypothesis (H₁). These hypotheses are given below.

- H₀: The groups are similar/identical.
- H₁: The groups are significantly different.

The retention of null hypothesis (H₀) requires that the p-value in the result of Chi-square test should be greater than the value of significance level (denoted by α). The value of α is fixed at 0.05, which means that the test has a confidence level of 95% i.e., we are 95% confident that the observed results are real and are not an error caused by randomness. The results of Chi-square test for various variables are tabulated in Table III.5.

Table III.5 Chi-square test of independence of various variables

Sr. No.	Variable Name	p-value	Null hypothesis retained?
Household characteristics			
01	Owner/Tenant	1	Yes
02	House/Apartment	0.44	Yes
03	Surface of household	0.24	Yes
04	Year of construction	0.04	No
05	Number of persons in the household	0.83	Yes
06	Presence in the household at least 3 days per week (during weekdays)	0.52	Yes
07	Types of electric heating	1	Yes*
Presence of appliances			
08	Electric water heater	0.49	Yes
09	Air Conditioning	0.82	Yes
10	Heat Pump	0.99	Yes*
11	PV panels	0.31	Yes*
12	Double glazing of windows	0.16	Yes*
Number of appliances in household			
13	Electric Oven	0.98	Yes*
14	Dishwasher	0.57	Yes
15	Washing Machine	0.30	Yes*
16	Cloth Dryer	0.26	Yes
17	Electric Vehicle	0.56	Yes*
Energy Consumption during weekdays			
18	Pressing Iron	0.49	Yes
19	Electric Oven	0.08	Yes

20	Dishwasher	0.64	Yes
21	Washing Machine	0.60	Yes
22	Cloth Dryer	1	Yes
23	Electric hotplates	0.50	Yes
24	Radiator	1	Yes
Energy Consumption during the weekend			
25	Pressing Iron	0.51	Yes
26	Electric Oven	0.51	Yes
27	Dishwasher	1	Yes
28	Washing Machine	1	Yes
29	Cloth Dryer	1	Yes
30	Electric hotplates	0.63	Yes
31	Radiator	1	Yes

The p-value of Chi-square test for each of these variables is greater than 0.05. The case is contrary for the variable “year of construction”. Therefore, the null hypothesis is retained for all the variables in Table III.5 except for the variable “year of construction”. The information regarding this variable is collected for the purpose of using thermal loss of the dwellings in the study. However, since the experiment did not consider the modes of space heating, therefore this variable is not important in the study.

It should be noted in Table III.5 that for some variables, an asterisk is placed in the fourth column, next to ‘Yes’. These asterisks represent a warning during the execution of test, indicating that the approximation of Chi-square test of independence may not be aptly correct. For a chi-square test, it is assumed that the values in the ‘expected’ column should be equal to or more than 5 in at least 80% of the cells, and no cell in the ‘expected’ column should have value less than 1 (Bewick et al., 2003; McHugh, 2013). The warning might be raised due to non-fulfillment of assumption for respective variables. Though this makes estimation with a warning, yet it gives a result regarding the similarity in the groups.

The social comparison of both groups depicts that there is no statistically significant difference between the groups for most of the socio-technical statistical variables. In the results, there lies a limitation for certain statistical variables due to the non-fulfillment of basic assumption, yet the end results with respect to the data of both groups depicts that there is no statistically significant difference. Since the null hypothesis is retained for most of the variables, therefore it can be concluded that the two groups are statistically similar for multiple social characteristics.

III.4 Comparison of normal energy consumption between two groups

It has been observed in the previous section that the control group and treated group are found identical for several statistical variables. Yet, it is still to be studied that whether the energy consumption of treated groups is like the control group, and if so, then to what extent. The normal (or non-alert days) energy consumption in this case refers to the energy consumption on the weekdays when no energy flexibility is activated. If the normal energy consumption looks similar (or have negligible difference), it validates that the control group can be used as a standard against which the effect of treatment on treated group can be quantified.

For this purpose, meta-analysis is studied between the groups. A meta-analysis is a statistical analysis of a number of analysis results in order to integrate their findings (Crits-Christoph, 1992; Glass, 1976). Its purpose is to address a research question as definitely as possible through combining findings of multiple studies (TARG Bristol, 2017). The meta-analysis of energy consumption for non-alert days is discussed in this section with the methodology given below.

- The consumption data is split for each session as follows.
 - o To analyze the normal consumption, the measured load curve of non-alert days of each session is taken. The number of non-alert days are given in Table III.2.
- The meta-analysis is performed on the measured load curve of period of interest. The reason being that we are interested in knowing the impact of the nudge signals during period of intervention on alert day. It is recalled here that:
 - o For green alerts, this period is between noon and 3 PM.
 - o For orange alerts, this period is between 6 PM and 8 PM.
- The measured load curve data is originally sampled at 30 minutes. The data is analyzed in the form of average energy consumption per hour (during the period of interest). i.e.
 - o The average energy consumption per hour is mathematically defined as:

$$E_{average} = \frac{\sum_{t=i}^n P_i}{n/2} \quad \text{Equation III.1}$$

Where P_i is the instantaneous power at instant 'i'.

- o For green alerts, the number of measurements between noon and 3 PM are 6. The measurements are added and then divided by 3 to obtain the average energy consumption per hour.

- Similarly for orange alerts, the number of measurements between 6 PM and 8 PM are 4. The measurements are added and then divided by 2 to obtain the average energy consumption per hour.
- For all the analyses given below, it is found that there is no difference between the results if we either take the energy consumption of the whole period of interest or take average energy consumption per hour during the period of interest. Since it is common practice to present energy consumed per hour, therefore, the average energy consumed per hour is taken as unit. Besides, it is also helpful to compare the average energy consumption per hour in the evening to that of average energy consumption per hour in the afternoon.
- The analysis is performed using the following three statistical tools.
 - **Kernel Density Estimation:** To visually compare the statistically distribution of energy consumption of both groups.
 - **Mann Whitney U Test:** To observe whether the data of both groups is statistically similar or different.

III.4.a Visualization of statistical distribution on normal days

This analysis gives a global picture of the energy consumption during the period of interest on non alert days. The kernel density estimation³⁶ (shortened as KDE) is used to visualize the statistical distribution. For non-parametric data, KDE is an estimate of probability density function of a random variable (i.e. energy consumption in this case). Based on a finite number of data, KDE attempts to infer characteristics of a population of data. In our case, the comparison is done by the superposition of KDE distribution of treated group on the KDE distribution of control group.

Figure III.2 illustrates it for the non-alert days of green session 1 and green session 2 respectively. The peak density of energy consumption for both groups is found to be less than 0.5 kWh. It suggests that it is highly probable that the energy consumption of a household remains less than 0.5 kWh. The peak density of control group is higher than the peak density of treated group during 1st session of green alerts. It might be caused due to re-adjustment towards normal life during post-confinement, when the subjects had to adjust

³⁶ The estimation is made using the visualization function of Kernel density estimation in “pandas” package (of python language). This visualization function uses a class “scipy.stats.gaussian_kde” from “scipy” python package at backend. For the estimation, the default parameters are used i.e. Gaussian kernel is used alongwith “scott” method as bandwidth estimator of density estimation.

their work hours at their workplace and in their home. However, comparing to the 1st session, the peak density of both groups is within proximity in 2nd session.

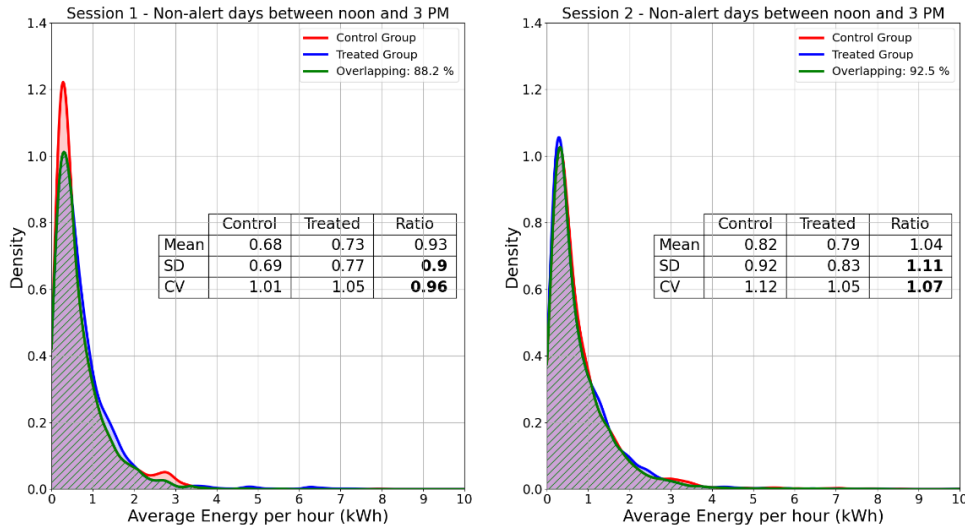


Figure III.2 The distribution of non-alert days of green sessions

The distribution of control group is overlapped by the distribution of treated group by 88.2% and 92.5% for green session 1 and green session 2 respectively. The difference of the mean values between both groups is minimal, whereas the coefficient of variation of both groups is also closer. Therefore, it can be concluded that the two groups are similar in terms of energy consumption during non-alert days of green sessions. The coefficient of variation is a ratio of standard deviation and mean value and it gives a measure of dispersion of data relative to the mean (Lovie, 2005). The third column entitled ‘ratio’ in the graphs below represents the values calculated as follows;

ROW, COLUMN	FORMULA
MEAN, RATIO	$\frac{\mu_C}{\mu_T}$
SD (STANDARD DEVIATION), RATIO	$\frac{\sigma_C}{\sigma_T}$
CV ($\frac{\sigma}{\mu}$), RATIO	$\frac{CV_C}{CV_T}$

Where μ is the mean value, σ is the standard deviation, CV is the coefficient of variation and the subscripts C and T represents control group and treated group respectively.

The distribution of non-alert days of orange sessions using KDE is illustrated in Figure III.3. The distribution of control group is overlapped by the distribution of treated group by more than 90%. Like the case of green sessions, the difference in mean values is minimal whereas the ratio of coefficient of variation is also closer to 1. Therefore, it can be

concluded that the two groups are similar in terms of energy consumption during non-alert days of orange sessions.

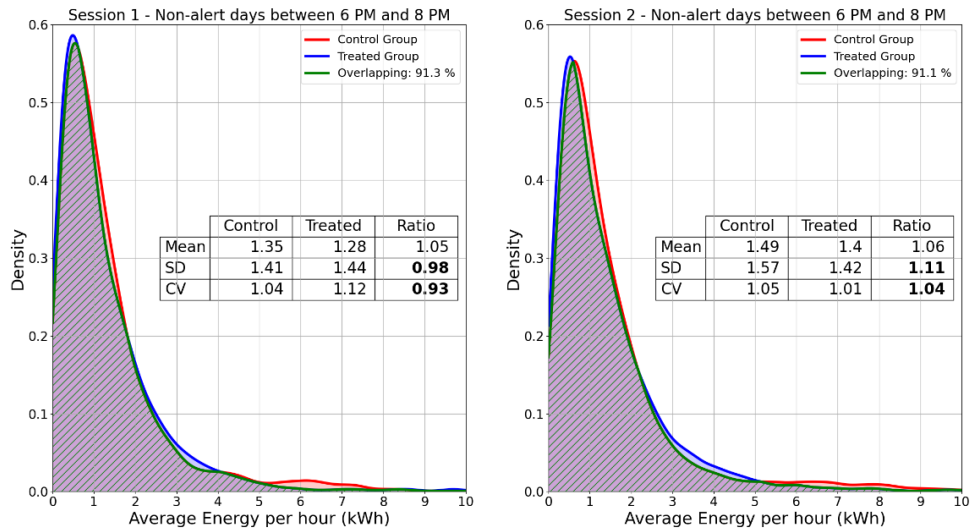


Figure III.3 The distribution of non-alert days of orange sessions

III.4.b Assessment of statistical difference in the normal energy consumption

This analysis gives a global demonstration of the energy consumption during the period of interest on non-alert days. As observed from the Figure III.2 and Figure III.3, the data is asymmetric and does not follow normal distribution. Therefore, a statistical test for non-parametric data, which is known as *Mann-Whitney U Test*³⁷ is performed. With respect to the alternative hypothesis (H1), the test is further classified into *one-sided* or *two-sided*. Here, our interest is to check whether the distribution underlying control group is similar (or significantly different) than the distribution underlying treated group. A *two-sided Mann-Whitney U Test* is done in this regard with the following hypotheses.

- Null hypothesis (Ho): The distribution (of energy consumption data) of control group is **similar to** the distribution (of energy consumption data) of treated group.
- Alternative hypothesis (H1): The distribution (of energy consumption data) of control group is **significantly different** than the distribution (of energy consumption data) of treated group.

The rejection of null hypothesis (Ho) represents the presence of statistically significant difference between the distributions of both groups. Since no intervention is made on non-alert days, therefore our interest is to find out whether the null hypothesis is retained for

³⁷ This test is also known as Wilcoxon rank sum test. The test was performed using the function "scipy.stats.mannwhitneyu()" of Python's SciPy package.

the energy consumption data of non-alert days. Statistically, it can be observed if the p-value of *two-sided Mann-Whitney U Test* is greater than $\alpha = 0.05$. The results of *two-sided Mann-Whitney U Test* is globally given below for green and orange sessions in Table III.6. Since the p-value for both types of sessions are greater than α , therefore it can be concluded that there does not exist any statistically significant difference in the energy consumption of both groups for non-alert days. In other words, the energy consumption of both groups is found similar for non-alert days.

Table III.6 The Mann-Whitney U Test for non-alert days of both types of sessions

Sr. No.	Sessions	Control Group			Treated Group			Mann-Whitney U Test	
		Obs	μ_c	σ_c	Obs	μ_T	σ_T	Stat	P-value
1	Green	9684	0.77	0.85	11794	0.77	0.81	56800317	0.50
2	Orange	16474	1.42	1.50	19285	1.34	1.43	164398493	1.19

where Obs = number of observations, μ = Mean value of observations, σ = standard deviation of observations.

III.5 Analysis of the impact of nudge signals

The visual analysis and statistical analysis in section III.4 concludes that the energy consumption of both groups during the period of interest on non-alert days is similar. It is now pertinent to perform the analysis of impact of nudge signals for the alert days. Firstly, a visual analysis of the energy consumption during the period of intervention on alert days is presented. It tells us whether the distribution curve of treated group is displaced with respect to the distribution curve of control group. Secondly, a statistical analysis is presented to find out whether the distribution underlying treated group is statistically less than (or greater than) the distribution underlying treated group. And finally, a linear regression is presented to find out the impact of each nudge signal in terms of percentage variation.

III.5.a Visualization of statistical distribution on alert days

This analysis presents a global comparison of energy consumption during period of intervention on alert days with respect to the energy consumption during the same period on non-alert days. Figure III.4 and Figure III.5 represents the distribution of energy consumption for green session 1 and green session 2 respectively. The distribution curve of treated group displaced towards left as compared to distribution curve of control group. For both green sessions, the ratio of mean energy consumption between control group and treated group decreased during alert days as compared to non-alert days. It merely suggests that the treated group increased their energy consumption during the period of intervention. In the case of green session 1, the ratio of coefficient of variation has negligible

difference between non-alert days and alert days. It means that the data is near-identically dispersed around the mean value. Despite, the difference of ratios is considerable for green session 2, the data of measured load curve of 2nd green alert was not registered by *ENEDIS* owing to a technical problem in its server. Therefore, Figure III.5 only represents the 1st green alert of 2nd session, which is insufficient for current and forthcoming analysis of green session 2.

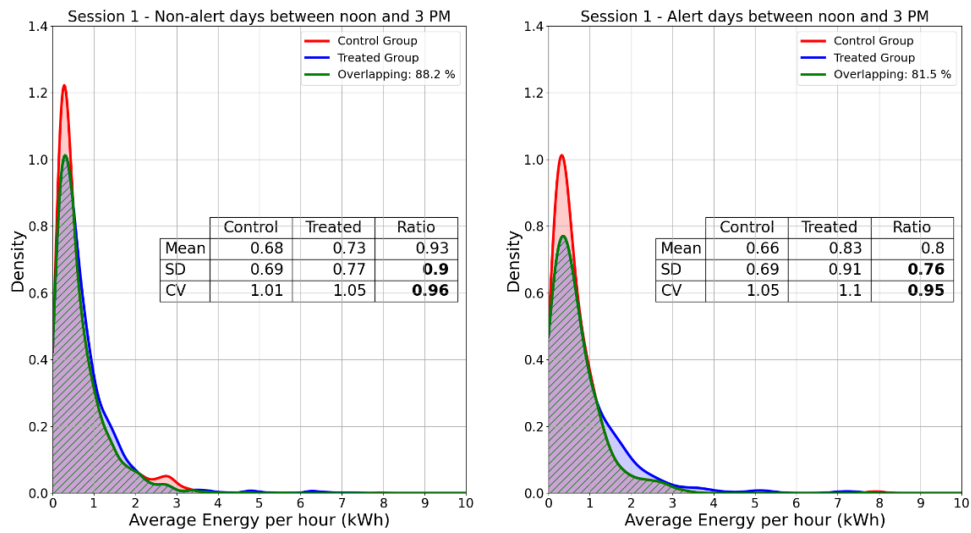


Figure III.4 The distribution of non-alert days and alert days of green session 1

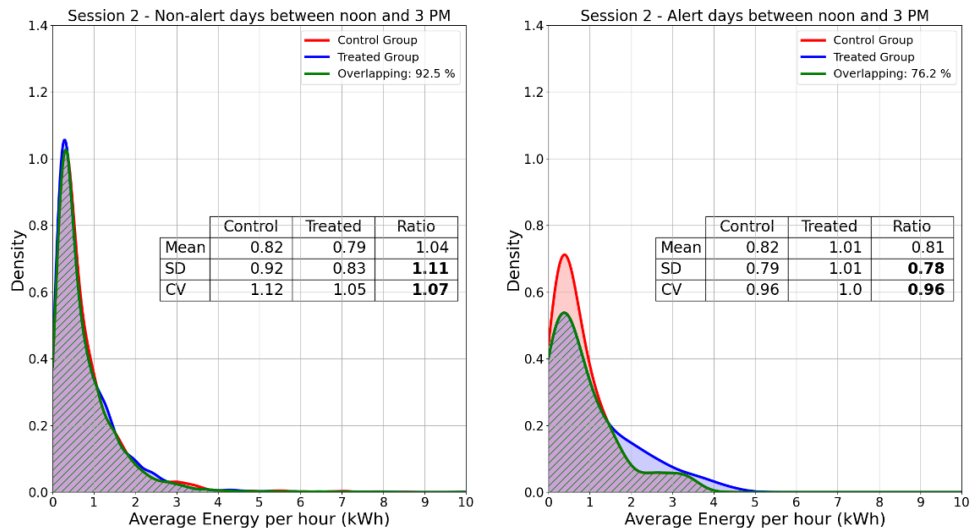


Figure III.5 The distribution of non-alert days and alert days of green session 2

Figure III.6 and Figure III.7 represents the distribution of energy consumption for orange session 1 and orange session 2 respectively. As compared to the mean consumption of control group, the mean consumption of treated group relatively decreased more for alert days as compared to non-alert days. On alert days, the distribution curve of treated group is displaced towards right (i.e. towards lower energy consumption) as compared to the

distribution curve of control group. Overall, the visual analysis between non-alert days and alert days gives a vague idea about the impact of nudge signal. Therefore, no clear conclusion can be drawn from this analysis and the results merit further analysis.

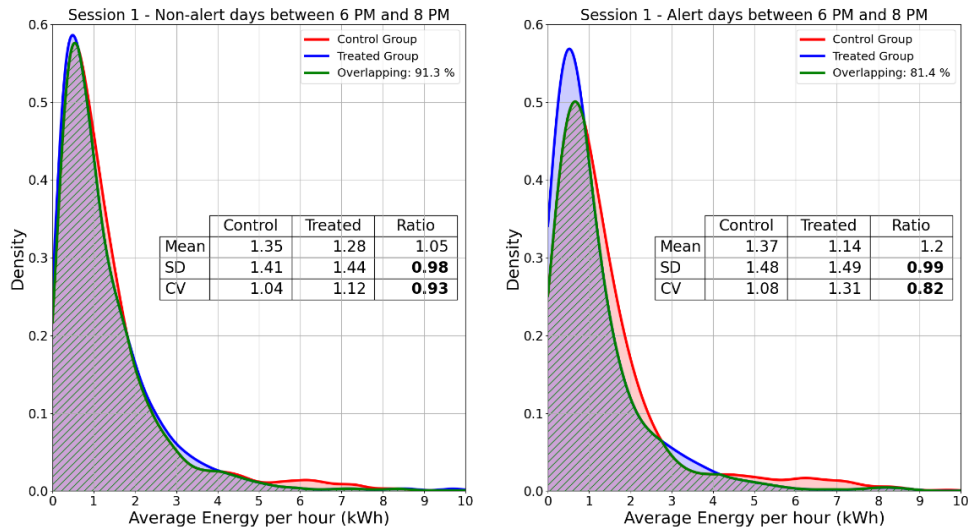


Figure III.6 The distribution of non-alert days and alert days of session 1 of type 'orange'

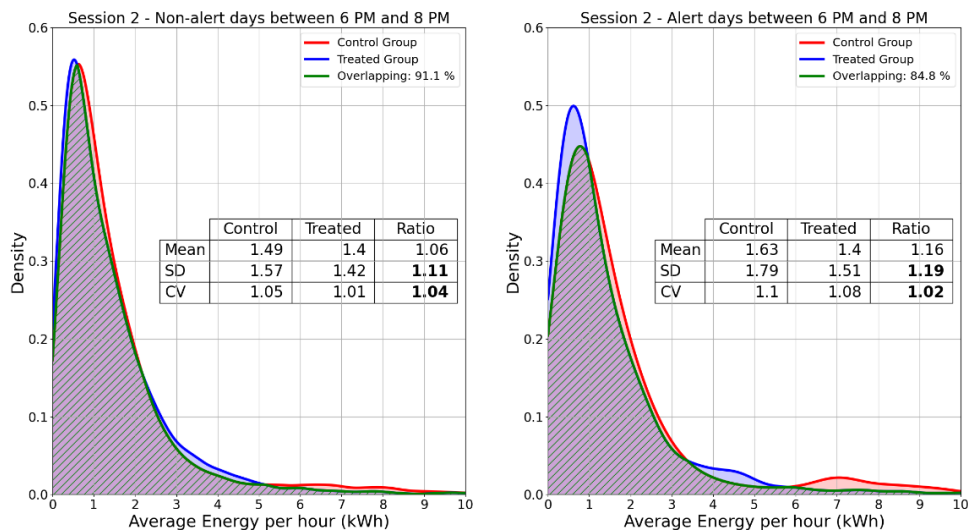


Figure III.7 The distribution of non-alert days and alert days of session 2 of type 'orange'

III.5.b Assessment of statistical difference in energy consumption on alert days

For this purpose, *one-sided Mann-Whitney U Test* is used with the following hypotheses.

- Null hypothesis (H₀): The distribution (of energy consumption data) of control group is statistically **greater** than the distribution (of energy consumption data) of treated group.

- Alternative hypothesis (H₁): The distribution (of energy consumption data) of control group is statistically **less** than the distribution (of energy consumption data) of treated group.

Since interventions are made for green alert days, therefore our interest is to find out whether the null hypothesis is rejected for the energy consumption data of alert days. Statistically, it can be observed if the p-value of *one-sided Mann-Whitney U Test* is less than $\alpha = 0.05$, which validates our alternative hypothesis.

Table III.7 The Mann-Whitney U Test for green alert days of both sessions

Sr. No	Alert Date	Control Group			Treated Group			Mann-Whitney U Test			$\mu_c - \mu_T$
		Obs	μ_c	σ_c	Obs	μ_T	σ_T	Stat	p-value	Null hypothesis retained	
Green Session 1											
1	27-07-2020	58	0.62	0.55	71	0.99	1.21	1736	0.06	Yes	-0.37
2	25-08-2020	45	0.82	0.63	34	0.81	0.64	772	0.53	Yes	0.01
3	14-09-2020	66	0.7	0.6	88	0.79	0.83	2871	0.45	Yes	-0.09
4	17-09-2020	65	0.69	1.03	88	0.77	0.68	2522	0.11	Yes	-0.08
5	22-09-2020	62	0.47	0.39	88	0.79	0.96	2253	0.04	No	-0.32
Green Session 2											
6	14-06-2021	72	0.82	0.78	85	1.01	1	2895	0.28	Yes	-0.19
7	06-09-2021	-	-	-	-	-	-	-	-	-	-
Aggregated Result											
Aggregated		368	0.69	0.71	454	0.86	0.93	76914	0.025	No	-0.17

The results for each alert of green sessions are tabulated above in Table III.7. Despite having negative difference of mean value of control group and treated group for most of green alerts, it is observed that the null hypothesis is only rejected for 5th green alert. However, the aggregated result of all 6 green alerts rejects null hypothesis i.e., the distribution (of energy consumption data) of control group is statistically less than the distribution (of energy consumption data) of treated group. It means that for all the green alerts globally, the treated group implemented load shifting for energy flexibility between noon and 3 PM. Yet, it seems a weak conclusion; considering that the retention of null hypothesis for 5 out of 6 green alerts is not supportive for aggregated results.

Table III.8 The Mann-Whitney U Test for orange alert days of both sessions

Sr. No	Alert Date	Control Group			Treated Group			Mann-Whitney U Test			$\mu_c - \mu_T$
		Obs	μ_c	σ_c	Obs	μ_T	σ_T	Stat	p-value	Null hypothesis retained	
Orange Session 1											
1	14-11-2019	78	1.41	1.37	92	1.32	2.39	4212	0.97	Yes	0.09
2	20-12-2019	77	1.25	1.4	90	0.92	0.74	3830	0.88	Yes	0.33
3	20-01-2020	77	1.65	1.76	92	1.3	1.57	4174	0.98	Yes	0.35
4	06-02-2020	79	1.41	1.6	91	1.11	1.31	4174	0.97	Yes	0.3
5	19-02-2020	79	1.47	1.61	91	1.19	1.29	4071	0.93	Yes	0.28
6	27-02-2020	78	1.26	1.33	91	1.23	1.32	3760	0.75	Yes	0.03
7	02-03-2020	78	1.38	1.42	90	0.93	1.07	4450	1.00	Yes	0.45
8	04-03-2020	79	1.22	1.37	91	0.98	1.06	4026	0.91	Yes	0.24
9	09-03-2020	78	1.32	1.32	90	1.26	1.91	3992	0.94	Yes	0.06
Orange Session 2											
10	30-11-2020	75	1.55	1.67	88	1.21	0.9	3449	0.69	Yes	0.34
11	02-12-2020	75	1.69	1.71	89	1.4	1.06	3511	0.72	Yes	0.29
12	04-12-2020	75	1.68	1.83	89	1.25	1.15	3801	0.94	Yes	0.43
13	07-12-2020	75	1.52	1.46	89	1.4	1.43	3664	0.86	Yes	0.12
14	09-12-2020	75	1.74	1.93	89	1.48	1.32	3468	0.67	Yes	0.26
15	18-12-2020	75	1.33	1.37	88	1.19	1.05	3513	0.76	Yes	0.14
16	04-01-2021	73	1.63	1.71	88	1.57	1.79	3506	0.84	Yes	0.06
17	06-01-2021	74	1.74	1.89	88	1.61	2.11	3600	0.88	Yes	0.13
18	08-01-2021	74	1.61	1.93	88	1.4	1.97	3798	0.97	Yes	0.21
19	11-01-2021	74	1.86	2.11	88	1.75	1.92	3578	0.86	Yes	0.11
20	26-01-2021	73	1.74	1.92	87	1.37	1.23	3420	0.80	Yes	0.37
21	11-02-2021	74	1.53	1.71	87	1.23	1.44	3496	0.83	Yes	0.3
Aggregated Result											
Aggregated		1595	1.52	1.66	1876	1.29	1.51	166724	1.00	Yes	0.23

The results for each alert of orange sessions are tabulated above in Table III.8. It is expected that the treated group curtail their load between 6 PM and 8 PM. Though, this can be observed by the positive difference in the mean of control group and treated group. Yet, the one-sided Mann-Whitney U Test validates it by retaining the null hypothesis, i.e. the distribution (of energy consumption data) of control group is stochastically greater than the distribution (of energy consumption data) of treated group. The null hypothesis is retained for energy consumption of individual alerts as well as for aggregated energy consumption of all alerts. It suggests that globally and for individual alerts, the treated group implemented load curtailment for energy flexibility between 6 PM and 8 PM.

III.5.c Quantification of change in the energy consumption during the period of intervention

It is now significant to quantify the change in the energy consumption during period of intervention on alert days. For this purpose, ordinary least square (OLS) regression test³⁸ is performed. The least square model is created by estimating the parameters in a regression model by minimizing the sum of the squared residuals. It can be applied to single or multiple explanatory variables and also categorical explanatory variables (Hutcheson & Sofroniou, 1999). For analysis, it is used to quantify the variation in energy consumption for all the alerts globally as well as for individual alerts. The variables for the test and origin of their data are tabulated in Table III.9.

Table III.9 The variables of ordinary least square test

Sr. No.	Variable Name	Description of variable and origin of data	Type of Data
Dependent Variable (Y)			
1	Ln (Period Energy)	For each subject, the natural log of energy consumption during the period of intervention on alert day.	Float (Decimal)
Independent Variables (X)			
2	Ln (Day Energy)	For each subject, the natural log of energy consumption during whole alert day.	Float (Decimal)
3	Sent SMS	The alert SMS sent to the subject of treated group on day 'D-1'. 1 for sent SMS and 0 for no sent SMS.	Boolean
4	Temperature	The average temperature in Grenoble on alert day. The unit of temperature is degree Celsius. The temperature is recorded at GreEn-ER. GreEn-ER is a research supporting living lab with microgrid interaction in Grenoble (Delinchant et al., 2016).	Float (Decimal)

It is recalled here that the original set of nudges consists of 3 types of SMS. However, we are only taking the alert SMS for analysis with the assumption that all the other SMS are also well received by treated group. Certain subjects of treated group did not receive alert SMS (owing to mal functioning of 3rd party dispatch server), therefore these subjects are considered as part of control group for respective alert. The global as well as individual impact of nudge signals for each alert is discussed in this section. The following general interpretations are helpful for understanding the change in the energy consumption. These

³⁸ The test was performed using the function "statsmodels.regression.linear_model.OLS" of Python's StatsModel package. The default parameters of the function are used.

interpretations are given with respect to the data type of each independent variable. The value of coefficient *coef* in the OLS regression gives us change in the energy consumption.

- The positive sign of coefficient *coef* indicates that the dependent variable Y increases with respect to the corresponding independent variable X. Similarly, a negative sign of coefficient *coef* indicates that the dependent variable Y decreased with respect to the corresponding independent variable X.
- If the dependent variable Y is log-transformed and independent variable X is in its original datatype, then:
 - o The value of coefficient *coef* is exponentiated i.e. $b = \exp(\text{coef})$.
 - o Then it is transformed into percentage using $a = (b - 1) * 100$.
 - o The results are interpreted as; the dependent variable Y will be changed by 'a' % for one unit rise in independent variable X.
- If both dependent variable Y and independent variable X are log-transformed, then the dependent variable Y will be changed by *coef*% with respect to 1% change in the independent variable X.

Figure III.8 summarizes the global results, whereas the detailed results are presented in annex C in tabular and graphical form. For green alerts, the increase in the energy consumption of treated group by the virtue of nudge signal is observed by 11.17% within a confidence interval of (0.9%, 21.4%) overall. This change in energy consumption is attributed to the nudge signal by 11.49% for 1st green session, whereas it is 7.09% for the 2nd green session. However, in the case of second green session, the results only take in to account one green alert, therefore due to less data, the results might not be statistically valid despite having a R² value of 0.707. The impact of exogenous variables like daily energy consumption and temperature is also found statistically significant during the period of intervention.

<i>Dependent variable: np.log(green_period_energy)</i>			
	Both Sessions	Green Session 1	Green Session 2
	(1)	(2)	(3)
Intercept	-3.764*** (-4.476 , -3.051)	-3.713*** (-4.620 , -2.806)	-0.008*** (-0.009 , -0.006)
Impact of SMS	0.112** (0.009 , 0.214)	0.115* (-0.002 , 0.232)	0.071 (-0.102 , 0.244)
Ln (Day Energy)	1.123*** (1.050 , 1.196)	1.111*** (1.021 , 1.202)	1.130*** (1.039 , 1.221)
Temperature	0.023*** (0.010 , 0.036)	0.026*** (0.010 , 0.042)	-0.158*** (-0.198 , -0.118)
Observations	899	612	156
R ²	0.681	0.653	0.707
Adjusted R ²	0.680	0.651	0.704
Residual Std. Error	0.539 (df=895)	0.547 (df=608)	0.548 (df=153)
F Statistic	314.809*** (df=3; 895)	198.935*** (df=3; 608)	10057.451*** (df=2; 153)
Note:	*p<0.1; **p<0.05; ***p<0.01		
	Standard Errors are clustered by subject ID		

Figure III.8 The statistical results of OLS regression for green sessions

We also analyze the change in the energy consumption of treated group during the period of intervention on green alert days versus the same period in previous non-alert days of the same session. Table III.10 gives us the results for both green sessions, where the change in energy consumption can be observed in the column ‘Difference on alert day’. e.g. the energy consumption of treated group between noon and 3 PM on the 1st alert day increased by 9.7% as compared to the previous non-alert weekdays. And it is statistically significant within the range (0.01, 0.05] of the *significant code*³⁹. However, the results are inverse for alert 3 and alert 5. By looking at the values in Table III.10 and Figure III.9, it can be deduced that no evidence of habit formation is found for load shifting against the green alerts.

³⁹ Significant codes: a) *** [0, 0.001] , b) ** (0.001, 0.01] , c) * (0.01 , 0.05]

Table III.10 The assessment of impact of green alert on alert day as compared to non alert days

Sr. No.	Alert Date	Obs	R ²	Condition Number	Difference on alert day	Log (Energy of day)	Temperature (°C)
Session 1 – Green Alerts							
1	27-07-2020	1969	0.728	370.19	0.097*	1.067***	-0.012
2	25-08-2020	2185	0.741	341.881	0.004	1.069***	-0.009*
3	14-09-2020	3547	0.727	317.336	-0.017	1.062***	0.002
4	17-09-2020	3716	0.721	319.488	0.004	1.06***	0.0
5	22-09-2020	4063	0.717	320.414	-0.033	1.066***	0.002
Session 2 – Green Alerts							
6	14-06-2021	6189	0.693	190.825	0.059	0.992***	-0.002

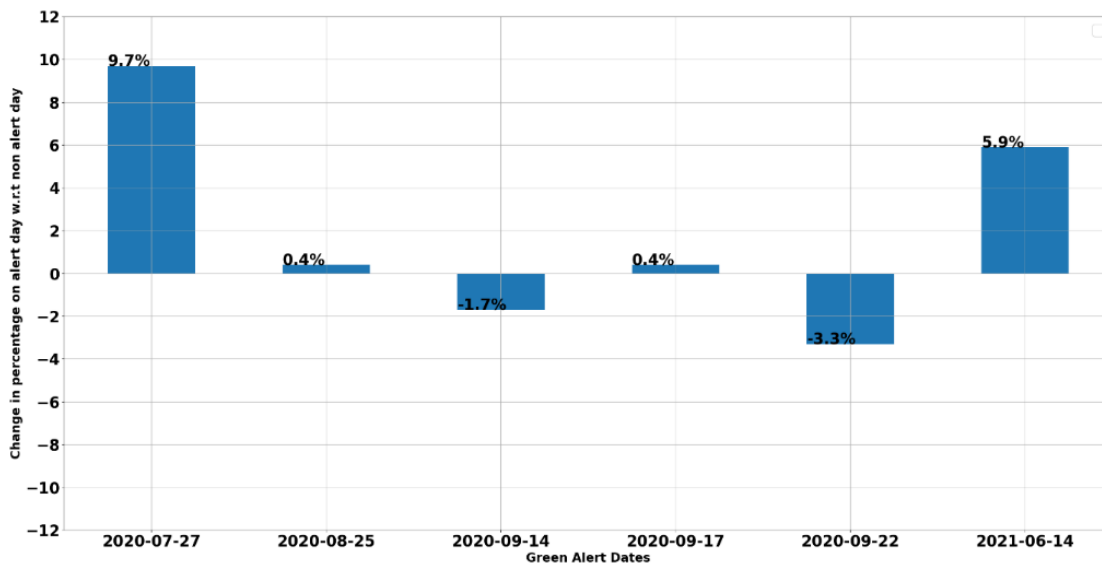


Figure III.9 The percentage change in the energy consumption on alert day as compared to non-alert day

Similar to the green alerts, the global impact of all the orange alerts comprising of both sessions (as well as individual session) is analyzed. Figure III.10 summarizes these results, whereas the detailed results are presented in annex C in tabular and graphical form. Overall, a decrease of 18.21% in the energy consumption of treated group is observed by the virtue of nudge signal within the confidence interval of (-25.8%, -10.6%). This change in energy consumption is attributed to the nudge signal by 22.78% for 1st orange session, whereas it is 14.68% for the 2nd orange session. The impact of exogenous variables like daily energy consumption is found statistically significant, yet the impact of temperature is found minimal.

A study of similar nature is done in the Monash residence halls in Australia (Jorgensen et al., 2021), in which the load curtailment is found in the range of 12-20% in response to behavioral science oriented treatment. The result of our study is closer to the study done in Australia, yet the protocol is different.

<i>Dependent variable: np.log(orange_period_energy)</i>			
	Both Sessions	Orange Session 1	Orange Session 2
	(1)	(2)	(3)
Intercept	-1.905*** (-2.355 , -1.454)	-2.004*** (-2.531 , -1.477)	-1.804*** (-2.301 , -1.307)
Impact of SMS	-0.182*** (-0.258 , -0.106)	-0.228*** (-0.318 , -0.137)	-0.147*** (-0.231 , -0.062)
Ln (Day Energy)	0.984*** (0.938 , 1.030)	0.992*** (0.938 , 1.045)	0.974*** (0.923 , 1.024)
Temperature	-0.007*** (-0.012 , -0.003)	-0.005 (-0.010 , 0.001)	-0.005 (-0.011 , 0.002)
Observations	3,440	1,508	1,932
R ²	0.726	0.704	0.741
Adjusted R ²	0.726	0.704	0.740
Residual Std. Error	0.521 (df=3436)	0.554 (df=1504)	0.492 (df=1928)
F Statistic	658.595*** (df=3; 3436)	488.978*** (df=3; 1504)	520.543*** (df=3; 1928)
Note:	*p<0.1; **p<0.05; ***p<0.01		
	Standard Errors are clustered by subject ID		

Figure III.10 The statistical results of OLS regression for orange sessions

Finally, we analyze the change in the energy consumption of treated group during the period of intervention on orange alert days versus the same period in previous non-alert days of the same session. Table III.11 gives us the results for both orange sessions, where the change in energy consumption can be observed in the column *difference on alert day*. e.g., the energy consumption of treated group on the 2nd orange alert day decreased by 23.8% as compared to the previous non-alert weekdays. And it is statistically significant within the range [0, 0.001] of the *significant code*⁴⁰. The *difference of alert day* is graphically presented in Figure III.11 and Figure III.12 for orange session 1 and orange session 2 respectively. Like the green alerts, no evidence of habit formation is found for both alerts. However, the difference of energy is considerable as compared to the green alerts.

⁴⁰ Significant codes: a) *** [0, 0.001] , b) ** (0.001, 0.01] , c) * (0.01 , 0.05]

Table III.11 The assessment of impact of orange alert on alert day as compared to non-alert days

Sr. No.	Alert Date	Obs	R ²	Condition Number	Difference on alert day	Log (Energy of day)	Temperature (°C)
Session 1 – Orange Alerts							
1	14-11-2019	1181	0,696	149.721	-0.09	1.058***	0.003
2	20-12-2019	4333	0,672	146.638	-0.238***	1.032***	0.003
3	20-01-2020	6934	0,693	129.438	-0.106	1.036***	0.005*
4	06-02-2020	8375	0,693	129.682	-0.134*	1.026***	0.004
5	19-02-2020	9462	0,689	131.124	-0.131*	1.018***	0.005**
6	27-02-2020	10082	0,687	131.659	-0.118	1.016***	0.004*
7	02-03-2020	10351	0,688	132.009	-0.24***	1.017***	0.004*
8	04-03-2020	10441	0,687	132.036	-0.172**	1.016***	0.004*
9	09-03-2020	10798	0,689	132.263	-0.091	1.016***	0.003
Session 2 – Orange Alerts							
10	30-11-2020	2536	0,702	159.553	-0.065	0.983***	0.007*
11	02-12-2020	2625	0,701	157.091	-0.045	0.982***	0.007*
12	04-12-2020	2713	0,705	157.56	-0.21***	0.986***	0.007*
13	07-12-2020	2889	0,708	150.887	-0.109	0.986***	0.009***
14	09-12-2020	2976	0,706	150.29	-0.052	0.978***	0.009***
15	18-12-2020	3672	0,704	145.66	-0.192***	0.981***	0.008***
16	04-01-2021	5060	0,726	134.413	0.018	1.008***	0.01***
17	06-01-2021	5148	0,727	133.575	-0.046	1.009***	0.01***
18	08-01-2021	5235	0,725	132.967	-0.107	1.006***	0.01***
19	11-01-2021	5409	0,727	131.205	0.041	1.007***	0.01***
20	26-01-2021	6609	0,718	128.838	0.102	0.991***	0.007***
21	11-02-2021	7889	0,718	131.477	-0.099	0.989***	0.007***

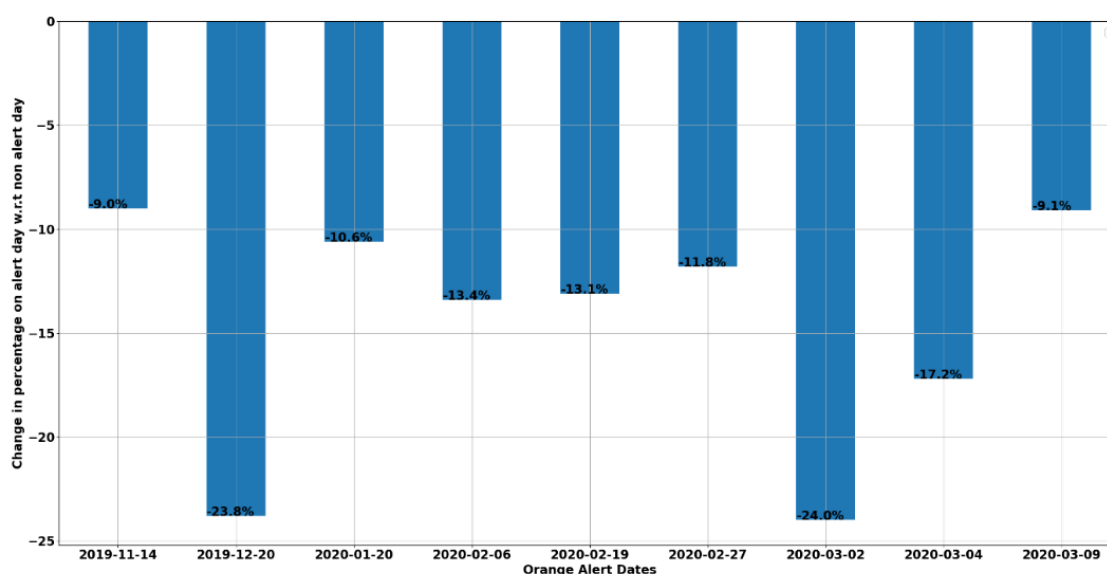


Figure III.11 The percentage change in the energy consumption on alert day as compared to non-alert day for 1st session of orange alerts

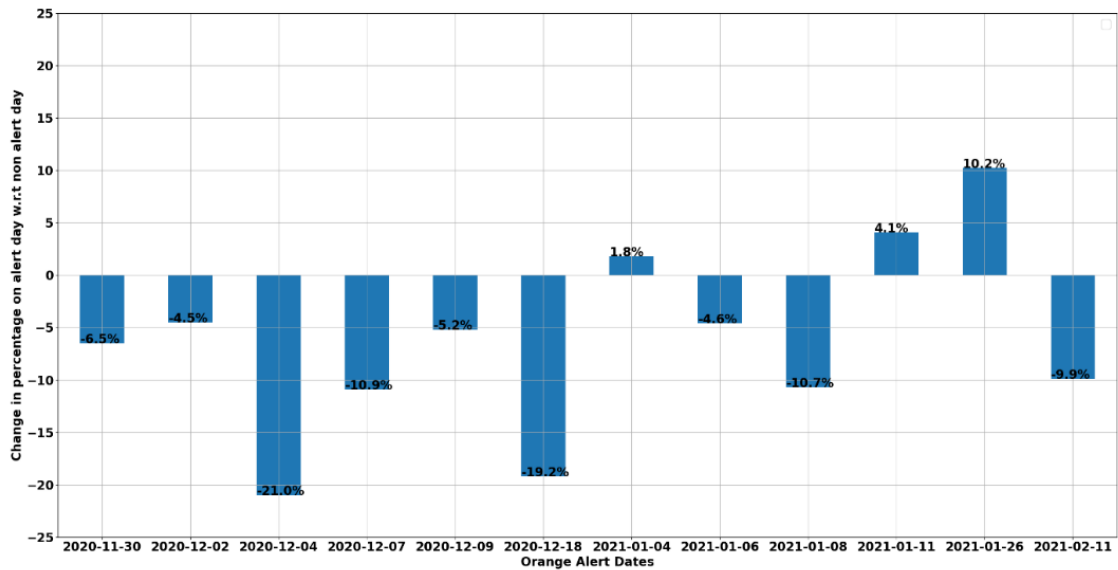


Figure III.12 The percentage change in the energy consumption on alert day as compared to non-alert day for 2nd session of orange alerts

III.6 Conclusion

Statistical analysis is a mean of measuring the impact of a treatment given to one group against a standard group. By using a statistical method, it is observed that there is no statistically significant difference between both groups with respect to most of the socio-technical statistical variables. The visual comparison of energy consumption of treated group with control group on non-alert days shows that the energy consumption is quite closer. It is also found that there does not exist any statistically significant difference between the energy consumption of control group with treated group. The similitude of energy consumption of both groups during non-alert days suggests that the impact of the nudge signal on alert days can be visualized and quantified by observing a difference in energy consumption.

Similar to the non-alert days, the comparison of energy consumption of control group and treated group is also done for alert days. The visual analysis shows that the distribution curve of treated group is displaced with respect to control group for alert days. This gives a mere conclusion that the nudge signals do have impact on the energy consumption during the period of intervention on alert days. In addition to this, the impact for each alert day is also measured to find out habit formation. The statistical test does not conform that all green alerts performed well, however it suggests that all the orange alerts does perform well.

Finally, it is observed that the treated group implemented load shifting of energy flexibility by 11.17% against the green alert SMS as compared to control group. Similarly, a load curtailment of energy flexibility by 18.21% is observed against the orange alert SMS.

Yet, by looking the energy consumption of treated group for each alert day with respect to previous non-alert weekdays, no significant (i.e., either horizontal or linear) pattern is observed regarding habit formation. In a nutshell, an impact of nudge signals is found and quantified for treated group as compared to control group. Yet no evidence of habit formation is found for implementation of energy flexibility by treated group.

Chapter IV

Reference Load Curve: An approach for indirect feedback

In this chapter, reference load curve is studied as a tool:

- to give an indirect feedback in graphical form
- to potentially replace control group and
- to categorize the individual with respect to their effort of energy flexibility according to a defined threshold.

IV.1 Introduction

The visualization of measured load curve in comparison to a baseline load curve serves as an easy and quick to understand feedback at a glance. “People can interpret various kinds of charts and tabular presentations if they are well designed” (Karjalainen, 2011). A residential consumer can himself analyze the impact of its effort conveniently before the next alert signal. In this case, the indirect feedback itself serves as a nudge for motivating the residential consumer. This baseline load curve can serve as a reference for measuring the impact on alert days. Hence, hereafter it will be called as reference load curve.

Here, it poses a scientific question as how to formulate a reference load curve that neither extremely underestimates or extremely over-estimates the effort of energy flexibility on alert day ‘D’. To explore the answer to this scientific question, multiple techniques are analyzed to formulate a reference load curve. These techniques include both naïve methods and some machine learning methods. By looking at the meta-analysis in chapter 3, it can be observed that a control group is inevitable to measure the impact of treatment. Therefore, we will try to answer another scientific question as whether the reference load curve can be used as an alternative to control group. Before defining the characteristics of a reference load curve, it is interesting to look at the conventional forms of indirect feedback. In this regard, a literature review of indirect feedback is given below.

IV.2 Literature review of Indirect Feedback

An indirect feedback is given to the customer by the energy utility after processing the raw data of energy consumption for a specific period of time; The purpose of indirect feedback is that the customer learns energy flexibility via reading and reflecting upon habitual energy usage (Darby, 2001). Despite the fact that the literature is heavily focused on direct feedback through either IHDs (Intelligent home devices) or web services, very few literature is available that talks about the mode of presentation of indirect feedback. Mostly, the indirect feedback is found in the form of frequent bills with either historic comparison (comparing the energy consumption with the past consumption of the same household), or normative comparison (comparing household energy consumption with that of a group of similar households) of energy consumed in kWh.

During the last decade of twentieth century, the indirect feedback was in the form of sending frequent billing with comparative or suggestive information. This includes more frequent energy bills with historical feedback (Wilhite & Ling, 1995), with normative feedback, with disaggregated feedback or with detailed annually or quarterly energy reports (Wilhite, 1999). Another form of indirect feedback is also tested by sending SMS and emails to the subjects in an experiment of Danish electricity supplier *SYD ENERGI* (Gleerup et al.,

2010). The SMS and emails are sent to the groups on daily basis, weekly basis or monthly basis. The purpose of the feedback is to inform the subject about deviation in energy consumption from their historical consumption level.

The famous study with home energy reports (HER) also used indirect feedback with historical comparison, normative comparison and injunctive norm (see section II.2.c). A study is also done to find out the consumer preference for feedback on household electricity consumption. It is observed in the study that many people find it difficult to understand the difference between power (W) and energy (kWh). Since the electricity bills are for energy consumption, therefore numerical feedback in the unit of W is not easily comprehensible for a non-professional. The study presents a variety of informational feedbacks and the consumers are found interested in the presentations of costs (over a period of time), appliance-specific consumption breakdown and historical comparison (Karjalainen, 2011). However, the diverse informational feedback causes the problem of information overload for the consumer. This makes it difficult for the consumer to understand the essence of feedback and to make decision accordingly for the future period of interventions.

In a nut shell, the state of the art suggests that the indirect feedback is either historic or normative (F. Wang et al., 2018). It is normally presented in numerical form e.g. kWh, cost, CO₂ emissions (Fischer, 2008), in the form of clustered baseline load curve of identical households (Abreu, Câmara Pereira, and Ferrão, 2012) or as average daily load curve (Ozawa, Furusato, and Yoshida, 2016). Following the idea of average daily load curve, some other naïve methods and machine learning models are statistically analyzed to formulate reference load curve of residential consumer.

IV.3 Formulation of Reference load curve

A reference load curve is a prediction of energy consumption for each instant on a non-alert day. Supposedly, a reference load curve serves to represent the usual behavior of energy consumption of a household on a non-alert day. Then the difference of area between the reference load curve and flexibility activated measured load curve gives a measure of energy flexibility on alert day. Besides, it also serves to graphically visualize the measure of energy flexibility, therefore serves as an easy to comprehend tool for indirect feedback of energy flexibility.

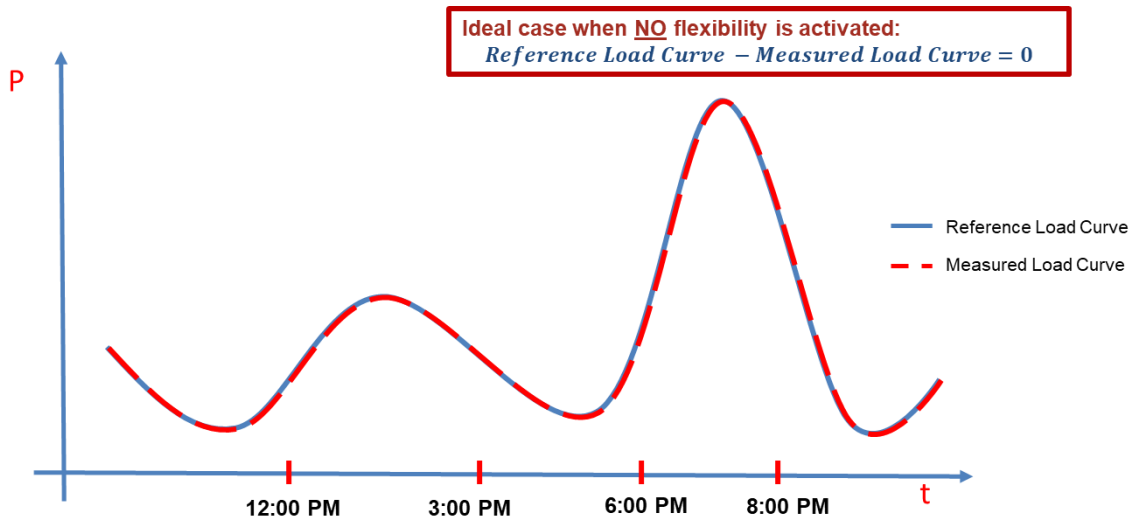


Figure IV.1 An ideal reference load curve

Figure IV.1 represents an ideal case of reference load curve. Ideally, it is expected from the reference load curve that the difference between reference load curve and measured load curve at each instant of measurement should be zero on non-alert day. In other words, the reference load curve exactly predict the energy consumption pattern of a household on a non-alert day. However, the formulation of reference load curve is practically based on predictive methods. Therefore, it may either under-estimate or over-estimate the energy consumption on a non-alert day. In other words, it suggests that practically, the difference between reference load curve and measured load curve at each instant of measurement is not zero (as illustrated in Figure IV.2). Since no technique fulfills the ideal condition of reference load curve, therefore it is required to choose a suitable method of formulating a reference load curve. A suitable method should formulate a reference load curve that should be as close as possible to measured load curve on non-alert days.

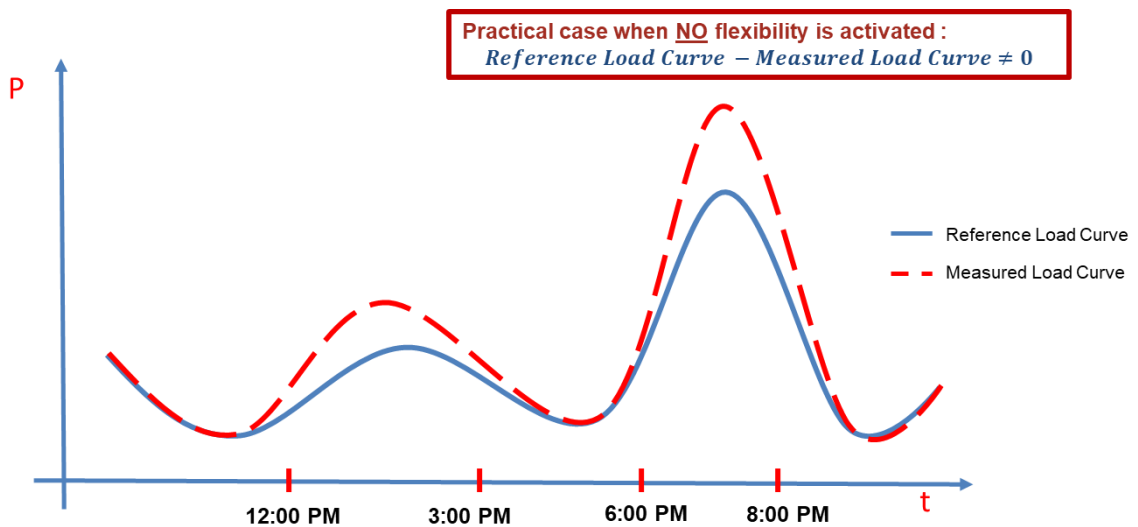


Figure IV.2 A practical reference load curve

IV.3.a Key Performance Indicators

Before discussing the techniques and their efficiency analysis for formulating reference load curve, it is necessary to define key performance indicators (KPIs). In this case, two indicators will be used i.e. energy difference indicator and coefficient of determination (also known as R² score). The energy difference indicator will later help us in understanding why alert type specific methods are defined in some technique. These KPIs are briefly discussed below.

IV.3.a.i Energy Difference Indicator

An energy difference indicator is the difference of reference load curve and measured load curve during the period of interest or intervention. Mathematically, it is given as follows:

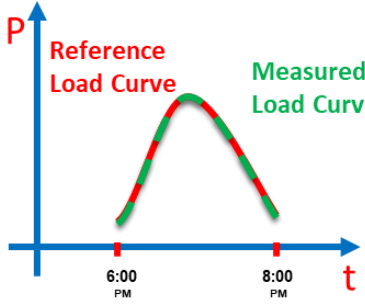
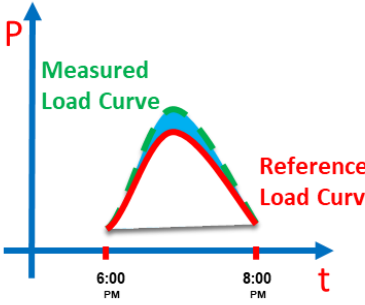
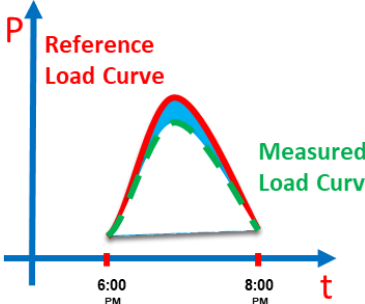
$$\text{Energy Difference Indicator}_{t_1}^{t_2} = \sum_{t_1}^{t_2} \text{Reference Load Curve} - \sum_{t_1}^{t_2} \text{Measured Load Curve} \quad \text{Equation IV.1}$$

where,

- t_1 : The starting time of period of interest/intervention. For type green, it is noon, whereas for type orange, it is 6 PM.
- t_2 : The ending time of period of interest/intervention. For type green, it is 3 PM, whereas for type orange, it is 8 PM.

In an ideal case, the energy difference indicator is zero for non-alert days. This is under the assumption that the reference load curve exactly predicts the (usual behavior of) energy consumption. For favorable feedback, we expect that the reference load curve should serve as nudge and its difference with measured load curve should be zero on non-alert days. However, it is not guaranteed that the prediction exactly estimates the measured load curve on non-alert days. The reference load curve either under-estimate or over-estimate the energy consumption on non-alert days practically. Table IV.1 illustrates the ideal and practical cases graphically. The energy difference is highlighted in blue color between the two curves. It should be noted that the underestimation (or overestimation) of reference load curve is illustrated with respect to the measured load curve.

Table IV.1 Ideal and practical case of energy difference indicator on non-alert days

Sr. No.	Case	Graphical Illustration	Sign of energy difference indicator
01	Ideal Case		Zero
02	Practical underestimation		Negative
03	Practical overestimation		Positive

Our interest in the case of green alert is to measure load shifting on alert days, therefore a practical reference load curve that minimally underestimates the measured load curve on non-alert days is suitable in this case. The value of energy difference indicator will be negative in this case. Alternatively, if a reference load curve that overestimates the measured load curve on non-alert days is selected to measure load shifting, it brings error in the calculation of energy flexibility for alert day.

Contrary to this, load curtailment is to be measured in response to orange alerts, therefore a practical reference load curve that minimally overestimates the measured load curve on non-alert days is suitable in this case. A reference load curve that underestimates measured load curve is not suitable in this case as it brings error in the calculation of energy flexibility for alert day. This problem is further discussed in section IV.4 regarding removal of error to measure energy flexibility on alert day.

In a nutshell, practically, a reference load curve that minimally underestimates the measured load curve on non-alert days is suitable to measure load shifting on green alert days. Whereas, a reference load curve that minimally overestimates the measured load curve on non-alert days is suitable to measure load curtailment on orange alert days. We will use this notion in section IV.3.b.ii and IV.3.b.iii to distinguishing the methods of formulation for green alert and orange alert.

IV.3.a.ii Coefficient of determination (R2 Score ⁴¹)

R-square is a statistical measure to find out about how close the data is to the fitted regression line. “The coefficient of determination can take values in the range $(-\infty, 1]$ according to the mutual relation between the ground truth and the prediction model” (Chicco et al., 2021). In complement to the energy difference indicator, the R^2 score suggests the suitability of the method. If the predicted value is equal to actual value, then R^2 score becomes 1. If a model disregards the input features and always predicts the expected value, then the R^2 score becomes 0.

IV.3.b Techniques studied to formulate reference load curve

Multiple techniques are studied in this regard to formulate a reference load curve. None of the technique requires non-supervised modelling, rather they are based on historical data of energy consumption. The techniques are explained in detail in their relevant subsection below, whereas Table IV.2 briefly presents the reason for choice of techniques.

Table IV.2 The choice of techniques to formulate reference load curve

Sr. No.	Technique	Remarks
Classical method		
01	Kernel Density Estimation	This technique serves to find the most probable value of energy consumption.
Naïve methods		
02	$Min(E_{D-1}:E_{D-5})$ $Max(E_{D-1}:E_{D-5})$	These methods are explored following kernel density estimation method. These methods are easily implementable and require less amount of near past energy consumption data of alert day ‘D’ as compared to classical method. Since these methods only takes the energy consumption data of previous 5 weekdays, therefore they also takes into account the variation in energy
03	$Mean[P_{D-1,i}, Min(P_{D-2,i}, P_{D-3,i}, P_{D-4,i}, P_{D-5,i})]$ $Mean[P_{D-1,i}, Max(P_{D-2,i}, P_{D-3,i}, P_{D-4,i}, P_{D-5,i})]$	

⁴¹ The R^2 score is calculated for each day in the data using the ‘r2_score’ method in the ‘metrics’ class of scikit-learn python package.

		consumption behavior (e.g. with respect to for example temperature).
Machine learning methods		
o5	Random Forest (Point to Point)	These methods are studies to find the impact of two architectural schemes of prediction.
o6	Random Forest (Sequence to Sequence)	

The following techniques are studied for the formulation of reference load curve. Among these techniques, two techniques have different statistical variables with respect to green alert and orange alert. Therefore, these techniques will be presented with respect to the type of alerts. It should be noted that the forthcoming analysis is done for the sole purpose of finding a suitable reference load curve that has near zero difference to the measured load curve on non-alert days.

IV.3.b.i The high-density value of Kernel Density Estimation of same timeslot

For an instance ‘i’, kernel density estimation is plotted based on the recorded consumption of previous ‘n’ weekdays. The purpose of doing this is to find out the most probable value of the same timeslot to put in reference load curve. In the distribution curve plotted through kernel density estimation, the consumption value with highest peak (i.e. density) is selected as the reference value of the instance ‘i’ in the reference load curve.

Figure IV.3 illustrated an example of orange period of interest. It shows an example where the kernel density estimation curve is plotted for the ‘n’ previous weekdays for timeslot 6:30 PM. The consumption value with highest density on kernel density estimation curve is taken as the reference value on reference load curve for timeslot 6:30 PM. The same procedure is used for the green period of interest.

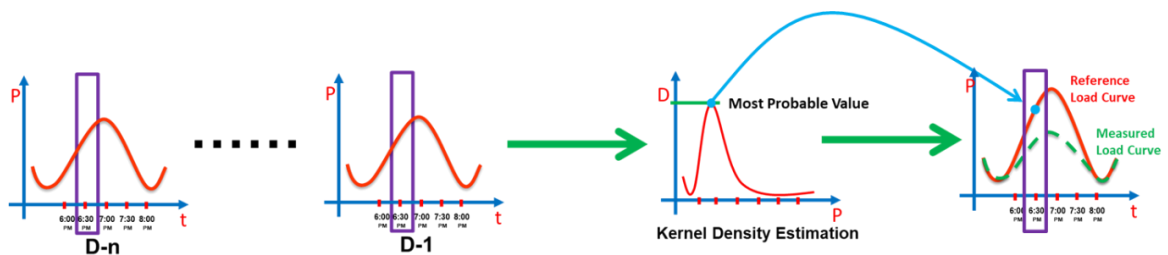


Figure IV.3 Formulation of reference load curve using Kernel Density Estimation

IV.3.b.ii Measured load curve of minimum (or maximum) energy consumption as reference load curve

This technique is type specific; therefore, we first discuss type green and later discuss type orange. As it is aforementioned that a practically suitable green reference load curve should have near zero yet negative error with measured load curve on non-alert day, therefore owing to this logic, the measured load curve having minimum energy

consumption between past 5 weekdays is taken as reference load curve. The choice of number of days can vary. In this case, the objective is to investigate the historical energy consumption during previous week, therefore 5 previous weekdays are chosen. Figure IV.4 represents measured load curve having minimum energy consumption between day ‘D-1’ and day ‘D-5’ as reference load curve on day ‘D’. It can be mathematically represented as:

$$Min(E_{D-1}:E_{D-5}) = Min \left(\sum_{t=i}^n P_{D-1,i}, \sum_{t=i}^n P_{D-2,i}, \sum_{t=i}^n P_{D-3,i}, \sum_{t=i}^n P_{D-4,i}, \sum_{t=i}^n P_{D-5,i} \right) \quad \text{Equation IV.2}$$

Where;

- Min() represents a function that gives the minimum value of a sample as an output.
- E_{D-1} is the energy consumption during period of interest on day ‘D-1’. Similarly, E_{D-5} is the energy consumption during period of interest on day ‘D-5’.
- $i \rightarrow n$ shows the range between 12:30 PM and 3:00 PM with half hourly samples.
- $P_{D-1,i}$ represents the power consumption at instantaneous value ‘i’ on day ‘D-1’. Similar for others days from ‘D-2’ to ‘D-5’.

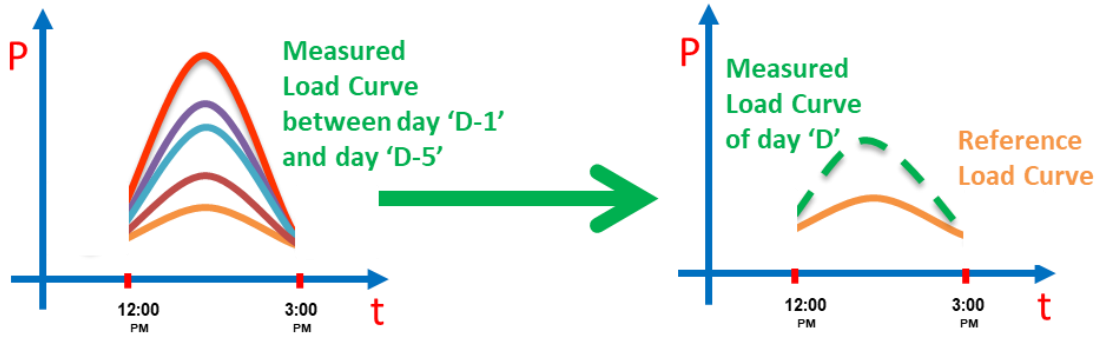


Figure IV.4 Selecting the minimum energy measured load curve as reference load curve

Using the same logic that the orange reference load curve should have minute positive error on non-alert days, the measured load curve having maximum energy consumption between past 5 weekdays is taken as reference load curve. Figure IV.5 represents measured load curve having maximum energy consumption between day ‘D-1’ and day ‘D-5’ as reference load curve on day ‘D’. It can be mathematically represented as:

$$Max(E_{D-1}:E_{D-5}) = Max \left(\sum_{t=i}^n P_{D-1,i}, \sum_{t=i}^n P_{D-2,i}, \sum_{t=i}^n P_{D-3,i}, \sum_{t=i}^n P_{D-4,i}, \sum_{t=i}^n P_{D-5,i} \right) \quad \text{Equation IV.3}$$

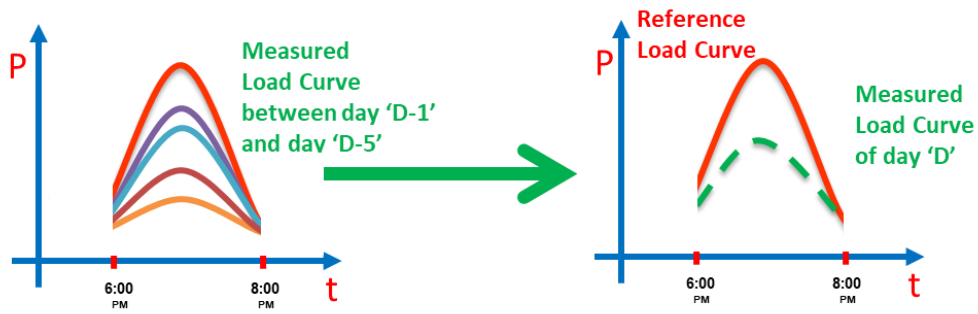


Figure IV.5 Selecting the maximum energy measured load curve as reference load curve

IV.3.b.iii Average of measured load curve of day 'D-1' and the day having minimum (or maximum) energy consumption between day 'D-2' and day 'D-5'

The technique presented in previous section is modified to formulate a new reference load curve. For a green reference load curve, an average is taken between the measured load curve of day 'D-1' and the measured load curve having minimum energy consumption between day 'D-2' and day 'D-5'. Figure IV.6 represents the green reference load curve.

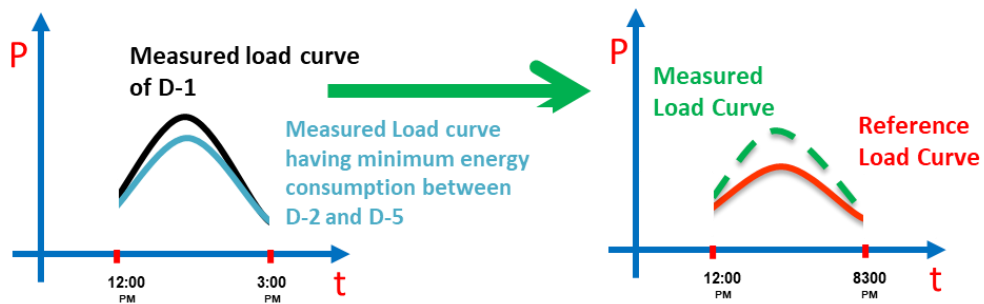


Figure IV.6 Green reference load curve formulation by taking average of measured load curve of day 'D-1' and day having minimum consumption between day 'D-2' and day 'D-5'

Similar to the green reference load curve, an orange reference load curve is also formulated. However, for an orange reference load curve, an average is taken between the measured load curve of day 'D-1' and the measured load curve having maximum energy consumption between day 'D-2' and day 'D-5'.

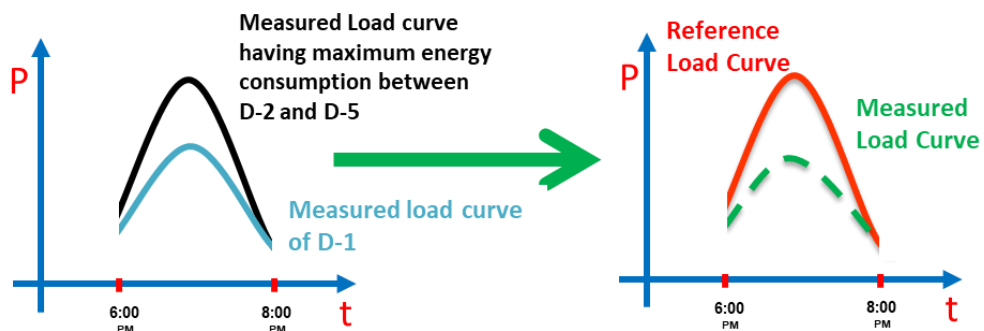


Figure IV.7 Orange reference load curve formulation by taking average of measured load curve of day 'D-1' and day having maximum consumption between day 'D-2' and day 'D-5'

IV.3.b.iv Random Forest Point to Point

In this technique, random forest is used to predict each instance (of period of interest) with respect to the rolling historical values of previous 5 weekdays of the same instance. The training of the model is done using all the historical data of same session. Figure IV.8 represents this technique for formulation of green reference load curve. e.g. to predict the reference value at 12:30 PM on day ‘D’, the measured energy consumption value at 12:30 PM during previous 5 weekdays is given as an input to the random forest prediction model. Similarly, the reference value at 1:00 PM, 1:30 PM, 2:00 PM, 2:30 PM and 3:00 PM are calculated to formulate the green reference load curve of day ‘D’.

Input (X)					Target (Y)
Day D-5 12:30 PM	Day D-4 12:30 PM	Day D-3 12:30 PM	Day D-2 12:30 PM	Day D-1 12:30 PM	Day D 12:30 PM
.
.
.
.
.
Day D-5 3:00 PM	Day D-4 3:00 PM	Day D-3 3:00 PM	Day D-2 3:00 PM	Day D-1 3:00 PM	Day D 3:00 PM

Figure IV.8 Random Forest Point to Point technique for creation of green reference load curve

Input (X)					Target (Y)
Day D-5 6:30 PM	Day D-4 6:30 PM	Day D-3 6:30 PM	Day D-2 6:30 PM	Day D-1 6:30 PM	Day D 6:30 PM
.
.
.
.
.
Day D-5 8:00 PM	Day D-4 8:00 PM	Day D-3 8:00 PM	Day D-2 8:00 PM	Day D-1 8:00 PM	Day D 8:00 PM

Figure IV.9 Random Forest Point to Point technique for creation of orange reference load curve

Similarly, the orange reference load curve is also studied using this technique. Figure IV.10 represents an illustrative example of model training and prediction of instance 6:30 PM at a day ‘D’. It should be noted that no additional parameter (e.g. temperature etc.) is added to the model, rather the historical consumption is only used for the prediction of reference load curve.

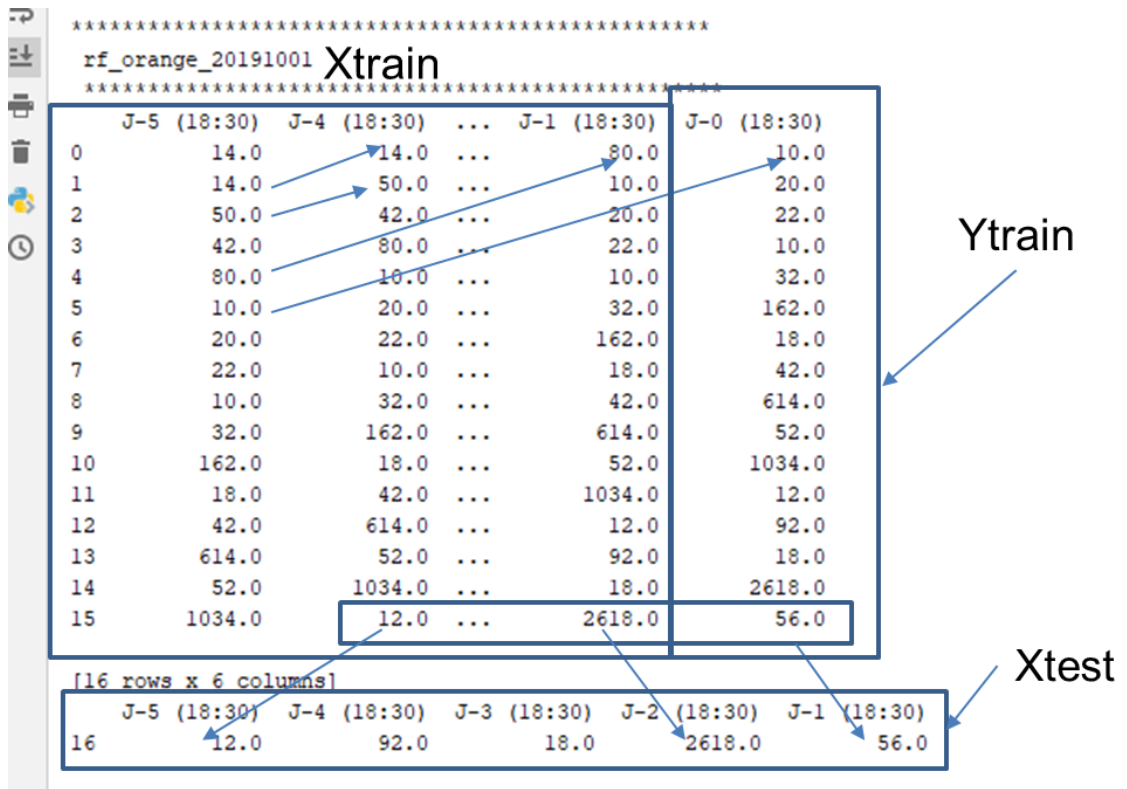


Figure IV.10 Training the model and predicting the instance 6:30 PM for a day 'D'

IV.3.b.v Random Forest Sequence to Sequence

This technique is like the 'Random Forest Point to Point' technique presented in previous section. The only difference is that the prediction of instance is replaced by prediction of a sequence. i.e. rather than predicting each instance and then formulating reference load curve, a sequence is predicted using the sequences of previous 5 weekdays. Figure IV.11 and Figure IV.12 are representing the techniques for green reference load curve and orange reference load curve respectively. Like random forest point to point, it should be noted that no additional parameter (e.g. temperature etc.) is added to the model, rather the historical consumption is only used for the prediction of reference load curve.

Input (X)									Target (Y)					
Day D-5 12:30 PM	Day D-5 1:00 PM	Day D-5 1:30 PM	Day D-5 2:00 PM	Day D-5 2:30 PM	Day D-5 3:00 PM	Day D-4 12:30 PM	Day D 12:30 PM	Day D 1:00 PM	Day D 1:30 PM	Day D 2:00 PM	Day D 2:30 PM	Day D 3:00 PM

Figure IV.11 Random Forest Sequence to Sequence for creation of green reference load curve

Input (X)									Target (Y)			
Day D-5 6:30 PM	Day D-5 7:00 PM	Day D-5 7:30 PM	Day D-5 8:00 PM	Day D-4 6:30 PM	Day D-4 7:00 PM	Day D-4 7:30 PM	Day D 6:30 PM	Day D 7:00 PM	Day D 7:30 PM	Day D 8:00 PM

Figure IV.12 Random Forest Sequence to Sequence for creation of orange reference load curve

IV.3.c Analysis for the suitability of a technique to formulate the reference load curve

In this section, we analyze the suitability of the techniques with respect to both key performance indicators. The reference load curves are calculated using the above techniques with the following characteristics.

- For analysis, the data of measured load curve for 175 households is used.
- Since the original data is sampled at 30 minutes, so we used the original sampling for analysis.
- The reference load curve is calculated for all the weekdays in both sessions of type green and type orange. The data of alert days is omitted to only analyze the non-alert days.
- Only the data corresponding to the duration of sessions is taken for analysis. The only exception is in the case of green session where the last date of data is 2nd July, 2021 instead of 30th September, 2021. It is because the fact that the analysis is done multiple times and it was lastly revisited in the month of July, 2021. No change in the results was observed during the multiple rounds of analysis.
- For each technique, the distribution of the energy difference indicator is presented in the form of box plot. The outliers are not presented with the boxplots. The reason for doing this is that the boxplots are strongly pinched by the outliers around OX axis on the graph. Thus, the boxplots become very small and could not be easily analyzed with naked eye.
- For each subject and each day, R^2 is calculated between measured load curve and reference load curve. The R^2 is only calculated for the period of interest. For each technique, the distribution of R^2 is presented in the form of box plot. Similar to the above case, the outliers are not presented with the boxplots.

IV.3.c.i Analysis of green reference load curves

The distribution of energy difference indicator for green reference load curves can be observed in the Figure IV.13. Recalling the characteristics of a suitable green load curve, it is expected from a green reference load curve that the difference should be negative yet arbitrarily closer to zero. In the term of a boxplot, these characteristics can be translated in such a way that:

- The boxplot is less dispersed around zero. It suggests that the reference load curve is closer to the measured load curve.
- Ideally, the boxplot should be beneath zero. Practically, the distribution of most of the data should be negative.

By observing the boxplots in Figure IV.13, it can be observed that the random forest techniques are not suitable as green reference load curve. The reason is that more than 50% data distribution of the boxplot is positive, suggesting that the predicted value of reference load curve is mostly higher than the measured load curve. Besides, they are widely dispersed as compared to the green reference load curve of other three methods. By comparing 2nd and 3rd boxplots, it is observed that the 3rd boxplot is less dispersed as compared to 2nd boxplot. Therefore according to the energy difference indicator, the 1st and 3rd method are suitable to formulate the reference load curve.

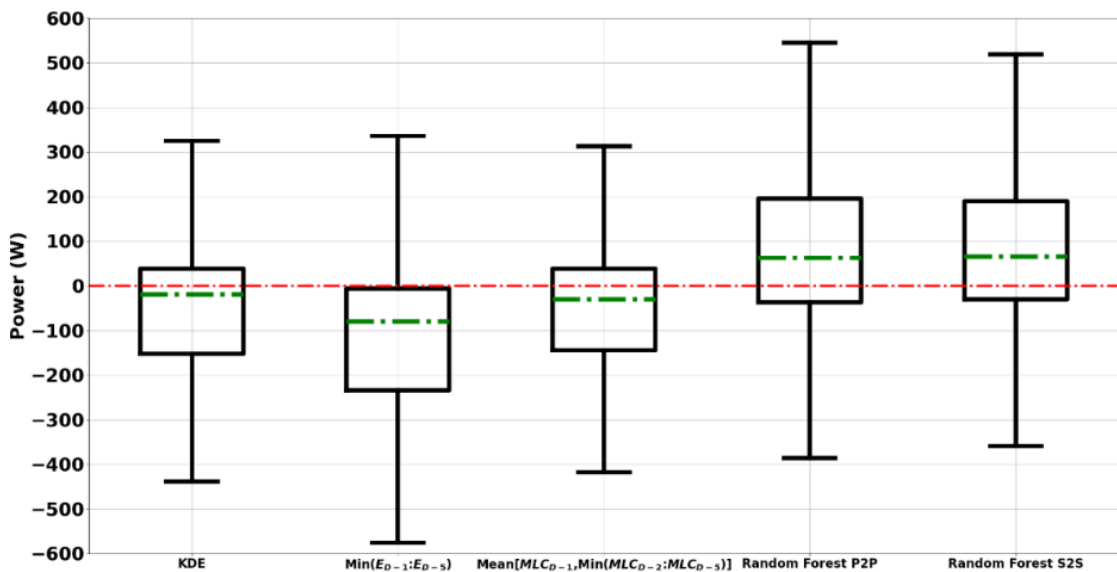


Figure IV.13 Distribution of energy difference indicator for green reference load curves

The R^2 score distribution boxplots for both random forest techniques (in Figure IV.14) also conforms the results of energy difference indicator in Figure IV.13. Though, the R^2 score of 2nd and 3rd method are close, yet the less dispersed data favors 3rd method over 2nd method. Among all the methods, the KDE methods performed better than others. Yet, it is easier to explain the underlying mechanism of 3rd method to our subjects as compared to KDE method. Therefore, in our experiment we chose the 3rd method albeit that the KDE method seems more promising for the formulation of a reference load curve to measure the load shifting of energy flexibility.

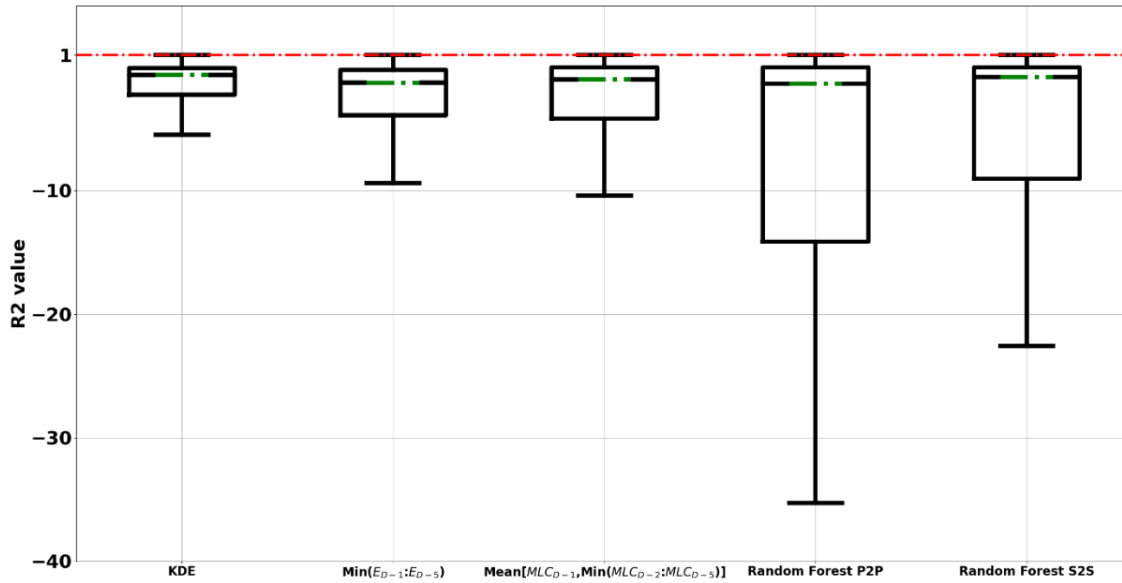


Figure IV.14 Distribution of R2 score for green reference load curves

IV.3.c.ii Analysis of orange reference load curves

The distribution of energy difference indicator for orange reference load curves can be observed in the Figure IV.15. Recalling the characteristics of a suitable orange load curve from section IV.3.a.i, it is expected from an orange reference load curve that the difference should be positive yet closer to zero. In the term of a boxplot, these characteristics can be translated in such a way that:

- The boxplot is less dispersed around zero. It suggests that the reference load curve is closer to the measured load curve.
- Ideally, the boxplot should be above zero. Practically, the distribution of most of the data should be positive.

It can be observed in Figure IV.15 that the kernel density estimation technique is not suitable for orange reference load curve. The reason is that around 75% of data distribution is negative, suggesting that the predicted value of reference load curve is mostly lower than the measured load curve. The 2nd boxplot is predominately positive yet highly dispersed as compared to others. Surprisingly, the random forest techniques give similar results albeit different values in their sample space and the two boxplots are also less dispersed. The random forest models are outperforming other methods with respect to energy difference indicator.

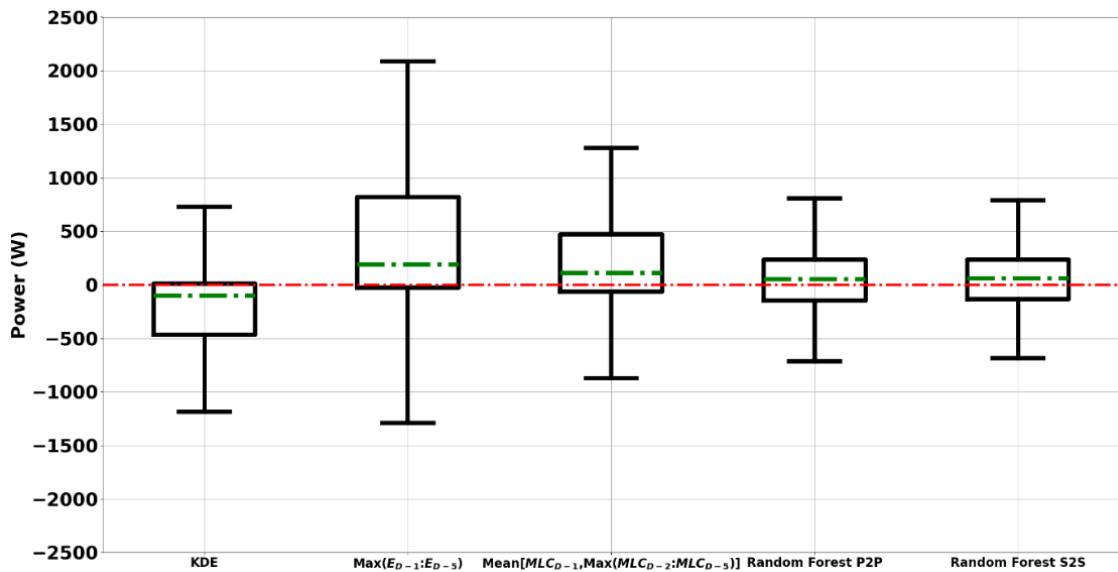


Figure IV.15 Distribution of energy difference indicator for orange reference load curves

The R^2 score boxplot of kernel density estimation in Figure IV.16 suggests it to be the most suitable reference load curve formulating method, yet the result of energy difference indicator is contradictory to this deduction. The R^2 score distribution of 2nd method is in the range of $[-180,1]$, which is lower comparing to the others. The 2nd lower R^2 score distribution is of 3rd method, which is in the range of $[-70,1]$. On the other hand, the random forest sequence to sequence method is found to be better than the random forest point to point method according to R^2 score distribution. It is evident that the random forest sequence to sequence method of formulating reference load curve outperforms other methods to measure the load curtailment of energy flexibility. Yet, since the machine learning methods are explored at a later stage during experiment, so the 3rd method was chosen to formulate the reference load curve for the reasons that:

- It has the best results among the first 3 methods.
- It is easy to explain to the subjects.
- It is symmetric to the method chosen for the case of green alerts.

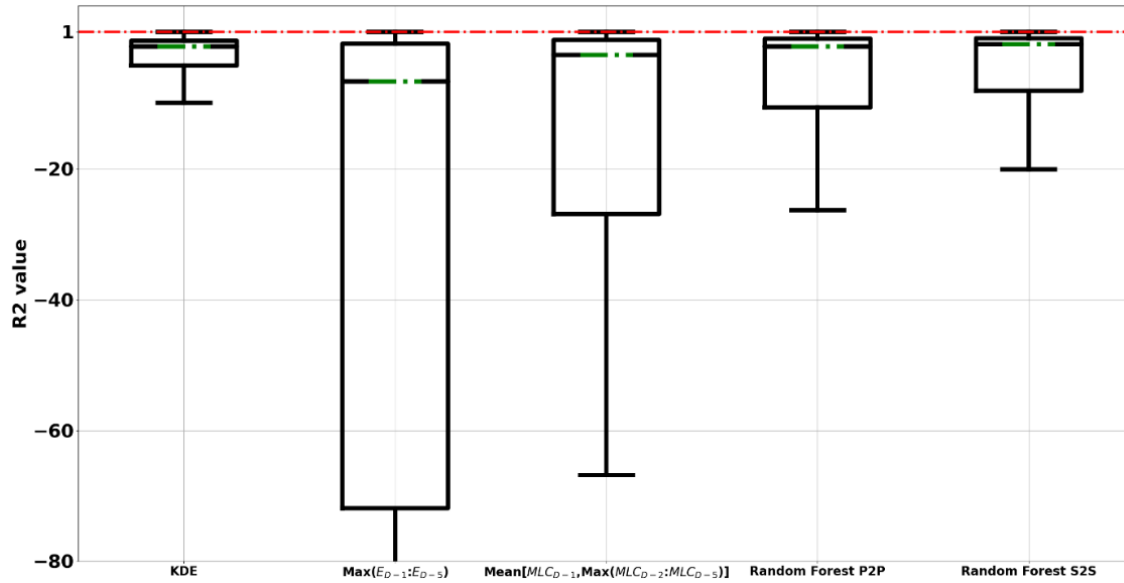


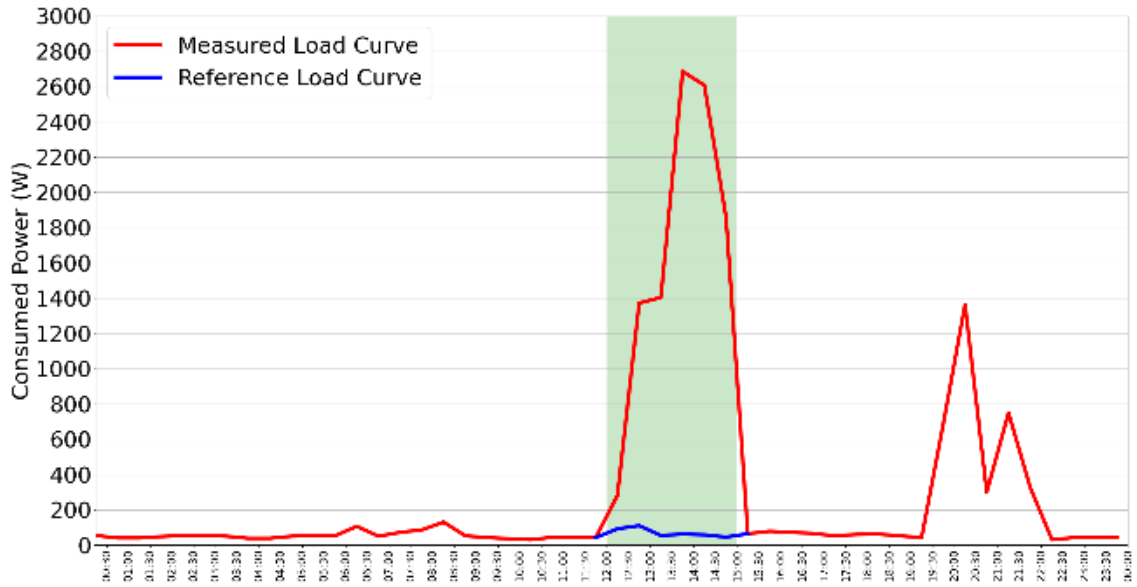
Figure IV.16 Distribution of R^2 score for orange reference load curves

IV.4 Reference load curve in the form of feedback

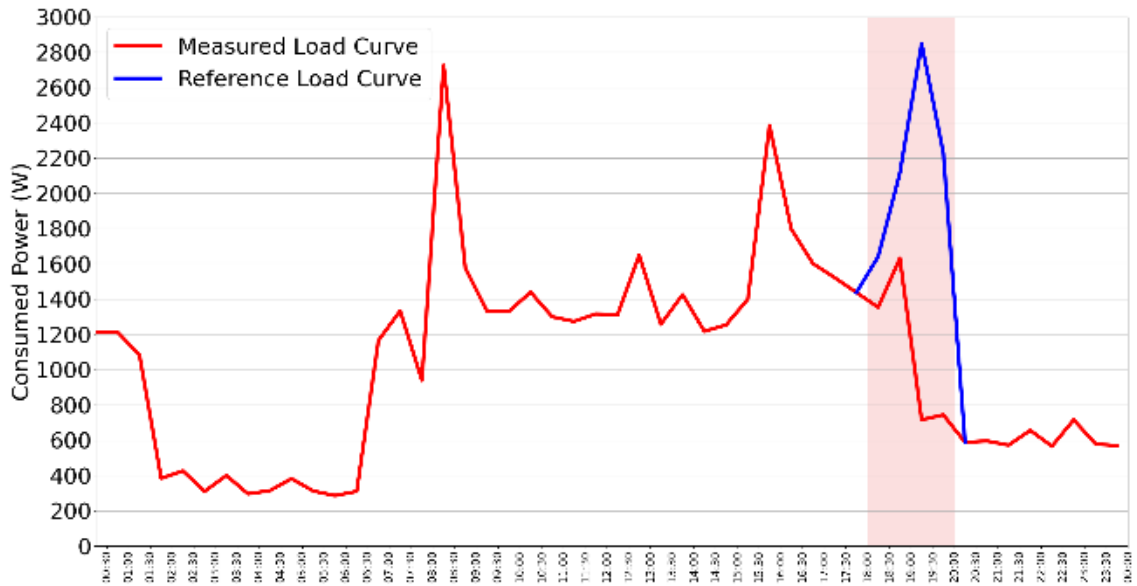
In our experiment, the initial purpose of reference load curve is to present it as an easy and quick to understand graphical feedback for alert days. Figure IV.17 demonstrates the use of reference load curve for both type of alerts as user interface design of energy feedback.

Formerly, the two methods presented in section IV.3.b.ii and section IV.3.b.iii are analyzed whereas latterly, the two random forest methods are also analyzed. Since the following method (already presented in section IV.3.b.iii) is found suitable in the former analysis, therefore, it is retained to formulate alert day reference load curve as feedback.

- For green alert day 'D', the average of measured load curve of 'D-1' and measured load curve having lowest energy consumption between day 'D-2' and day 'D-5'.
- For orange alert day 'D', the average of measured load curve of 'D-1' and measured load curve having highest energy consumption between day 'D-2' and day 'D-5'.



(a)



(b)

Figure IV.17 Indirect feedback using reference load curve for (a) Green Alert and (b) Orange Alert

As compared to displays and bills, the web-based feedback is found more effective as a tool of feedback for load curtailment (Vassileva et al., 2012). On day ‘D+2’ of each alert day ‘D’ and subjecting to the provision of measured load curve of alert day, an image of reference load curve superposed on measured load curve (of alert day) is uploaded on the personal space of subjects of treated group on *Etud-Elec* website. A total number of 2141 feedback images are created and presented to subjects of treated group individually on their personal account on *Etud-Elec* site. This does not include the 2nd green alert of 2nd session, for which the measured load curve data was not recorded on *ENEDIS* server owing to a technical problem at their end. On day ‘D+3’ after each alert, the subjects of treated group

are sent an SMS stating that they can view feedback of their consumption on alert day ‘D’ by logging in to their personal account.

Through the log of *Etud-Elec* site, it is observed that the subjects of treated group only visited their personal account 57 times in total. In percentage term, it makes 2.7% (57/2141) of the total feedback images. The lower number of visits on website might be owing to the time of day of receiving the SMS. Generally, the SMS was sent at noon on day ‘D+3’, however for certain alerts, it was sent at any time of day. The low number of views of indirect feedback makes it insufficient to statistically analyze the impact of reference load curve as a feedback. Besides serving the purpose of embellishing graphical feedback, the reference load curve also serves to categorize the subjects of treated group with respect to their response towards the nudge signals. However, before doing so, it is significant to determine whether the reference load curve is equally effective.

IV.5 Impact of error $\epsilon_{\text{non-alert}}$ in the calculation of energy flexibility

We observed the impact of under-estimation and over-estimation of reference load curve (with respect to measured load curve) for non-alert days in section IV.3.a.i. If an energy difference (owing to underestimation or overestimation) persists for non-alert days, its impact is counted in the calculation of energy flexibility. Therefore, it should be removed for corrected measure of energy flexibility. For the sake of convenience, we call;

- The underestimation (or overestimation) on non-alert days as error $\epsilon_{\text{non-alert}}$.
- The energy flexibility achieved on alert day as $E_{\text{flexibility,alert}}$.

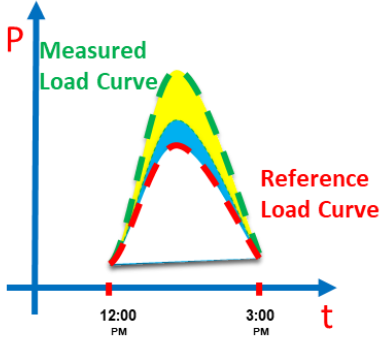
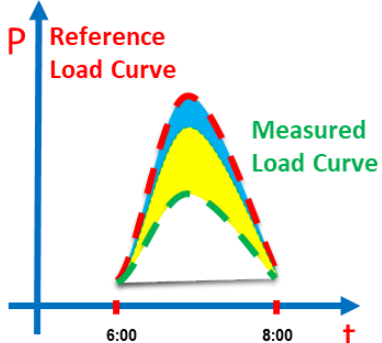
To measure the energy flexibility $E_{\text{flexibility,alert}}$ on alert day, the error $\epsilon_{\text{non-alert}}$ should be removed from the energy difference of reference load curve and measured load curve on alert day. Mathematically, these can be given as;

$$\epsilon_{\text{non-alert}}|_{t_1}^{t_2} = E_{\text{reference,non-alert}}|_{t_1}^{t_2} - E_{\text{measured,non-alert}}|_{t_1}^{t_2} \quad \text{Equation IV.4}$$

$$E_{\text{flexibility,alert}}|_{t_1}^{t_2} = \left(E_{\text{reference,alert}}|_{t_1}^{t_2} - \epsilon_{\text{non-alert}}|_{t_1}^{t_2} \right) - E_{\text{measured,alert}}|_{t_1}^{t_2} \quad \text{Equation IV.5}$$

where t_1 and t_2 are extreme timestamps of the period of interest (or intervention). The case of underestimating green load curve as well as overestimating orange load curve is illustrated graphically in Table IV.3 below. It should be noted that the area highlighted in blue represents $\epsilon_{\text{non-alert}}$, whereas the area highlighted in yellow represents $E_{\text{flexibility,alert}}$.

Table IV.3 Green and orange load curves to measure energy flexibility $E_{flexibility,alert}$

Sr. No.	Case	Graphical Illustration	Sign of energy difference indicator
01	Underestimating reference load curve (for green alerts)		$\epsilon_{non-alert} = \text{Negative}$ $E_{flexibility,alert} = \text{Negative}$
02	Overestimating reference load curve (for orange alerts)		$\epsilon_{non-alert} = \text{Positive}$ $E_{flexibility,alert} = \text{Positive}$

IV.6 Analysis regarding reference load curve as an alternative of control group

As mentioned earlier, it is observed in the section of meta-analyses that the energy consumption of treated group is compared to that of control group. The analyses are done on group level, rather than on individual level for each subject. The meta-analysis emphasizes the requirement of control group to measure the impact of treatment. Here it raises a scientific question whether it is possible to replace the control group with another tool that is able to measure the impact of treatment on individual level.

To explore this question, we put the reference load curve in place of control group and compare it with the results in section III.5.b. In this regard, the following two methods of formulation of reference load curve are analyzed. The results of former method is presented and discussed below whereas the results of later method is presented in Annex D and briefly discussed here.

- The average of measured load curve of ‘D-1’ and the measured load curve having
 - o Lowest energy consumption between day ‘D-2’ and day ‘D-5’ (for green alert).

- Highest energy consumption between day ‘D-2’ and day ‘D-5’ (for orange alert).
- Random forest sequence to sequence
- *One-sided Mann-Whitney U Test* is used with the following hypotheses.
- Null hypothesis (H₀): The distribution (of energy consumption data) of control group is statistically **greater** than the distribution (of energy consumption data) of treated group.
- Alternative hypothesis (H₁): The distribution (of energy consumption data) of control group is statistically **less** than the distribution (of energy consumption data) of treated group.

In the case of green alert days, our interest is to find out whether the null hypothesis is rejected for the energy consumption data of alert days. Statistically, it can be observed if the p-value of *Mann-Whitney U Test* is less than $\alpha = 0.05$, which validates our alternative hypothesis. Table IV.4 shows the results for the reference load curve used in the experiment, where ‘ μ ’ represents mean value, ‘ σ ’ represents the standard deviation, subscript ‘r’ and subscript ‘t’ represents reference load curve and measured load curve of treated group respectively. It should be noted that the effect of error $\epsilon_{\text{non-alert}}$ is removed from the reference load curve of alert days in all the analyses given below.

Table IV.4 The Mann-Whitney U Test for green alert days of both sessions using $Mean[P_{D-1,i}, Min(P_{D-2,i}, P_{D-3,i}, P_{D-4,i}, P_{D-5,i})]$ reference load curve in place of control group

S r. N o	Alert Date	Reference load curve			Treated Group			Mann-Whitney U Test			$\mu_r - \mu_t$
		Obs	μ_r	σ_r	Obs	μ_t	σ_t	Stat	p-value	Null hypothesis retained	
Green Session 1											
1	27-07-2020	70	0.71	0.67	71	0.99	1.21	2296	0.22	Yes	-0.28
2	25-08-2020	7	0.66	0.45	34	0.81	0.64	110	0.38	Yes	-0.15
3	14-09-2020	82	0.71	0.76	88	0.79	0.83	3607	0.5	Yes	-0.08
4	17-09-2020	88	0.71	0.54	88	0.77	0.68	3894	0.53	Yes	-0.06
5	22-09-2020	88	0.68	0.49	88	0.79	0.96	4053	0.7	Yes	-0.11
Green Session 2											
6	14-06-2021	85	0.74	0.62	85	1.01	1	3328	0.19	Yes	-0.27
7	06-09-2021	-	-	-	-	-	-	-	-	-	-
Aggregated Result											
Aggregated		420	0.71	0.62	454	0.86	0.93	93171	0.28	Yes	-0.15

The results of Table IV.4 are found considerably different than the results of Table III.7. A comparison of control group with treated group (in Table III.7) suggests that the distribution of control group is statistically less than the distribution of treated group for 5th green alert and all green alerts combined. On the other hand, the results in Table IV.4 shows that the distribution of reference load curve is statistically greater than the distribution of treated group for all alerts (and at aggregated level). The results in Table D.1 also shows same results. Therefore, the results do not emphasize that a reference load curve can be used as a replacement of control group to measure the load shifting of energy flexibility.

Table IV.5 The Mann-Whitney U Test for orange alert days of both sessions using $Mean[P_{D-1,i}, Max(P_{D-2,i}, P_{D-3,i}, P_{D-4,i}, P_{D-5,i})]$ reference load curve in place of control group

Sr. No	Alert Date	Reference load curve			Treated Group			Mann-Whitney U Test			$\mu_r - \mu_T$
		Obs	μ_r	σ_r	Obs	μ_T	σ_T	Stat	p-value	Null hypothesis retained	
Orange Session 1											
1	14-11-2019	91	1.68	2.44	91	1.32	2.41	4968	0.99	Yes	0.36
2	20-12-2019	90	1.03	1.04	90	0.92	0.74	4415	0.85	Yes	0.11
3	20-01-2020	92	1.24	1.35	92	1.3	1.57	4317	0.59	Yes	-0.06
4	06-02-2020	91	1.37	1.4	91	1.11	1.31	4896	0.98	Yes	0.26
5	19-02-2020	91	1.29	1.11	91	1.19	1.29	4680	0.94	Yes	0.1
6	27-02-2020	90	1.19	1.17	91	1.23	1.32	4307	0.73	Yes	-0.04
7	02-03-2020	90	1.19	1.22	90	0.93	1.07	4828	0.99	Yes	0.26
8	04-03-2020	91	1.2	1.4	91	0.98	1.06	4698	0.94	Yes	0.22
9	09-03-2020	90	0.88	1.32	90	1.26	1.91	3501	0.06	Yes	-0.38
Orange Session 2											
10	30-11-2020	88	1.39	1.1	88	1.21	0.9	4236	0.86	Yes	0.18
11	02-12-2020	89	1.53	1.47	89	1.4	1.06	4073	0.63	Yes	0.13
12	04-12-2020	89	1.37	1.3	89	1.25	1.15	4204	0.76	Yes	0.12
13	07-12-2020	89	1.29	1.19	89	1.4	1.43	3886	0.42	Yes	-0.11
14	09-12-2020	89	1.31	1.11	89	1.48	1.32	3814	0.34	Yes	-0.17
15	18-12-2020	88	1.47	1.4	88	1.19	1.05	4410	0.94	Yes	0.28
16	04-01-2021	88	1.53	1.68	88	1.57	1.79	3992	0.64	Yes	-0.04
17	06-01-2021	88	1.64	1.54	88	1.61	2.11	4297	0.9	Yes	0.03
18	08-01-2021	88	1.62	1.9	88	1.4	1.97	4301	0.9	Yes	0.22
19	11-01-2021	88	1.46	2.09	88	1.75	1.92	3361	0.06	Yes	-0.29
20	26-01-2021	87	1.37	1.26	87	1.37	1.23	3818	0.54	Yes	0
21	11-02-2021	87	1.4	1.63	87	1.23	1.44	4036	0.78	Yes	0.17
Aggregated Result											
Aggregated		1874	1.35	1.49	1875	1.29	1.51	1866790	1	Yes	0.06

In the case of orange alert days, our interest is to find out whether the null hypothesis is retained for the energy consumption data of alert days. Table IV.5 shows the results where the test is performed on the reference load curve used in this experiment as compared to the measured load curve of treated group. The number of observations for both groups are same. The difference of mean values ' $\mu_r - \mu_T$ ' in Table IV.5 is found to be slightly higher than the difference of mean values ' $\mu_C - \mu_T$ ' in Table III.8; albeit having more observations of reference load curve in Table IV.5 as compared to observations of control group in Table III.8. Yet, the results of *one-sided Mann-Whitney U Test* in Table IV.5 are found closer to the results of Table III.8.

The technique of random forest sequence to sequence is found to be the best technique to formulate a reference load curve for load curtailment (i.e. orange alerts) in our analysis presented in section IV.3.c.ii. The results of this technique is presented in Table D.2 in appendix D. The value of ' $\mu_r - \mu_T$ ' for certain alerts in Table D.2 is negative, although the *one-sided Mann-Whitney U Test* results show that the null hypothesis is retained for these alerts. Overall, the results of Table IV.5 and table D.2 in comparison with Table III.8 emphasizes that the reference load curve can replace control group to measure the load curtailment of energy flexibility.

In a nutshell, the *one sided mann whitney test* does not emphasize that a reference load curve can replace control group to measure load shifting. However, the result of same test emphasizes that the reference load curve can replace control group to measure load curtailment. Since it is demonstrated that the reference load curve can effectively measure load curtailment, therefore in the following section, an analysis of measure of energy flexibility (i.e. load curtailment) with respect to the reference load curve is presented for individual subjects.

IV.7 Measure of response of individual subjects towards orange alerts with respect to the reference load curve

It is discussed earlier that the energy flexibility against a treatment is measured through the comparison of treated group against a control group. This analysis does not give information on individual level for each subject of treated group. Therefore, the results of group are generalized for each subject. Contrary to this, the reference load curve is calculated for each individual subject. Subjected to the validation of reference load curve, it permits to analyze the response of each subject towards the call of energy flexibility. Therefore, the subjects of treated group can be identified as those who:

- acted positively in response to nudge signals.
- did not act positively in response to nudge signals.

- consumption did not change (or negligibly changed) in response to nudge signals OR the response to nudge signals cannot be determined.

To analyze this problem, the following methodology is adopted.

- The reference load curve that we used in our experiment is taken for analysis.
- Since the formulation of reference load curve is based on the historical consumption of previous 5 weekdays of alert day 'D', therefore for the sake of comparison, the energy consumption data of these days is taken.
- An average error $\epsilon_{\text{non-alert,average}} |_{6 \text{ PM}}^{8 \text{ PM}}$ is calculated for the error $\epsilon_{\text{non-alert,D-n}} |_{6 \text{ PM}}^{8 \text{ PM}}$ of each non-alert day between 'D-1' to 'D-5', where $n = 1$ to 5 .
- For each subject of treated group, the $E_{\text{flexibility,alert}} |_{6 \text{ PM}}^{8 \text{ PM}}$ is calculated using Equation IV.5, in which $\epsilon_{\text{non-alert}} |_{6 \text{ PM}}^{8 \text{ PM}}$ is replaced by $\epsilon_{\text{non-alert,average}} |_{6 \text{ PM}}^{8 \text{ PM}}$.
- The results are presented in the unit of average energy per hour (Wh) during the period of intervention. The results are only presented for those subjects of treated group whose reference load curve is successfully calculated for all 21 orange alerts.
- For each subject and orange alerts altogether, the results are presented in the form of boxplots.

To identify how the subjects of treated group responded towards the call of energy flexibility, it is imperative to fix a threshold. There is no universal rule for fixing threshold. Therefore, a sensitivity analysis is performed for multiple threshold values and is presented in Table IV.6. It is observed that the number of undetermined subjects increases with increasing the threshold value, whereas the values of other two variable consequently decreases. It suggests that a suitable threshold should be selected after due diligence of the appliances and the potential of energy flexibility.

Table IV.6 Sensitivity analysis of thresholds of identifying responses

Sr. No.	Response	Criterion	Number of identified subjects
Lower threshold = -50Wh, Upper threshold = 50Wh			
01	Positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} > 50 \text{ Wh}$	36
02	Not positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} < -50 \text{ Wh}$	27
03	Cannot be determined	$-50 \text{ Wh} \leq E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} \leq 50 \text{ Wh}$	16
Lower threshold = -100Wh, Upper threshold = 100Wh			
04	Positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} > 100 \text{ Wh}$	29
05	Not positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} < -100 \text{ Wh}$	18
06	Cannot be determined	$-100 \text{ Wh} \leq E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} \leq 100 \text{ Wh}$	32
Lower threshold = -150Wh, Upper threshold = 150Wh			

07	Positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} > 150 \text{ Wh}$	25
08	Not positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} < -150 \text{ Wh}$	10
09	Cannot be determined	$-150 \text{ Wh} \leq E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} \leq 150 \text{ Wh}$	44
Lower threshold = -200Wh, Upper threshold = 200Wh			
10	Positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} > 200 \text{ Wh}$	20
11	Not positive	$E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} < -200 \text{ Wh}$	8
12	Cannot be determined	$-200 \text{ Wh} \leq E_{\text{flexibility,alert}} _{6 \text{ PM}}^{8 \text{ PM}} \leq 200 \text{ Wh}$	51

For the sake of presentation, the threshold values of -200Wh/200Wh are explored. The average value of distribution in the box plot is considered for analysis of $E_{\text{flexibility,alert}}|_{6 \text{ PM}}^{8 \text{ PM}}$. It is recalled here that the difference is calculated between reference load curve and measured load curve, therefore a positive value indicates load curtailment whereas a negative value indicated no positive effort for load curtailment during the period of intervention.

Based on the selected threshold, Figure IV.18 demonstrates the distribution of energy flexibility of the subjects who acted positively for most of alerts. The black dots represent the outliers in the distribution made for boxplot, whereas the blue dotted line represents the threshold value of 200 Wh. It can be observed that the median value of subject 399 and subject 404 lies beneath the threshold line, yet the outliers made the average energy flexibility of these subjects greater than the threshold value. Therefore, these subjects are considered as positively acting subjects against the nudge signal of load curtailment. In total, 20 (out of 79) subjects are found in this category.

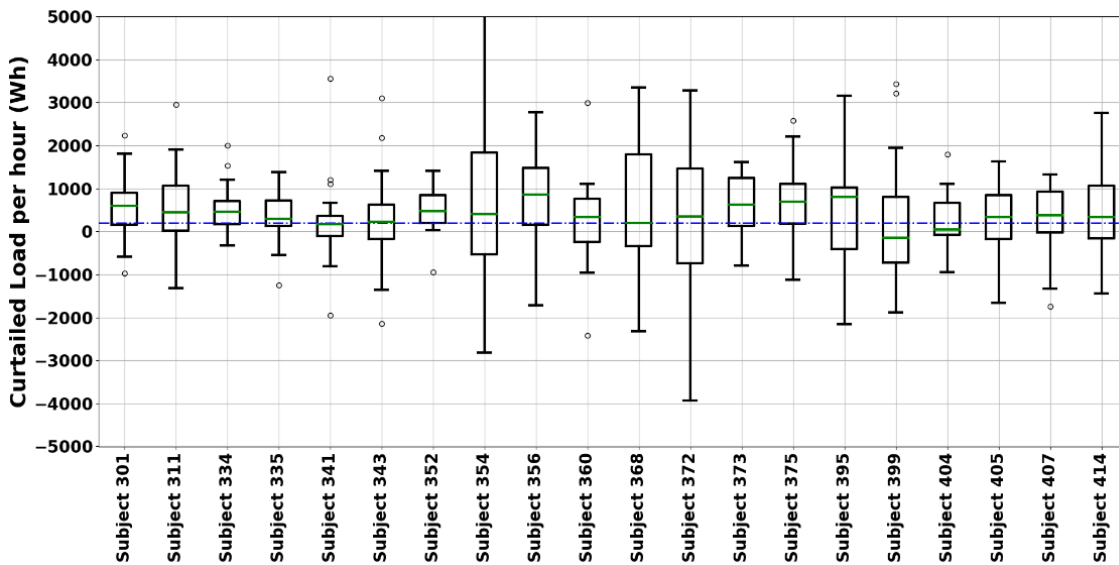


Figure IV.18 The energy flexibility of subjects who curtailed their load for most of the orange alerts

According to the selected threshold values, Figure IV.19 demonstrates the distribution of energy flexibility of the subjects who did not acted positively for most of alerts. The blue dotted line represents the threshold value of -200 Wh. In total, 8 (out of 79) subjects are found in this category. A particular case in this category is of subject 388, who achieved an average load curtailment of more than 2 kWh (per hour) for 5 orange alerts, however the case is contrary for most of the rest of the orange alerts.

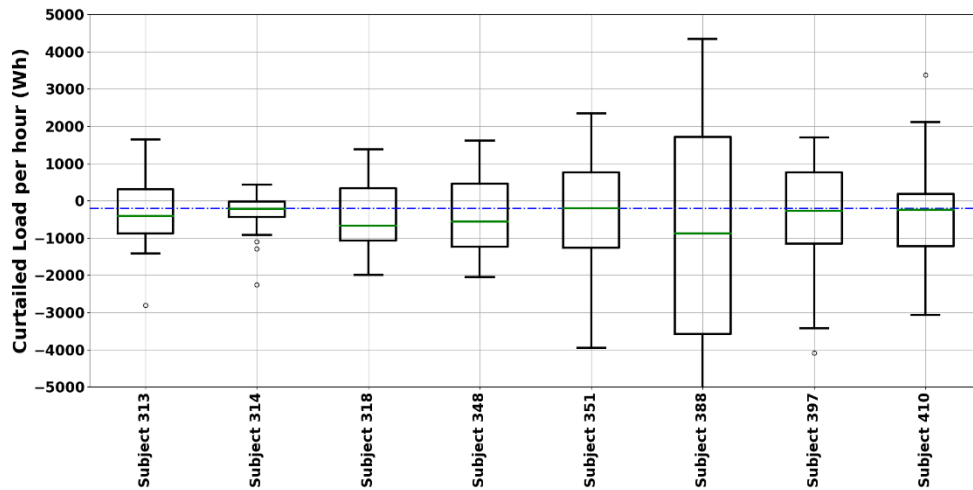


Figure IV.19 The energy flexibility of subjects who did not remain positive for most of the orange alerts

Finally, Figure IV.20 demonstrates the distribution of energy flexibility of the subjects who remained within the range of [-200 Wh, 200Wh]. We can further split it in to 2 sub-categories. The first category represents those whose energy consumption remains stable during alert days as compared to non-alert days. In other words, we can say that the effect is zero or negligible. In this regard, those subjects are identified whose average energy flexibility is in the range of [-25Wh, 25Wh]. In total, 9 (out of 79) subjects are identified in this category. For the rest of the 42 subjects, the energy flexibility activity cannot be determined with respect to the fixed threshold.

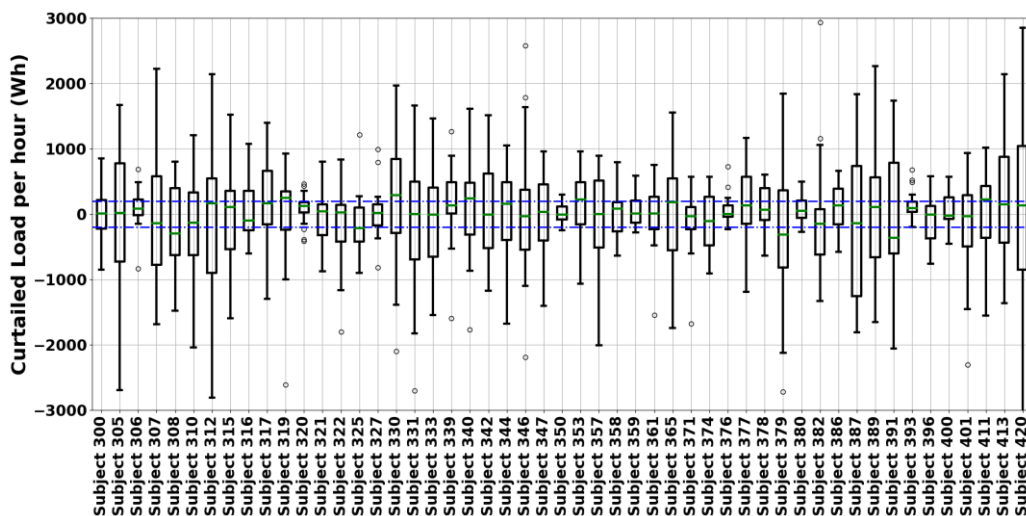


Figure IV.20 The energy flexibility of subjects whose effort of energy flexibility cannot be determined

IV.8 Conclusion

A reference load curve is a predicted load curve on non-alert day. It is assumed that it represents the normal behavior of energy consumption of a household. Its superposition with flexibility driven measured load curve of alert day gives a visual assessment of the effort of energy flexibility. Ideally, the difference of reference load curve with measured load curve should be zero on non-alert day, yet practically it is not possible. Therefore, multiple techniques are analyzed to formulate a reference load curve. By analyzing the distribution of each technique with respect to KPIs, it is concluded that the classical method (using kernel density estimation) performs better than the others for measuring the impact of load shifting on alert day. In the case of load curtailment, it is relatively emphasized by the R2 score indicator that a machine learning based method performed better than the others. Yet it is a future prospect to improve the performance of machine learning model with the introduction of supplementary parameters.

The original purpose of formulating reference load curve is to send as indirect feedback towards the treated group and analyze the impact of this feedback. However, the smaller number of views of indirect feedback makes it insufficient to statistically analyze the impact of reference load curve as feedback. Besides, an aspect of reference load curve is also explored where it is analyzed whether the reference load curve can replace control group in energy flexibility programs. Through our analysis, we are unable to conclude that the reference load curve could replace control group for load shifting, yet it is found effective to replace the control group for load curtailment.

Traditionally, the effort of energy flexibility is assessed on group level by comparing treated group with a control group. However, a suitable and valid reference load curve can serve to measure the effort of energy flexibility at individual level. This individual assessment serves to identify the subjects that have or have not positively responded towards the nudge signal of energy flexibility. Therefore, in complement to the statistical meta-analysis that quantifies the energy flexibility at group level, this instrument serves to quantify and categorize the individual effort of energy flexibility.

Chapter V

The quest of finding potential appliances for energy flexibility in line with subject commitment

This chapter presents a proof of concept regarding detection of potential appliances for energy flexibility in line with the subject commitment. For this purpose, an open source NILM algorithm is explored on multiple datasets of appliance level energy consumption.

V.1 Introduction

In the previous two chapters, we explored the impact of nudge signals on the energy flexibility of the treated group, and we explored an approach where a reference load curve can serve in energy flexibility programs. However, it is interesting to look for the potential appliances for energy flexibility among the appliances that are committed by the treated group. In this chapter, we discuss a proof of concept (and the factors bringing constraints to it) to look for the appliances that have potential for energy flexibility.

We recall here the section I.3.b in which an analysis of various appliances is done. This analysis is performed based on the annual energy consumption and the nominal power of the appliances. It is deduced from this theoretical analysis that the white appliances (i.e. washing machine, dishwasher and cloth dryer) might have highest potential of energy flexibility within a period of intervention. This is owing to their highest nominal power as compared to others. However, it can be observed by using non intrusive load monitoring (NILM). Generally, the energy disaggregation is done on load curves with a much higher granularity (e.g. 1 second, 8 seconds etc.); however the granularity in our experiment is much lower (i.e. 30 minutes). Therefore, we would like to implement it on a 30 minutes resolution. In the following section, the state of the art of non-intrusive energy disaggregation is presented. Whereas, in the rest of the chapter, we will discuss a proof of concept that can be used to detect potential appliances for energy flexibility using the NILM energy disaggregation technique.

V.2 Non-Intrusive Load Monitoring

Energy disaggregation is the process of acquiring the device level energy consumption data; correlated with the global energy consumption data of smart meter. It can be either hardware centric or software defined approach. The former is known as intrusive load monitoring (or in short ILM); in which energy consumption recording sensors are connected to the appliances. In this case, it is easy to acquire the energy consumption data of appliances individually at the expense of costly hardware materials and propriety communication protocols. The later is known as non-intrusive load monitoring (or in short NILM⁴²); in which the appliance level energy consumption data is extracted from the global energy consumption data. Rather than the installation of costly sensors, this approach contends on the machine learning techniques to disaggregate the global consumption data.

⁴² It is also sometimes referred as non-intrusive appliance load monitoring or NIALM in short.

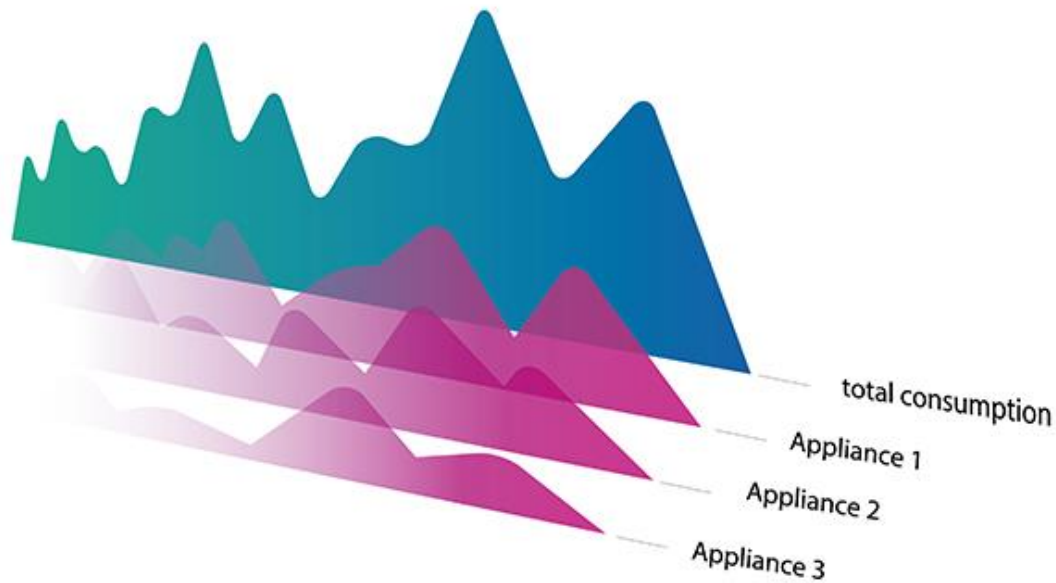


Figure V.1 Non-intrusive load monitoring (Parson, 2017)

NILM is initially studied in 1992 through a detailed analysis of current and voltage of smart meter data (Hart, 1992). Since the mass deployment of smart meters, the research work in NILM become more significant. The state of the art discusses two aspects of NILM, namely; open-source datasets and the algorithms. These are briefly discussed below.

V.2.a Energy disaggregation datasets

The intrusive load monitoring significantly complements the NILM approach to gather data. Using the ILM approach, public (and private) datasets of energy consumption data are collected. These datasets are used to train NILM models and to validate these models via performance metrics. In various research works, the data is collected for a sampling period of as low as 1 second and as high as 30 minutes throughout the world. Table V.1 illustrates some open-source datasets for residential sector.

Table V.1 Some open source NILM datasets

Sr. No.	Name of dataset	Country of origin	Number of households	Duration of data (yrs=years, m=months)	Sampling period (min=minutes, sec=seconds)
01	REDD (Kolter & Johnson, 2011)	USA	6	Several months	1 sec
02	DEDDIAG (Wenninger et al., 2021)	Germany	15	3.5 yrs	1 sec
03	UKDALE(J. Kelly & Knottenbelt, 2015)	UK	5	> 1yr	6 sec

04	REFIT (Murray et al., 2017)	UK	20	2 yrs	8 sec
05	Individual household electric power consumption Data Set (Hebrail & Berard, 2012)	France	1	4 yrs	1 min
06	AMPDs (Makonin et al., 2016)	Canada	1	2 yrs	1 min

V.2.b NILM Techniques

NILM is widely considered as a single channel blind source separation problem. Most of the research work tackles NILM as a regression problem, in which the individual models of appliances predict the energy consumption based on the global consumption as input. The state of the art also generally prefers supervised learning over non-supervised learning. However, some research work also tackles it as a classification problem, where the appliance (ON/OFF) state is detected and classified using the energy signature of the appliance (Basu et al., 2015). In essence, it should be a combination of both in which the appliance state should be detected and then a prediction regarding its consumption should be made.

The techniques of NILM are categorized as optimization problem, probability statistics, graph signal processing, machine learning and deep learning (Zhuang et al., 2018). Some research work of these techniques are given as follows:

- **Optimization problem:** Aided linear integer programming (Bhotto et al., 2017) and evolutionary optimization algorithm (Egarter et al., 2013) are among optimization techniques studied for NILM.
- **Probability Statistics:** Hidden Markov model is a widely studied probability statistics technique for NILM (H. Kim et al., 2011; X. Wang et al., 2018; Zhong et al., 2014). Among probability statistics techniques, Viterbi algorithm (Zeifman, 2012) and Bayesian approach (Srinivasarengan et al., 2013) are also studied.
- **Graph Signal Processing:** NILM research work is also explored using graph signal processing (He et al., 2018; Stankovic et al., 2014; Zhao et al., 2015, 2016).
- **Machine Learning:** The machine learning techniques are used for pattern recognition in NILM. Random Forest (Y. Liu et al., 2021; Wu et al., 2019) and support vector machine (Moradzadeh et al., 2020; Schirmer & Mporas, 2019) are among the machine learning techniques that are studied.
- **Deep Learning:** Various deep learning architectures are also extensively researched for the purpose of non-intrusive load monitoring. In this regard, various architectures of auto-encoder (Bonfigli et al., 2018; Tsai et al., 2018), recurrent

neural networks (J. Kim et al., 2016; Le et al., 2016) and convolutional neural network (D. Kelly, 2016; Penha & Castro, 2017) are studied.

The application of recurrent neural network is also compared with convolutional neural network (Pedro Paulo Marques do Nascimento, 2016). Most of the research in the field of energy disaggregation yields individual model for each appliance. However, the research is also done in order to make a unique model for disaggregation of multiple appliances (Faustine et al., 2020). However, this research is limited.

V.2.c Remarks on the state of the art

Following observations are made in the state of the art of non-intrusive load monitoring.

- The research on energy disaggregation approaches is extensively tested on 4 publicly available datasets. i.e. REDD, REFIT, AMPDs and UK-DALE.
- It is not found that energy disaggregation models are trained on an ensemble of datasets. Rather, the models are trained on individual datasets separately.
- The energy disaggregation algorithms are seldom available publicly for reproducibility of results.
- There does not exist a universal energy disaggregation algorithm. Rather, each research article explores the efficiency of its proposed algorithm. The efficiency is presented against one (or multiple) benchmark algorithms using chosen key performance indicators.
- The models trained and tested on 1 dataset might not be equally efficient for other datasets. This is an open research question under investigation regarding transfer learning. This point will be further investigated in section V.6 and section V.7.
- A number of benchmark algorithms tested on publicly available datasets are available by the name of *NILM-TK*⁴³ (Batra et al., 2014). However, it is found difficult to reproduce the results of these benchmark algorithms for a dataset other than those available in toolkit. Besides, it has the limitation of disaggregating for 5 highest consuming appliances in the household.

Before discussing the proof of concept using NILM, the methodology of exploring an open-source algorithm is presented in section below.

⁴³ NILM-TK is publicly available at <https://github.com/nilmk/nilmk>.

V.3 Methodology

The following schematic diagram shows the methodology adapted to evaluate the models for energy disaggregation.



Figure V.2 Schematic diagram of energy disaggregation of measured load curves on alert days

V.4 Committed appliances by the subjects

As mentioned in Section II.3.c.ii, the information regarding the presence of diverse appliances present in the households is collected. In addition to this, the information regarding the time of use of appliances with respect to defined timeslots is also collected. The purpose of collecting this information and having a special session with the subjects is to make the energy consumption and the nominal power of appliances visible to the subjects. Once the subjects of treated group are enlightened about the impact of individual appliance, they are requested to choose appliances from a given set that they can commit to use (in response to green alert) or not use (in response to orange alert).

Figure V.3 shows the frequency of committed appliances by the treated group for both green alerts and orange alerts respectively. For instance, the washing machine is the appliance that is mostly committed by the subjects of treated group for both green alerts and orange alerts. The commitment frequency of washing machine is followed by dishwasher and electronic gadgets. It is significant to see that the treated group committed the cooking appliances mostly for load curtailment. It is the same case for space heating appliances, however the frequency of commitment of space heating is less as compared to the cooking appliances.

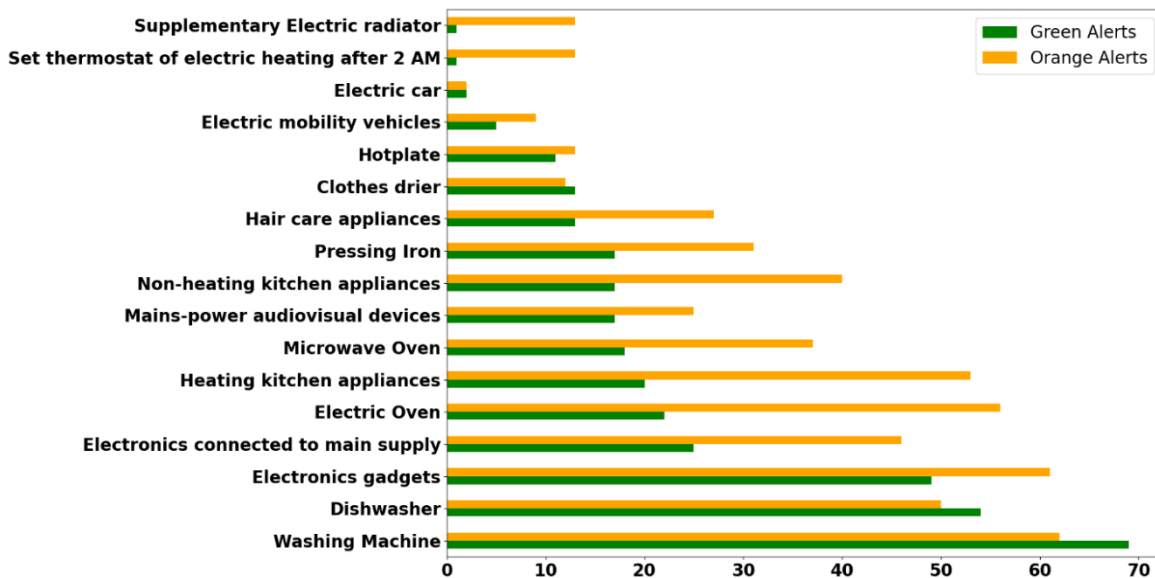


Figure V.3 Frequency of appliances committed for green alert

The subjects of treated group are given a choice to change their default commitment before the alert day. However as already mentioned in section IV.4, it is observed that the subjects of treated group very rarely logged in to their personal account on Etude-Elec. A total of 8 modifications were made for all the alerts. The low number of modifications of indirect feedback makes it insufficient to statistically analyze the impact of modified commitment against default commitment. Therefore, for the sake of argument, we consider that the subjects of treated group stick to their default commitment reminded to them by the commitment SMS.

Since the frequency of appliance commitment of the experiment is analyzed, it is now significant to explore an energy consumption dataset. This dataset serves to train the model of respective appliances for energy disaggregation. The dataset must have maximum number of appliances in common with the committed appliances of the experiment. For this purpose, IRISE dataset is used, which is demonstrated in the following section. This database is not open-source, therefore it is not mentioned in the Table V.1.

V.5 The characteristics of IRISE dataset

IRISE dataset is a subset from European database of residential consumption as part of *REMODECE project*⁴⁴ (Enertech, 2008), for residential monitoring to decrease energy use and carbon emissions in Europe (Basu et al., 2015). The dataset is private in nature. As part of *IRISE* dataset, the energy consumption data of 98 houses for a span of a year is recorded. The data is sampled by 10 minutes. The dataset has data of site consumption as well as energy consumption of individual appliances. The site consumption also includes the unmeasured load which were either intentionally (or non-intentionally) added to the dataset. Therefore, in addition to the site consumption, the aggregated consumption is calculated by summing up the energy consumption of known appliances at each instance.

V.5.a Committed appliances of treated group in the IRISE dataset

Before using *IRISE* dataset, it is significant to find whether the energy consumption data of the committed appliances by the treated group is available in the *IRISE* dataset. In this regard, the following 6 appliances are exactly found in the *IRISE* database. It is more probable that these appliances could be easily traced in the energy consumption data of treated group by using a trained model on *IRISE*. The frequency of presence of these appliances in *IRISE* households is given in Figure V.4.

- Electric Oven
- Washing Machine
- Dishwasher
- Microwave Oven
- Hotplate
- Cloth dryer

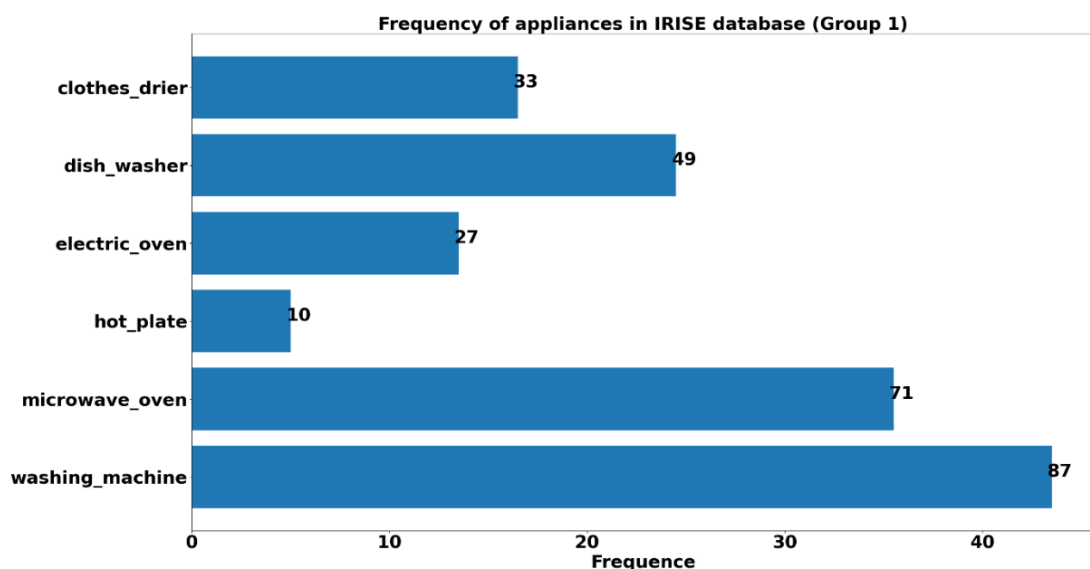


Figure V.4 Frequency of matched appliances in IRISE dataset with committed appliances

⁴⁴ Information regarding integration of IRISE dataset in REMODECE project:
https://remodece.isr.uc.pt/database/Campaign_Irise.html

The following 5 committed appliances are found closer to certain appliances found in *IRISE* database. In the EXPESIGNO dataset, these appliances are grouped under the name of a) Electronics connected to main supply, b) Heating kitchen appliances, c) Mains-power audiovisual devices, d) Set thermostat of electric heating after 2 AM and e) Supplementary Electric radiator. The frequency of presence of these appliances in *IRISE* households is given in Figure V.5.

- Computer site
- Electric Deep Fryer
- TV
- Electric Cooker
- Electric Heating

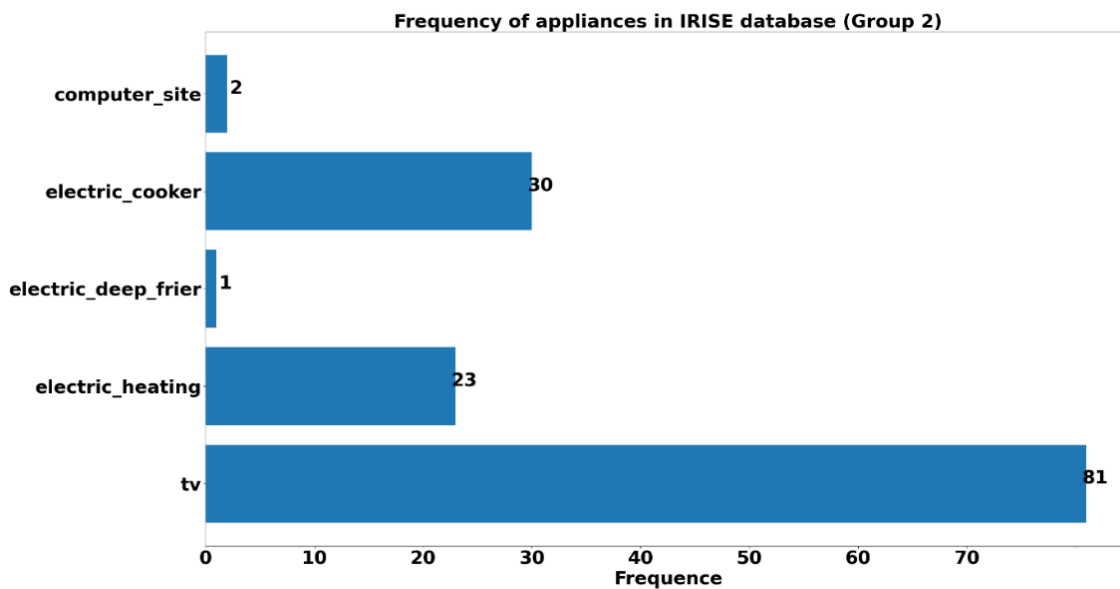


Figure V.5 Frequency of closely matched appliances in IRISE dataset with committed appliances of treated group

Beside the group of appliances, the following 3 appliances are not found in *IRISE* database. Since the *IRISE* dataset is recorded between 1999 and 2001, therefore it is obvious that the energy consumption data of electric vehicle and electric mobility could not be found in the *IRISE* dataset.

- Pressing Iron
- Electric vehicle
- Electric mobility

V.5.b The statistics of appliances in IRISE dataset

Before going further, it is interesting to look at the statistical variables of each appliance. For each appliance, these variables are calculated for their respective data present in *IRISE* households. Table V.2 represents the statistics of *IRISE* appliances at 10 minutes sampling. Since zeros represent that the appliance is *OFF*, therefore it is omitted during the calculation of mean, median and standard deviation. The aggregated consumption in the table represents the aggregated consumption of the appliances below in

all the houses in which the appliances are present. The average window size represents the average of the various lengths of duration when the appliance consumed energy.

Table V.2 Statistics for IRISE appliances at 10 minutes sampling

Appliance	Count	ON (%)	OFF (%)	Mean (W)	Standard Deviation (W)	Median (W)	Max (W)	Average window size
Aggregated	5103252	x	x	176	365	76	x	x
Hotplate	12817	2	98	109	87	88	604	3
Electric Heating	594695	44	56	151	187	90	10924	23
Electric Oven	137399	9	91	56	93	21	4246	29
Electric Cooker	179874	11	89	52	95	3	1259	9
Clothes drier	102373	5	95	137	138	103	1287	11
Dishwasher	85566	3	97	174	161	119	628	5
Microwave	153982	4	96	39	49	23	587	3
TV	1703741	24	76	11	16	9	856	13
Washing Machine	301406	6	94	73	107	22	5866	7

The data of energy consumption in our (EXPESIGNO) experiment is collected by LINKY smart meter, which is sampled at 30 minutes. To keep the homogeneity in the sampling rate of IRISE with EXPESIGNO, the power consumption data of IRISE appliances are down sampled at 30 minutes by taking the mean value for half hourly timestamps. The statistics of IRISE appliances are calculated and presented in Table V.3.

Table V.3 Statistics for IRISE appliances at 30 minutes sampling

Appliance	Count	ON (%)	OFF (%)	Mean (W)	Standard Deviation (W)	Median (W)	Max (W)	Average window size
Aggregated	1716530	x	x	174	293	78	x	x
Hotplate	6759	4	96	69	61	49	434	3
Electric Heating	211074	47	53	142	174	77	3641	35
Electric Oven	48718	10	90	52	77	22	1421	11
Electric Cooker	72977	14	86	43	73	6	623	5
Clothes drier	39331	5	95	119	122	87	1082	5
Dishwasher	39754	5	95	126	105	126	469	3
Microwave	105928	8	92	19	29	11	355	3
TV	647996	27	73	9	10	8	856	7
Washing Machine	129679	8	92	57	74	22	1955	3

By implementing down sampling (by taking average) on the data, it is common to find the maximum value of energy consumption different than the maximum value of original sampling. It is because the maximum value of original sampling is averaged with its adjacent values, therefore it is modified. However, it is expected that the mean consumption should remain same irrespective of the down-sampling. By comparing Table V.2 and Table V.3, it is found out that the mean values vary to a considerable extent for hotplate, dishwasher and microwave. The mean values are highlighted in yellow in the tables. A possible explanation of this variation lies in the consumption pattern of these appliances. These appliances have very short time of peak consumption, while the rest of the time they have low consumption during their cycle of operation. Therefore, the mean value varies for these appliances upon down sampling.

V.6 Reproducibility of an open-source energy disaggregation algorithm

To train and test models of energy disaggregation on *IRISE* dataset, an open-source algorithm known as *Transfer-NILM*⁴⁵ is used. It uses *sequence to point*⁴⁶ architecture of neural networks. This algorithm deals with the transfer learning of energy disaggregation. The transferability of model is tested for appliance level (i.e. model trained on one appliance is used to do energy disaggregation on another appliance) and cross dataset level (i.e. model of an appliance trained on one dataset and tested on the same appliance from another dataset) (D’Incecco et al., 2020).

Before training models on *IRISE* dataset using *transfer-NILM* algorithms, the claimed results in the research article should be verified. For this purpose, the results are reproduced for washing machine and microwave oven on the open-source *REFIT*⁴⁷ dataset (already mentioned in Table V.1). To check the cross-dataset results, the training is done on *REFIT* dataset whereas it is tested on *REDD*⁴⁸ dataset (already mentioned in Table V.1). The length of data subsets is given in Table V.4. The data of *REDD* is down sampled and the same length is cropped as of *REFIT* dataset. Previously, we have observed that the *IRISE* dataset has a default resolution of 10 minutes, whereas to bring homogeneity with the resolution of *LINKY* data, *IRISE* dataset is resampled at 10 minutes. Similarly, to bring homogeneity in the resolution of *REFIT* (by default at 8 seconds) and *REDD* (by default at 1 seconds), the *REDD* dataset is down sampled to a resolution of 8 seconds.

⁴⁵ It is available at <https://github.com/MingjunZhong/transferNILM>.

⁴⁶ The python libraries of tensorflow and keras are used for algorithm formulation.

⁴⁷ Data source: <https://pureportal.strath.ac.uk/en/datasets/refit-electrical-load-measurements-cleaned>

⁴⁸ Data source: <http://redd.csail.mit.edu/>

Table V.4 The length of data subsets in REFIT and REDD dataset

Sr. No.	Appliance	REFIT and REDD data sampled at 8 seconds		
		Training Subset (Million)	Validation Subset (Million)	Test Subset (Million)
01	Microwave	18.22	5.43	6.76
02	Washing Machine	43.47	5	6

The following performance metrics are used to match the reproducible results with the claimed results. These performance metrics are chosen since they are used in the original article of the algorithm. Therefore, it is easier to compare the results.

- **Mean Absolute Error (MAE):** It evaluates the absolute difference between the prediction \hat{x}_t and ground truth x_t at each instance and calculates the mean value.

$$MAE = \frac{1}{T} \sum_{t=1}^T |\hat{x}_t - x_t| \quad \text{Equation V.1}$$

- **Signal aggregate Error (SAE):** It indicates the relative error of the total energy. 'r' denotes the total consumption of the appliance and \hat{r} the predicted total energy.

$$SAE = \frac{|\hat{r} - r|}{r} \quad \text{Equation V.2}$$

- **Energy per day (EpD):** This metric indicates the absolute error of the predicted energy used in a day, which is typically useful when the household users are interested in the total energy consumed in a period. Where $e = \sum_t x_t$ denotes the energy consumed in a day period and D is the total number of days.

$$EpD = \frac{1}{D} \sum_{n=1}^D |\hat{e} - e| \quad \text{Equation V.3}$$

- **Normalized disaggregation error (NDE):** This metric measures the normalized error of the squared difference between the prediction and the ground truth of the appliances.

$$NDE = \frac{\sum_{i,t} (x_{it} - \hat{x}_{it})^2}{\sum_{i,t} x_{it}^2} \quad \text{Equation V.4}$$

V.6.a Training and testing on same dataset (REFIT)

REFIT dataset is sampled at 8 seconds. It consists of a total of 43 million data points in the training subset of washing machine. For processing 43 million data points, it requires a compatible hardware solution in memory and computing capacity. That's why 2 approaches are compared i.e. a complete dataset (of 43 million data points) processed on a remote server with parallelization capacity on GPU and a reduced subset for training on personal computer. The following results are observed regarding the time of training without and with GPU installed on the server. NVIDIA RTX6000 GPU is installed on the server machine. The results show that with a GPU installed on a computer machine, the time taken to train a model considerably reduce to a greater extent.

Table V.5 The comparison of model creation with and without GPU

Sr. No.	Parameter	Without GPU	With GPU
01	Appliance	Washing Machine	Washing Machine
02	Length of training dataset used	5 (out of 43) million	43 (out of 43) million
03	Length of validation dataset used	5 (out of 5) million	5 (out of 5) million
04	Epochs	5	10 (early stopping at 5 th epoch)
05	Time taken	Around 13 hours	Around 6 hours

Here the results of washing machine (WM) and microwave oven (MO) are presented. The results of washing machine are graphically presented in Figure V.6 below. The purpose of energy disaggregation (as stated earlier) is to disaggregate the smart meter measured load curve into the consumption of respective appliances. Here, we reproduce the results to test that the disaggregated load curve of the appliance recognizes well the state of the appliance by comparing it to the ground truth value of the appliance. Not much difference is spotted for MAE and NDE, however difference in SAE and EpD is observed. Overall, it can be observed that the results are close. It demonstrates that the results are reproduced as they were claimed in (D’Incecco et al., 2020).

Table V.6 Trained and tested on REFIT dataset at sample period of 8 seconds

Appliance	Results of (D’Incecco et al., 2020)				Reproduced results			
	MAE	SAE	EpD (Wh)	NDE	MAE	SAE	EpD (Wh)	NDE
Microwave Oven	12.66	0.17	95.78	0.71	9.5	0.007	105.07	0.71
Washing Machine	16.85	2.61	319.11	0.54	32.09	0.16	418.16	0.58

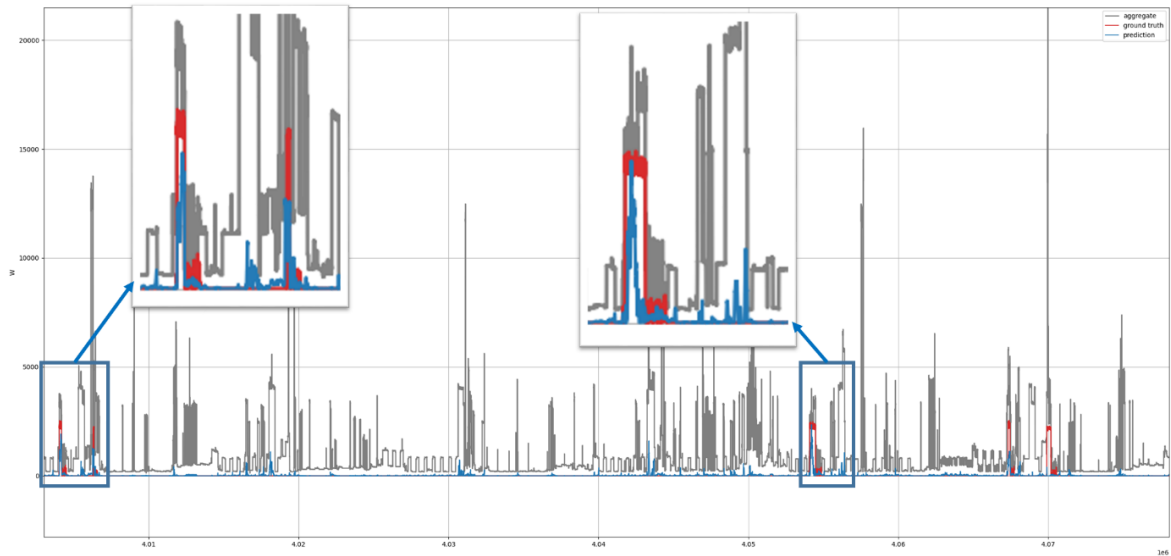


Figure V.6 Energy disaggregation of washing machine via model trained on REFIT dataset and tested on REFIT house no. 8. Grey curve represents aggregated consumption, red curve represents the actual consumption of washing machine and blue curve represents the predic

V.6.b Training on REFIT and testing on REDD dataset: transfer learning

The *REDD* dataset is sampled at 1 seconds for smart meter, whereas 3 seconds for appliances. To keep homogeneity with *REFIT*, the *REDD* dataset is resampled to 8 seconds. It is significant to look at the statistical values of appliances in Table V.7 before reproducing results. Since the same appliance in different datasets does not belongs to same brand and same model, therefore their energy consumption pattern is different. Besides, the appliances are not operated all the time, therefore they have different time of utilization. By looking at Table V.7, it can be observed that the appliances in REDD dataset have low utilization time (i.e. ON time), yet high mean and standard deviation value as compared to the appliances in REFIT dataset. It is illustrated in Figure V.7 as well, in which the peak consumption of washing machine for REDD dataset is higher than that of REFIT dataset. Besides, there is also impact of other appliances in the aggregated consumption. Therefore, these constraints bring difficulty in the transfer learning of energy disaggregation.

Table V.7 The statistical characteristics of appliances in REFIT and REDD dataset

Appliance	On (%)	OFF (%)	Mean (W)	Standard Deviation (W)
REFIT dataset (sampled at 8 seconds)				
Microwave Oven	6.6	93.4	9	92
Washing Machine	10.94	89.06	147	478
REDD dataset (sampled at 8 seconds)				
Microwave Oven	2.03	97.97	358	613
Washing Machine	0.69	99.31	2282	743

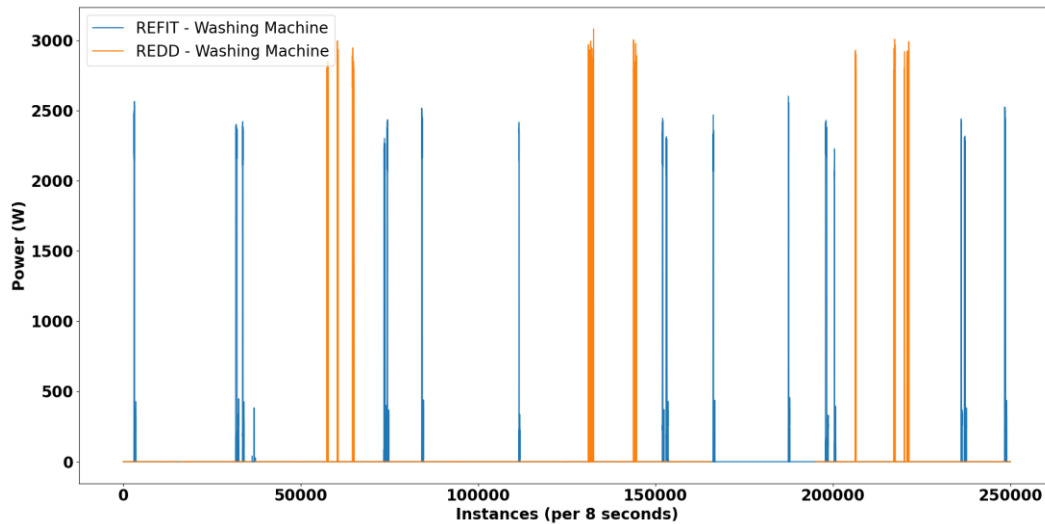


Figure V.7 The comparison of peak energy consumption of washing machine in REFIT dataset with REDD dataset

The models of washing machine and microwave that are previously trained on REFIT dataset is used to disaggregate data of REDD dataset. The reproduced results are found closer to the claimed results (D’Incecco et al., 2020) in both cases. The result of washing machine is presented in Figure V.8.

Table V.8 Trained on REFIT and tested on REDD dataset at sample period of 8 seconds

Appliance	Results of (D’Incecco et al., 2020)				Reproduced results			
	MAE	SAE	EpD (Wh)	NDE	MAE	SAE	EpD (Wh)	NDE
Microwave Oven	23.10	0.36	208	0.71	23.81	0.30	221.03	0.76
Washing Machine	36.83	0.74	750.85	0.91	35.24	0.87	778.04	0.96

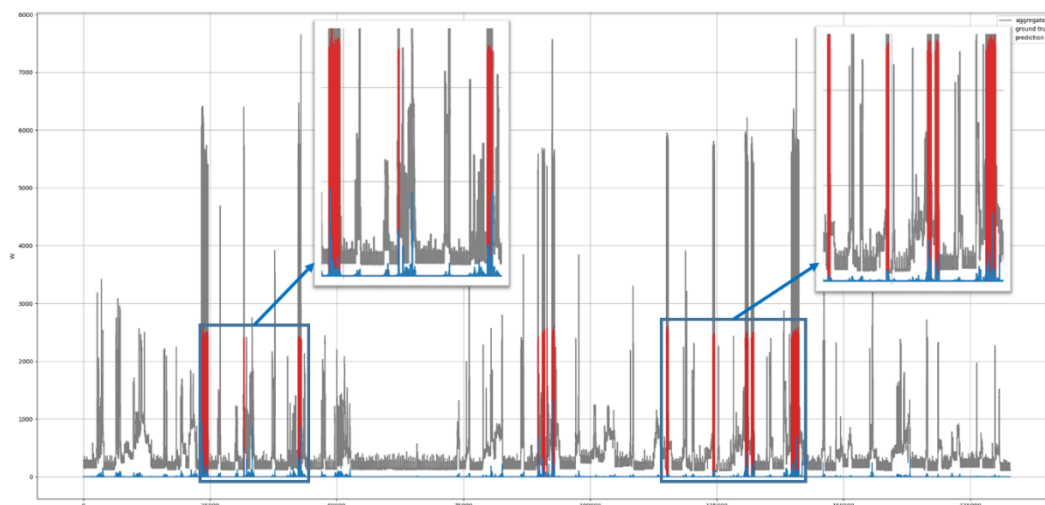


Figure V.8 Energy disaggregation of washing machine via model trained on REFIT dataset and tested on RED house no. 3. Grey curve represents aggregated consumption, red curve represents the actual consumption of washing machine and blue curve represents the predicted curve of washing machine

V.7 NILM on IRISE dataset

As the reproducibility of results using *transfer-NILM* algorithm is verified, it is now significant to create disaggregation models for the appliances that are discussed in the section V.5.a. These are the appliances that are present in the EXPESIGNO dataset and are either found in the IRISE dataset or certain appliances in IRISE dataset are found closer to the appliances in EXPESIGNO. It should be noted that *transfer-NILM* algorithm makes unique prediction model for each appliance. Before training models on *IRISE* dataset, the data is checked for the presence of outliers. Certain outliers are found in the data which represents the abnormal energy consumption of appliance at a certain instance. The outliers might be caused owing to the mal communication of sensor. These outliers are removed by the technique mentioned in the section below.

V.7.a Removing outliers from IRISE dataset

For an appliance X, these outliers are removed using the following technique.

- The energy consumption of an appliance X is sorted in ascending order. The highest consumption being 'X_n'.
- A scatter plot is created. The ordinate and abscissa represent the energy consumption of appliance X. Figure V.9 illustrates this scatter plot.
- Starting from the highest value 'X_n', if the following equation holds,

$$X_n - X_{n-1} > \frac{\sum_{i=1}^5 X_{n-i} - X_{n-i-1}}{5} \quad \text{Equation V.5}$$

- i.e. the distance between two highest consumption values is greater than the average distances of subsequent 5 values, than either;
 - o Replace the outlier value with zero if the preceding and succeeding values in the unsorted vector are zeros.
 - o Replace the outlier value with the average of preceding and succeeding values in the unsorted vector, if both preceding and succeeding values are non-zero, otherwise replace the outlier with zero.

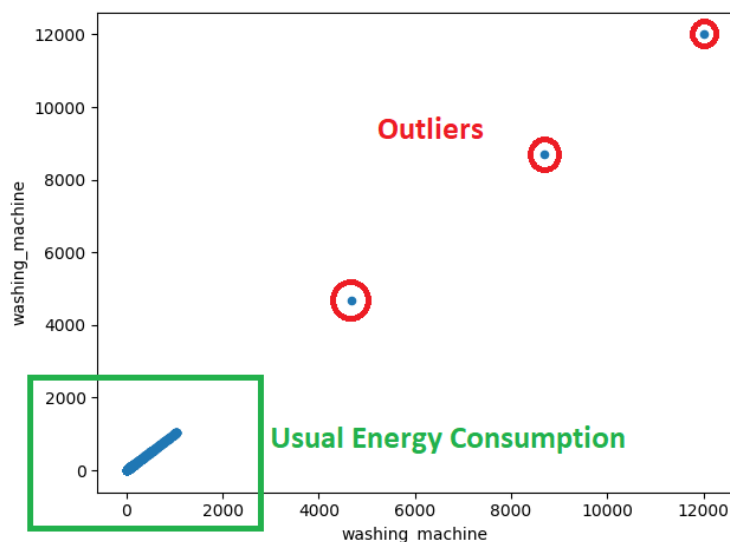


Figure V.9 Detection of outliers in washing machine of a household in IRISE dataset

V.7.b Predictive model creation based on IRISE dataset

To train a predictive model of an appliance for energy disaggregation, the energy consumption data of respective appliance and aggregated data is collected from ‘n’ households in IRISE dataset. where ‘n’ represents the total number of houses in which the data of respective appliance is found. The data is split for training, validation and testing of model. The data of ‘n-2’ household is kept for training; one household is kept for validation and one household is kept for test. The open-source algorithm with sequence to point neural network architecture (D’Incecco et al., 2020) as discussed in section V.6 is used to train the models. The reproducibility of the claimed result of this algorithm is already verified in the same section. The same procedure is adapted for IRISE dataset as it was earlier used for the energy disaggregation of REFIT and REDD dataset.

The models are trained for 100 epochs with early stopping. An epoch represents one training cycle of dataset, whereas by default the same model should be trained 100 times. However, with a threshold criterion fixed for early stopping, the algorithm checks if the decrease in validation loss between an epoch ‘n’ and ‘n-1’ is greater than the threshold value for three consecutive epoch runs. If the condition does not fulfill, then the training is stopped earlier and the model with the best validation score is saved. It is observed that the model training early stopped within the 5th and 10th epoch.

The split of households with respect to each appliance (discussed in section V.5.a) is given in Annex E. The training and validation loss is also given Annex E. whereas Table V.9 represents the length of data subsets (in number of rows) for data sampled at 10 minutes as well as data sampled at 30 minutes. It is obvious that the values for data down sampled at 30 minutes is one third the values of data originally sampled at 10 minutes.

Table V.9 The length of data subsets for training the predictive model of each appliance

Sr. No.	Appliance	Data sampled at 10 minutes			Data sampled at 30 minutes		
		Training Subset (Million)	Validation Subset (Million)	Test Subset (Million)	Training Subset (Million)	Validation Subset (Million)	Test Subset (Million)
01	Hotplate	0.387	0.054	0.053	0.126	0.018	0.018
02	Electric Heating	1.083	0.054	0.054	0.361	0.018	0.018
03	Electric Oven	1.304	0.054	0.053	0.435	0.018	0.018
04	Electric Cooker	1.458	0.054	0.053	0.488	0.018	0.018
05	Clothes drier	1.625	0.054	0.054	0.542	0.018	0.018
06	Dishwasher	2.485	0.054	0.053	0.832	0.018	0.018
07	Microwave	3.624	0.054	0.053	1.211	0.018	0.018
08	TV	4.157	0.054	0.053	1.391	0.018	0.018
09	Washing Machine	4.490	0.054	0.053	1.502	0.018	0.018

For the test of predictive models, the performance indicators are tabulated below in Table V.10. It is pertinent to discuss here that these performance indicators are different than the one used in section V.6. The purpose of former performance metrics is to verify the reproduction of the results by comparing it with the claimed results in the scientific publication of (D’Incecco et al., 2020). These performance metrics are given in the scientific publication and served to calculate the error in the prediction of energy consumption. Therefore, for the sake of convenience, the reproduced results are compared with the claimed results using the same performance metrics. In the case of detecting potential appliances for energy flexibility, the most significant element is to detect the state of appliance (i.e. ON/OFF) at each instant in the period of intervention. The prediction of extent of consumption is a matter of secondary importance in this regard. Therefore, our interest here is to detect the percentage of correctly predicted instances.

Table V.10 The performance indicators for testing the predicted models

Sr. No.	Performance Indicator	Remarks
01	True Positive (TP)	Actual state : ON Predicted state : ON
02	True Negative (TN)	Actual state : OFF Predicted state : OFF
03	False Positive (FP)	Actual state : OFF Predicted state : ON
04	False Negative (FN)	Actual state : ON Predicted state : OFF

05	F1 Score	<p>F1 score is a harmonic mean of precision⁴⁹ and recall⁵⁰. In the terms of the above 4 performance indicators, the F1 score is mathematically given as:</p> $F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = \frac{TP}{TP + \left\{ \frac{1}{2} \times (FP + FN) \right\}}$
-----------	-----------------	--

V.7.c The results of energy disaggregation in terms of performance indicators

Initially the median value (of measured consumption) is taken as a threshold for disaggregation results. These median values are given in Table V.2 and Table V.3 for resolution of 10 minutes and 30 minutes respectively. These median values are greater in value than 10W for the appliances except for electric cooker and TV. The threshold value brings sensitivity to results therefore it must be carefully selected by looking at the noise of the prediction. By looking at the results graphically, it is observed that the disaggregated curve does not have same amplitude as the measured curve of the appliance. Besides, it is also observed graphically that the disaggregated curve is found oscillatory within a specific amplitude around OX axis. Therefore, the threshold is set to 10W for all the appliances to suppress the effect of noise in the prediction on the results. The only exceptions in this regard are electric cooker and TV for which the threshold is set to 5W.

Table V.11 The results with respect to performance indicators for data sampled at 10 minutes

Sr. No.	Appliance	Threshold (W)	True Values		Performance Indicators				
			ON (%)	OFF (%)	True Positive (TP) (%)	True Negative (TN) (%)	False Positive (FP) (%)	False Negative (FN) (%)	F1 Score
01	Hotplate	10	2	98	1	93	5	1	0.17
02	Electric Heating	10	13	87	12	75	12	1	0.65
03	Electric Oven	10	3	97	3	90	7	0	0.45
04	Electric Cooker	5	4	96	2	83	13	2	0.21
05	Clothes drier	10	0.44	99.56	0.38	96.49	3.07	0.05	0.2
06	Dishwasher	10	2	98	2	94	4	0	0.5
07	Microwave	10	3	97	1	95	2	2	0.33
08	TV	5	27	73	24	53	20	3	0.68

⁴⁹ Precision is a ratio of true positives and the actual values. In terms of true positive (TP) and false positive (FP), the precision is mathematically given as $Precision = \frac{TP}{TP+FP}$

⁵⁰ Recall is a ratio of true positives and the predicted values. In terms of true positive (TP) and false negative (FN), the recall is mathematically given as $Recall = \frac{TP}{TP+FN}$

09	Washing Machine	10	3.62	96.38	1.73	90.88	5.72	1.67	0.32
-----------	-----------------	----	------	-------	------	-------	------	------	------

The results for 10 minutes sampling and 30 minutes sampling are presented in Table V.11 and Table V.12 respectively. For the sake of presentation, the values are rounded off, except for the highlighted appliances for whom it is found difficult to present results with rounded off numbers. Therefore, the results of these appliances are rounded off to 2 decimal digits.

Table V.12 The results with respect to performance indicators for data sampled at 30 minutes

Sr. No.	Appliance	Threshold (W)	True Values		Performance Indicators				
			ON (%)	OFF (%)	True Positive (TP) (%)	True Negative (TN) (%)	False Positive (FP) (%)	False Negative (FN) (%)	F1 Score
01	Hotplate	10	2.41	97.59	0.88	94.04	3.55	1.53	0.26
02	Electric Heating	10	18	82	17	58	24	1	0.58
03	Electric Oven	10	4	96	4	90	6	0	0.57
04	Electric Cooker	5	4	96	3	78	19	1	0.22
05	Clothes drier	10	0.54	99.46	0.45	94.36	5.10	0.08	0.15
06	Dishwasher	10	3	97	3	91	6	0	0.5
07	Microwave	10	6.68	93.32	1.25	91.63	1.68	5.43	0.26
08	TV	5	29	71	27	45	26	2	0.66
09	Washing Machine	10	4	96	2	88	8	2	0.29

It can be observed in Table V.11 that the ON (percentage) is very less as compared to OFF (percentage). The percentage value of false positive (FP) predictions is found greater (or equal to) the true positive (TP) predictions. It means that besides accurately predicting the ON instances, the model also predicts a large number of ON instances when the appliance was actually OFF. In addition to this, the percentage of false negatives (FN) is also (more or less) equal to the true positive (TP). False negative (FN) means that there were certain instances when the appliance was in ON state, however the model predicted that it is OFF. The same pattern is observed in the Table V.12 for the data sampled at 30 minutes. This makes a model less reliable as we expect that the number of true positives (TP) and true negatives (TN) should be much higher than false positives (FP) and false negatives (FN). Despite having considerable percentage of false positives (FP), the F1 score of certain appliances suggests them to be good models yet sensitive to the fixed threshold. The electric heating, electric oven, dishwasher and TV are notable in this regard.

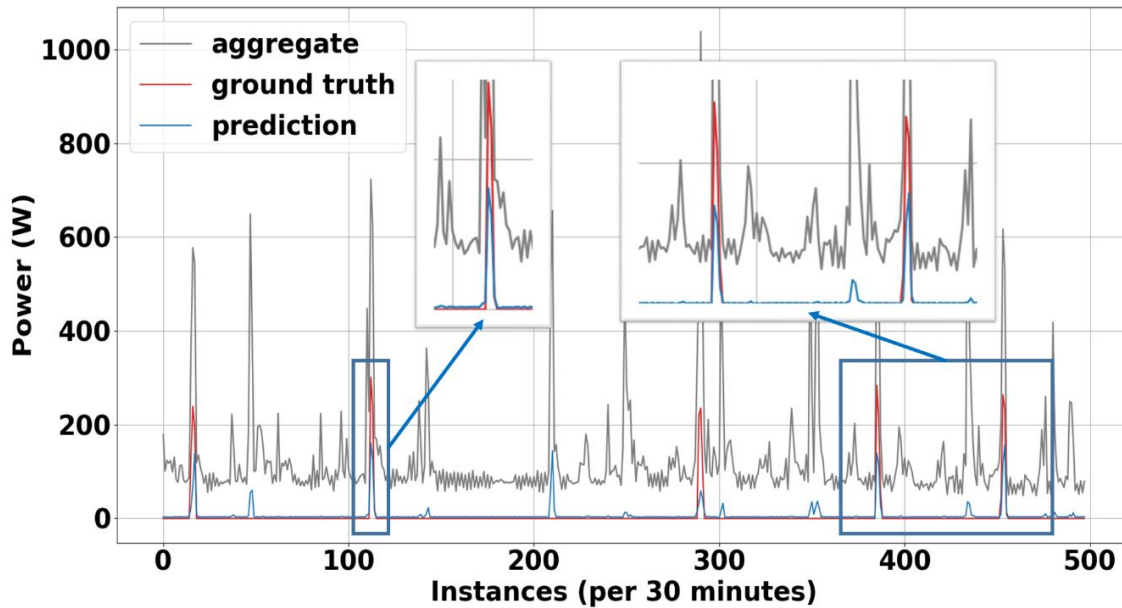


Figure V.10 Test result for dishwasher on IRISE house 997 at sampling rate of 30 minutes

V.8 Transferability on the EXPESIGNO dataset

The principal objective of creating predictive models of energy disaggregation is to use them on the energy consumption data of treated group during the period of intervention on alert days. In this way, the transfer learning of the predictive models could be analyzed. The created predictive models are based on deep learning. Due to its ability to identify shared factors present in the input, deep learning is deemed good for transfer learning. “Although preliminary experiments have shown much potential of deep learning in transfer learning, applying deep learning to this field is relatively new and much more needs to be done for improved performance” (Chen & Lin, 2014).

The proof of concept is presented and tested on a household of *IRISE* dataset for each appliance. The proof of concept works, though more reliable predictive models are sought. We need more reliable models to detect more true positives and less false positives so that transfer learning of predictive models from an open-source dataset (or ensemble of datasets) towards the EXPESIGNO dataset can be analyzed. For validation of predictive models, the performance of predictive models will be analyzed by using the intrusive load monitoring data of 4 households in EXPESIGNO dataset. It should be noted that instantaneous energy consumption data is recorded as appliance level for these 4 households as part of a side project of EXPESIGNO (known as DEEP-EXPESIGNO). Following the validation, the measured load curve during period of intervention will be disaggregated to identify potential appliances.

V.9 Conclusion

Energy disaggregation (particularly non-intrusive load monitoring) can serve as a helpful tool to analyze the impact of energy flexibility at appliance level. This does not only validate the potentiality of appliances for energy flexibility, but it also helps in verifying the impact of commitment for energy flexibility (in case commitment exists). In our experiment regarding nudge-based energy flexibility, the white appliances are mostly committed appliances.

Several interventions are made in our experiment. The principal objective is to analyze the impact of nudges at appliance level, however as a starting point, a proof of concept is presented. This proof-of-concept deals with the non-intrusive load monitoring through supervised learning using a dataset of recorded consumption at global level as well as at appliance level. The results of an open source NILM algorithm are reproduced to verify the claimed results before training the predictive models using IRISE dataset. The algorithm produces unique model for each appliance with respect to the given parameters.

The test results of IRISE dataset (by models trained on same dataset) depicts that the models can equally (or to a larger extent) predict false positives as compared to true positives. The F1 score of certain appliances are found well albeit these appliances having more false positives as compared to true positives. Though the models are not found reliable enough to implement on our experiment case, yet we found that NILM can serve as a promising tool to measure the potential of appliances during period of interventions. Here we have explored only a small part of the scientific objectives of such a study. There still needs many points to be addressed in order to ensure the quality of the results that we will be able to obtain on our EXPESIGNO experiment.

Chapter VI

Conclusion and Future Prospects

This chapter presents a general conclusion of this thesis in light of preceding chapters and present future prospects for research work.

VI.1 General Conclusion

Energy flexibility has been in practice worldwide in the form of demand response programs since many decades. These programs are implemented on the electric grid predominantly energized by conventional fossil fuel plants. However, the energy mix is in transition and it tends towards renewable energy sources through distributed generation. Where the integration of distributed renewable generation mitigates climate change to benefit environment, it also poses technical constraint of intermittent production (leading to possible network congestion). To keep energy balance in a future energy mix which is predominantly energized by renewable energy sources, it is inevitable to implement the different modes of energy flexibility.

Residential sector (being the highest consumer of electricity) remained the focal point of demand response. It still remains a potential sector to implement energy flexibility. Feedback is another important aspect of demand response. It can be either direct (i.e., in real time) or indirect (after the period of consumption). Direct energy flexibility is an incentive-based approach to switch loads at remote end. It is significant for industrial sector; however, it is not desirable for residential consumers since it is intrusive. Besides, there exist concerns among consumers regarding privacy and balanced use of direct load control. Contrary to this, the indirect energy flexibility offers price-based approaches that acts as extrinsic motivator for the consumers. However, there are certain factors that hinders the decision making of a residential consumer to implement price-based energy flexibility. It includes privacy, cognitive difficulty in understanding non-linear price system, information imperfection in the price-based signal, lack of motivation and uncertainty about potential consequences. **Chapter I** begins with a discussion on the changing landscape of energy mix and the constraints of traditional energy flexibility mechanisms, which leads us to the significance non-monetary energy flexibility mechanism and a brief introduction of implementation of its gist in this thesis.

Regarding the indirect energy flexibility, the cognitive burden of acting on price signal is on consumer which enables the consumer to make a controlled and calculated decision. The dual process theory of thinking suggests that the dynamic pricing instigates system 2 of thinking, which operates consciously, in control, with high effort and low-capacity process. It makes the consumer to think about the implementation and consequences of responding to the price signal with certain questions i.e., a) why to act? b) when to act? c) how to act? and d) how did it go? This makes it riskier to implement price based indirect energy flexibility. Besides, the distributed generation of renewable energy is considered as a common pool resource. In the anticipated energy mix predominant with renewable energy sources, the usual energy consumption will cause scarcity of energy,

either through intermittency of renewable energy or high demand at a certain time. Therefore, it is inevitable that the individuals should implement flexibility in energy consumption to avoid future dilemma. The practice of energy flexibility in present will eventually become a habit in future.

Chapter II begins with a discussion on the socio-cognitive constraints of human being. As a complement to the traditional demand response mechanisms, behavioral insights based energy flexibility mechanism is then presented with the introduction of “Nudge” (and eventually “Green Nudge”) as a tool. Nudge provides indirect suggestions to influence the decision making, whereas green nudge is an extension of nudge technique which is attributed to environmental cause. A review of nudges used in the field of energy worldwide is also presented. Most of these nudges used either a self-comparative or social comparative norm. In contrast to this, the architecture of a field experiment in residential sector using green nudges is presented. This non-monetary field experiment is performed with 175 French households (within the constraints of GDPR); categorized in control group and treated group. Two types of interventions are conceived i.e. orange alerts for load curtailment and green alerts for load shifting. The nudge signals (assisting with the curious questions) in the form of SMS are sent to the treated group for both types of interventions. The orange alerts are based on the forecasted grid congestion whereas the green alert requires prediction of load curve of residential sector, which is presented in the later section of this chapter.

Chapter III begins with the general statistics of alerts and a comparison of socio-technical variables of both groups. Both groups are found identical in all the statistical variables except the *year of construction*. However, this statistical variable does not have any concern with our experiment. Following this, the distribution of energy consumption of both groups is compared for normal days. The normal energy consumption of both groups is found similar by the visualization of distribution curve and statistical test of significant difference. The last section of this chapter deals with the meta-analysis of impact of nudge signals. It is observed that the distribution curve of treated group is shifted towards right for green alerts when superposed on the distribution curve of control group. In the case of orange alerts, the distribution curve of treated group is shifted towards left as compared to control group. It fairly gives a notion that the nudge signal has an impact on the energy consumption during the period of intervention on alert days.

The statistical test concludes that the distribution of consumption of treated group for all green alerts is greater than the distribution of control group, yet the result of individual green alert does not support this. On contrary, the result of individual orange alert supports the aggregated result that the distribution of consumption of treated group

for all green alerts is less than the distribution of control group. The variation in the energy consumption is quantified for both types of alerts. On aggregated level, the treated group implemented load shifting of energy flexibility by 11.17% in response to green alerts, whereas the group implemented load curtailment by 18.21% in response to orange alerts. The result regarding load curtailment of 18.21% sounds better compared than the study of moral suasion which yielded 8% (Ito et al., 2018), whereas it is closer to the results of a study of similar nature which yielded load curtailment in the range of 12-20% (Jorgensen et al., 2021). It is worth mentioning that the subjects of our experiment are environment conscious, therefore, the results of green nudges could be different for a bigger sample of subjects having different environment related ideologies.

Chapter IV presents a tool that can serve to provide feedback as well as to detect the effort of energy flexibility for each intervention. This tool is named as reference load curve. The original purpose of reference load curve is to provide the treated group graphical indirect feedback of each intervention and to measure its impact on the energy flexibility. However, the low number of views by the treated group is insufficient to measure this impact. Yet to present the feedback, the question raised as how to formulate reference load curve. In this regard, a classical method, two naïve methods and two machine learning methods are analyzed by energy difference indicator and R^2 score. According to the indicators, a classical reference load curve based on kernel density estimation is found suitable for measuring load shifting while a machine learning based reference load curve is found suitable for measuring load curtailment. In our experiment, we selected a naïve method to measure both types of energy flexibility since its results were best among the formerly studied methods and it is easy to explain to the treated group.

Unlike the comparison of control group and treated group, the reference load serves to measure the energy flexibility at individual level. To do so, it is imperative to first validate whether the reference load curve can potentially replace control group. An analysis is done in this regard, and it is observed that the reference load curve can replace control group for load curtailment. The later section presents the utility of reference load curve to detect the individual effort of energy flexibility and to categorize the consumers into groups i.e., a) who acted positive in response to nudge signal, b) who did not act positive in response to nudge signal and c) those for whom the impact is negligible or cannot be determined.

Chapter V demonstrates the frequency of committed appliances for both (green and orange) alert by the treated group. It serves to identify the appliances that were potential to implement energy flexibility during the period of intervention. A proof of concept is presented using *NILM* energy disaggregation. The idea is to use *NILM* for disaggregating the measured load curve of our experiment to identify the potentiality of an

appliance during the period of intervention. Yet many analyses are required to explore in order to ensure the quality of the results that we will be able to obtain on our experiment.

In a nutshell, the green nudge based non-monetary energy flexibility mechanism is found effective for energy flexibility of load shifting as well as load curtailment. The designed mechanism answers the curious questions of residential consumer upon receiving an energy flexibility alert. This mechanism complements the existing mechanisms of direct load control and dynamic pricing. The results for downward flexibility are found more concrete as compared to the results of load shifting. A reference load curve can replace the control group to detect the individual effort of energy flexibility. To identify the potentiality of an appliance, the NILM energy disaggregation can be used, however it needs to be explored in order to ensure the quality of results on our experiment.

VI.2 Future Prospects

This thesis offers several future prospects in multiple axes. The axes can be categorized based on use of other nudges in behavioral science, architecture development, reference load curve and data analysis in depth. We discuss each of these below.

VI.2.a Use of other nudges of behavioral science for energy flexibility

In our study, we used a set of green nudges for motivating people to implement energy flexibility. It includes information (of forecasted grid condition), default commitment, reminder and feedback. However, several other nudges should also be explored to implement energy flexibility. For instance, the intuitive metrics can be used to express information and/or the consequences of energy flexibility can be reframed in terms people care about (Yoeli et al., 2017).

In addition to this, the nudge of personalized target can be explored. These nudges are designed for everyone separately. Each individual subject could be communicated about the impact of switching an appliance present in his/her household. The impact of feedback as a nudge is an interesting subject to explore. We have seen in our experiment that the subjects are seldom interested in looking at their feedback. For instance, anchoring could be used during the information sessions in which the subjects should be informed about the importance of feedback. The feedback could be intertwined with few Yes/No questions to be responded by the subject. This in return also give data about the extent of easiness of implementing energy flexibility for each subject which could then be analyzed statistically.

VI.2.b Architecture development for experiment

There are certain developments that can be made as future prospect in the architecture. Some of these are outlined as follows.

- The orange alerts in our experiment are based on forecasted network congestion ($PP1/PP2$) on French national grid. Like *EJP* which is notified with respect to the part of the France (north, south, east or west), the forecasted network congestion can be further disintegrated at department level (in France). For local energy communities, it is essential to be more realistic and problem-oriented that the forecasted network congestion of its mini-grid should be used.
- Regarding the green alerts, a threshold of nebulosity is used as one of the criteria. It represents the solar energy, whereas criterion should also be explored for other types of renewable energy e.g., threshold for wind speed etc.
- Auto-consumption could be taken into account while sending nudges.
- In the architecture of our experiment, we used the modelled residential load curve (of national level) to evaluate a criterion of green alerts. This load curve is periodic within the constraints of its coefficient values. However, it would be more significant to model the residential load curve at experiment level. It is possible that this modelled load curve could not be as periodic as the national residential load curve.
- In the architecture of our experiment, we only took into account the smoothed national temperature for the prediction of modelled residential load curve. The historical data of pseudo-radiation (in percentage) is also given at national level. The impact of this variable could be explored for the prediction of residential load curve. Besides, we used random forest for prediction (with good results), however other techniques should also be explored, for instance deep learning (though it needs a large amount of data).

VI.2.c Reference Load curve

Reference load curve is a significant tool that has been explored for the sake of this experiment at a greater extent. However, according to our information, the work in this regard is preliminary and offers a lot of prospects. Some of these are outlined here.

- Since reference load curve is a prediction, therefore methods other than *Random Forest* should also be explored. For instance, deep learning with a sufficient amount of data might give better results than machine learning models.
- Only two performance indicators are explored in this thesis. Other performance indicators should also be explored for better presentation of suitability of a method to formulate reference load curve.

- With our analysis, we cannot show that the reference load curve can replace control group to detect and measure load shifting. However, this area of research deserves a lot of improvement and several methods of formulating reference load curve should be analyzed in this regard in future.
- Based on a sensitivity analysis of the thresholds for detecting positively acting subjects, a most suitable threshold value should be selected.

VI.2.d In-depth data analysis of potential appliances for energy flexibility

In Chapter V, a proof of concept regarding in-depth analysis of potential appliances for energy flexibility is presented. It offers several future prospects, of which some are mentioned below.

The impact of resolution on NILM should be evaluated in depth. It should be evaluated in depth that whether the models trained on higher resolution are equally (or near equally) effective to disaggregate the aggregated load curve of lower resolution.

The impact of training NILM models on an ensemble of open-source datasets (for a fixed value of resolution) should be evaluated. With respect to the performance indicators, the impact of training on individual datasets should be compared with the training on ensemble of datasets.

Finally, the models should be tested on the EXPESIGNO dataset to identify and quantify the potentiality of appliances during the period of intervention. A naïve method of quantification in terms of percentage could be using the formula below.

$$\text{Potentiality of appliance} = \frac{\text{Number of times the curtailment (or shifting) of appliance is identified during period of alert}}{\text{Total number of subjects who committed this appliance}} * 100$$

ANNEX A

A.1 Web interface of Etude Elec Site

The screenshot shows the login page of the EtudElec website. At the top left is the GAEL logo, and at the top right is the UGA (Univ. Grenoble Alpes) logo. The main heading is "EtudElec". Below it, a message states "Le site de l'étude menée par le laboratoire GAEL". A notification says "Votre session a expiré, veuillez vous reconnecter". The login form includes a "Numéro de téléphone" field with the value "0123456789" and a "Code sujet" field with the value "123". A "CONNEXION" button is positioned below the fields. At the bottom, there are links for "Conditions d'utilisation" and "Fiche d'informations". At the very bottom of the page, logos for ANR, La Région Auvergne-Rhône-Alpes, and the Agence Nationale de la Recherche are displayed.

A.2 Homepage of personal account

The screenshot shows the homepage of the EtudElec website for a personal account. The browser address bar shows "https://etudelec.univ-grenoble-alpes.fr". The page is divided into a left sidebar and a main content area. The sidebar contains a navigation menu with the following items: "Accueil", "Livret personnel sur vos consommations 2019/2020", "Questionnaire Juin 2020", "Vos Équipements", "Vos Engagements" (with sub-items "Période Verte" and "Période Orange"), "Retour d'informations" (with sub-item "Courbes de consommation"), "Guide de l'étude", and "Contact". The main content area features the "EtudElec" heading and a paragraph explaining the study: "Cette étude est menée par le Laboratoire d'Économie Appliquée de Grenoble (GAEL) et le laboratoire de Génie Electrique de Grenoble (G2ELab), dans le cadre d'un programme de recherche publique financé par la région Auvergne Rhône-Alpes et l'Agence Nationale de la Recherche." This is followed by another paragraph: "Cette étude porte sur les enjeux de la consommation électrique des ménages dans ce contexte de développement des énergies renouvelables (qui représentent actuellement que 18% de la production d'électricité en France). Plus particulièrement nous sommes intéressés par les conditions d'une plus grande flexibilité au niveau de la consommation des ménages. Cela signifie que les consommations électriques quotidiennes devraient être :". A bulleted list follows: "• Réduites : lorsque la production par les ENR est insuffisante." and "• Favorisées : lorsque la production par les ENR est importante." Below this, another paragraph states: "Cette flexibilité permettra de consommer des énergies vertes, de participer à la transition énergétique et de lutter contre le réchauffement climatique en limitant le recours à des centrales thermiques." The final sentence reads: "Notre objectif est de mesurer les possibilités d'une plus grande flexibilité chez les ménages."

A.3 Default commitments for green alert

Vos engagements - Période Veri X

https://etudelec.univ-grenoble-alpes.fr/engagements/periode-verte

Accueil
Livret personnel sur vos consommations 2019/2020
Questionnaire Juin 2020
Vos Équipements
Vos Engagements
Période Verte
Période Orange
Retour d'informations
Courbes de consommation
Guide de l'étude
Contact

Vos Engagements - Période Verte

ENTRE 12H ET 15H, EN SEMAINE, VOUS VOUS ENGAGEZ À :

- Utiliser le four
- Utiliser le micro-ondes
- Utiliser les plaques électriques
- Utiliser le lave-vaisselle
- Utiliser des appareils électriques de cuisine chauffant (friteuse, robot cuiseur, bouilloire...)
- Utiliser des appareils électriques de cuisine non-chauffant (mixeur, batteur...)
- Utiliser le lave-linge
- Utiliser le sèche-linge
- Repasser le linge
- Utiliser des appareils électriques pour les soins capillaires
- Mettre le thermostat du chauffage électrique principal à plus de 20°C
- Utiliser le radiateur d'appoint
- Utiliser des appareils électroniques sur secteur (ordinateur fixe, imprimante...)
- Utiliser des appareils audiovisuels sur secteur (TV, hifi, consoles de jeu fixes...)
- Recharger les batteries des appareils audiovisuels et électroniques (tablettes, smartphone)
- Recharger votre voiture électrique
- Recharger votre scooter électrique
- Recharger vos véhicules de mobilité électrique (vélo, trottinette...)

REENREGISTRER

A.4 Default Commitment of Orange Alert

Vos engagements - Période Ora X

https://etudelec.univ-grenoble-alpes.fr/engagements/periode-orange

Accueil
Livret personnel sur vos consommations 2019/2020
Questionnaire Juin 2020
Vos Équipements
Vos Engagements
Période Verte
Période Orange
Retour d'informations
Courbes de consommation
Guide de l'étude
Contact

Vos Engagements - Période Orange

ENTRE 18H ET 20H, EN SEMAINE, VOUS VOUS ENGAGEZ À NE PAS :

- Utiliser le four
- Utiliser le micro-ondes
- Utiliser les plaques électriques
- Utiliser le lave-vaisselle
- Utiliser des appareils électriques de cuisine chauffant (friteuse, robot cuiseur, bouilloire...)
- Utiliser des appareils électriques de cuisine non-chauffant (mixeur, batteur...)
- Utiliser le lave-linge
- Utiliser le sèche-linge
- Repasser le linge
- Utiliser des appareils électriques pour les soins capillaires
- Mettre le thermostat du chauffage électrique principal à plus de 20°C
- Utiliser le radiateur d'appoint
- Utiliser des appareils électroniques sur secteur (ordinateur fixe, imprimante...)
- Utiliser des appareils audiovisuels sur secteur (TV, hifi, consoles de jeu fixes...)
- Recharger les batteries des appareils audiovisuels et électroniques (tablettes, smartphone)
- Recharger votre voiture électrique
- Recharger votre scooter électrique
- Recharger vos véhicules de mobilité électrique (vélo, trottinette...)

REENREGISTRER

A.5 Feedback of the alert



ANNEX B

B.1 Household Characteristics

B.1.i Owner/Tenant

This variable concerns the status of households regarding their housing (owner or tenant).

Table B.1: Distribution of household status according to group

Status	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Owner	62	78.5	74	77.9
Tenant	17	21.5	21	22.1
Total	79	100	95	100

B.1.ii House/Apartment

This variable concerns the type of housing of the household (apartment or house).

Table B.2: Distribution of the type of housing of households according to the groups

Status	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
House	36	54.4	50	52.6
Apartment	43	45.6	45	47.4
Total	79	100	95	100

B.1.iii Surface of household

This variable represents the living area of the household of both groups. In order to use this variable, it was necessary to make groupings into classes. The surface areas are grouped by classes of 50, whereas the last grouping is between from 150 to 400.

Table B.3: Distribution of liveable area by group

Surface Area	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
[0-50[5	6.3	9	9.4
[50-100[45	57.0	51	53.7
[100-150[24	30.4	24	25.3
[150-400]	5	6.3	11	11.6
Total	79	100	95	100

B.1.iv Year of construction

This variable provides information on the period of construction of the household.

Table B.4: Distribution of the period of construction of the household according to the groups

Year of Construction	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Avant 1949	9	11.4	7	7.4
1949-1961	2	2.5	13	13.7
1962-1974	31	39.2	23	24.2
1975-1989	14	17.7	16	16.8
1990-2012	18	22.8	25	26.3
Après 2012	5	6.3	11	11.6
Total	79	100	95	100

B.1.v Number of persons in the household

This variable concerns the number of persons per household.

Table B.5: Distribution of the number of people in the household according to the groups

Number of Persons	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
1	8	10.10	13	13.70
2	20	25.30	25	26.30
3	24	30.40	22	23.20
4	16	20.30	22	23.20
5 or more	11	14.00	13	13.70
Total	79	100	95	100

B.1.vi Presence in the household at least 3 days per week (during weekdays)

This variable corresponds to the household response for the question; "In your dwelling is there a person during the day at least 3 days a week (excluding weekends)?"

Table B.6: Distribution of the presence of a person in the dwelling at least 3 days a week

Presence 3 days per week	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	36	45.6	49	51.6
No	43	54.4	46	48.4
Total	79	100	95	100

B.1.vii Types of electric heating

This variable provides information on the type of heating used by households.

Table B.7: Distribution of different types of heating according to group

Type of Heating	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Gas	53	67.1	64	67.4
Electricity	7	8.9	9	9.5
Liquid Fuel	4	5.1	5	5.3
Granular Wood	1	1.3	5	5.3
Inverter AC	0	0.0	2	2.1
Other	14	17.7	10	10.5
Total	79	100	95	100

B.2 Presence of appliances

B.2.i Electric water heater

This variable addresses whether or not the household has a water heater.

Table B.8: Distribution of water heater ownership by group

Presence of water heater	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	22	72.2	21	22.1
No	57	27.8	74	77.9
Total	79	100	95	100

B.2.ii Air Conditioning

This variable provides information on the presence of air conditioning in households.

Table B.9: Distribution of presence of air conditioning by group

Presence of air conditioner	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	5	6.3	8	8.4
No	74	93.7	87	91.6
Total	79	100	95	100

B.2.iii Heat Pump

This variable provides information on presence of heat pump in the household.

Table B.10: Distribution of presence of heat pump by group

Presence of heat pump	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	4	5.1	6	6.3
No	75	94.9	89	93.7
Total	79	100	95	100

B.2.iv PV panels

This variable allows us to see whether or not households have photovoltaic panels.

Table B.11: Distribution of presence of PV panel by group

Presence of PV panels	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	0	0	3	3.2
No	79	100	92	96.8
Total	79	100	95	100

B.2.v Double glazing of windows

It is interesting to know about the presence of double glazed windows in the dwelling of households.

Table B.12: Distribution of presence of double glazing by group

Presence of double glazed windows	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
Yes	66	83.5	84	88.4
No	5	6.3	8	8.4
Partially	8	10.1	3	3.4
Total	79	100	95	100

B.3 Number of appliances in household

B.3.i Electric Oven

This variable refers to the number of ovens present per household.

Table B.13: Distribution of number of ovens according to group

Number of electric ovens	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
0	2	2.5	2	2.1
1	70	88.6	85	89.5
2	7	8.9	8	8.4
Total	79	100	95	100

B.3.ii Dishwasher

This variable provides information on the number of dishwashers per household.

Table B.14: Distribution of number of dishwashers by group

Number of dishwashers	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
0	16	20.3	15	15.8
1	63	79.7	80	84.2
Total	79	100	95	100

B.3.iii Washing Machine

This variable shows the number of washing machines present in the homes.

Table B.15: Distribution of number of washing machines according to group

Number of washing machines	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
0	1	1.3	0	0
1	77	97.5	95	100
2	1	1.3	0	0
Total	79	100	95	100

B.3.iv Cloth Dryer

This variable provides information on the number of clothes dryers owned by households.

Table B.16: Distribution of number of clothes dryers by group

Number of cloth dryers	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
0	61	77.2	65	68.4
1	18	22.8	30	31.6
Total	79	100	95	100

B.3.v Electric Vehicle

This variable tells us whether or not households have an electric vehicle.

Table B.17: Distribution of number of electric vehicle by group

Number of cloth dryers	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
0	79	100	93	97.9
1	0	0	2	2.1
Total	79	100	95	100

B.4 Energy Consumption during weekdays

B.4.i Pressing Iron

This variable concerns with the hour of use of the iron during a weekday.

Table B.18: Distribution of iron use during a weekday

Time of use of pressing iron	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	10	22.7	10	16.9
9 AM to Noon	5	11.3	7	11.8
Noon to 3 PM	1	2.3	3	5.1
3 PM to 6 PM	3	6.8	7	11.8
6 PM to 8 PM	7	15.1	14	23.7
8 PM to 11 PM	15	34.1	17	28.8
11 PM to 6 AM	1	2.3	1	1.7
Don't know	2	4.5	0	0
Total	79	100	95	100

B.4.ii Electric Oven

This variable corresponds to the use of the oven on weekdays.

Table B.19: Distribution of electric oven use during weekday

Time of use of electric oven	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	4	5.3	4	3.4
9 AM to Noon	8	10.5	9	7.6
Noon to 3 PM	6	7.9	11	9.3
3 PM to 6 PM	3	3.9	7	5.9
6 PM to 8 PM	49	64.4	65	55.0
8 PM to 11 PM	6	7.9	21	17.7
11 PM to 6 AM	0	0	1	0.8
Don't know	0	0	0	0
Total	79	100	95	100

B.4.iii Dishwasher

This variable corresponds to dishwasher use on weekdays.

Table B.20: Distribution of dishwasher use on weekdays

Time of use of dishwasher	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	5	7.9	8	8.6
9 AM to Noon	5	7.9	9	9.7
Noon to 3 PM	5	7.9	9	9.7
3 PM to 6 PM	3	4.8	6	6.5
6 PM to 8 PM	6	9.5	5	5.4
8 PM to 11 PM	31	49.2	43	46.2
11 PM to 6 AM	8	12.7	13	14.0
Don't know	0	0	0	0
Total	79	100	95	100

B.4.iv Washing Machine

This variable refers to the use of the washing machine on weekdays.

Table B.21: Distribution of washing machine use on weekdays

Time of use of washing machine	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	5	6.7	6	5.5
9 AM to Noon	18	24.0	20	18.2
Noon to 3 PM	1	1.3	13	11.8
3 PM to 6 PM	15	20.0	16	14.5
6 PM to 8 PM	17	22.7	28	25.5

8 PM to 11 PM	14	18.7	16	14.5
11 PM to 6 AM	3	4.0	9	8.2
Don't know	2	2.7	2	1.8
Total	79	100	95	100

B.4.v Cloth Dryer

This variable refers to the use of the cloth dryer on weekdays.

Table B.22: Distribution of cloth dryer use on weekdays

Time of use of cloth dryer	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	0	0	2	8.0
9 AM to Noon	3	21.4	6	24.0
Noon to 3 PM	0	0	3	12.0
3 PM to 6 PM	1	7.1	0	0
6 PM to 8 PM	2	14.3	5	20.0
8 PM to 11 PM	5	35.7	8	32.0
11 PM to 6 AM	1	7.1	1	4.0
Don't know	2	14.3	0	0
Total	79	100	95	100

B.4.vi Electric hotplates

This variable refers to the use of the induction plates for the purpose of cooking on weekdays.

Table B.23: Distribution of induction plates use on weekdays

Time of use of hotplate	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	5	8.9	8	9.9
9 AM to Noon	8	14.3	10	12.3
Noon to 3 PM	10	17.9	13	16.0
3 PM to 6 PM	1	1.8	0	0
6 PM to 8 PM	27	48.2	37	45.7
8 PM to 11 PM	5	8.9	12	14.8
11 PM to 6 AM	0	0	1	1.2
Don't know	0	0	0	0
Total	79	100	95	100

B.4.vii Radiator

This variable refers to the use of the electric radiator for space heating on weekdays.

Table B.24: Distribution of electric radiator use on weekdays

Time of use of electric radiator	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	6	28.6	11	50.0
9 AM to Noon	1	4.8	0	0
Noon to 3 PM	0	0	0	0
3 PM to 6 PM	0	0	0	0
6 PM to 8 PM	12	57.1	9	40.9
8 PM to 11 PM	2	9.5	1	4.5
11 PM to 6 AM	0	0	0	0
Don't know	0	0	1	4.5
Total	79	100	95	100

B.5 Energy Consumption during the weekend

B.5.i Pressing Iron

This variable concerns with the hour of use of the iron during a weekend.

Table B.25: Distribution of iron use during a weekend

Time of use of pressing iron	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	1	2.6	2	3.4
9 AM to Noon	10	26.3	15	25.4
Noon to 3 PM	2	5.2	1	1.7
3 PM to 6 PM	5	13.1	15	25.4
6 PM to 8 PM	9	23.6	10	16.9
8 PM to 11 PM	6	15.7	13	22.0
11 PM to 6 AM	0	0.0	1	1.7
Don't know	5	13.1	2	3.4
Total	79	100	95	100

B.5.ii Electric Oven

This variable corresponds to the use of the oven on weekend.

Table B.26: Distribution of electric oven use during weekend

Time of use of electric oven	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	1	2.6	2	3.4
9 AM to Noon	10	26.3	15	25.4
Noon to 3 PM	2	5.3	1	1.7

3 PM to 6 PM	5	13.2	15	25.4
6 PM to 8 PM	9	23.7	10	16.9
8 PM to 11 PM	6	15.8	13	22.0
11 PM to 6 AM	0	0	1	1.7
Don't know	5	13.2	2	3.4
Total	79	100	95	100

B.5.iii Dishwasher

This variable corresponds to dishwasher use on weekends.

Table B.27: Distribution of dishwasher use on weekends

Time of use of dishwasher	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	1	1.3	4	4.2
9 AM to Noon	10	1.3	14	14.7
Noon to 3 PM	13	17.3	16	16.8
3 PM to 6 PM	12	16.0	8	8.4
6 PM to 8 PM	5	6.7	6	6.3
8 PM to 11 PM	24	32.0	31	32.6
11 PM to 6 AM	10	13.0	11	11.6
Don't know	0	0	5	5.3
Total	79	100	95	100

B.5.iv Washing Machine

This variable refers to the use of the washing machine on weekend.

Table B.28: Distribution of washing machine use on weekend

Time of use of washing machine	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	6	6.1	6	5.3
9 AM to Noon	34	34.7	49	43.0
Noon to 3 PM	16	16.3	14	12.3
3 PM to 6 PM	17	17.3	16	14.0
6 PM to 8 PM	9	9.2	11	9.6
8 PM to 11 PM	9	9.2	10	8.8
11 PM to 6 AM	3	3.1	6	5.3
Don't know	4	4.1	2	1.8
Total	79	100	95	100

B.5.v Cloth Dryer

This variable refers to the use of the cloth dryer on weekend.

Table B.29: Distribution of cloth dryer use on weekend

Time of use of cloth dryer	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	0	0	2	8.0
9 AM to Noon	3	16.0	6	24.0
Noon to 3 PM	2	11.0	5	20.0
3 PM to 6 PM	4	22.0	1	4.0
6 PM to 8 PM	3	16.0	4	16.0
8 PM to 11 PM	4	22.0	4	16.0
11 PM to 6 AM	0	0	2	8.0
Don't know	2	11.0	1	4.0
Total	79	100	95	100

B.5.vi Electric hotplates

This variable refers to the use of the induction plates on weekend.

Table B.30: Distribution of induction plates use on weekend

Time of use of hotplate	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	3	4.6	8	7.3
9 AM to Noon	11	16.9	21	19.1
Noon to 3 PM	18	27.7	26	23.6
3 PM to 6 PM	3	4.6	4	3.6
6 PM to 8 PM	24	36.9	37	33.6
8 PM to 11 PM	6	9.2	14	12.7
11 PM to 6 AM	0	0	0	0
Don't know	0	0	0	0
Total	79	100	95	100

B.5.vii Radiator

This variable refers to the use of the electric radiator for space heating on weekdays.

Table B.31: Distribution of electric radiator use on weekdays

Time of use of electric radiator	Control Group		Treated Group	
	Number of subjects	Percentage (%)	Number of subjects	Percentage (%)
6 AM to 9 AM	4	20.0	4	17.4
9 AM to Noon	2	10.0	7	30.4
Noon to 3 PM	0	0	0	0
3 PM to 6 PM	1	5.0	1	4.3
6 PM to 8 PM	11	55.0	8	34.8
8 PM to 11 PM	2	10.0	1	4.3
11 PM to 6 AM	0	0	0	0
Don't know	0	0	2	8.7
Total	79	100	95	100

ANNEX C

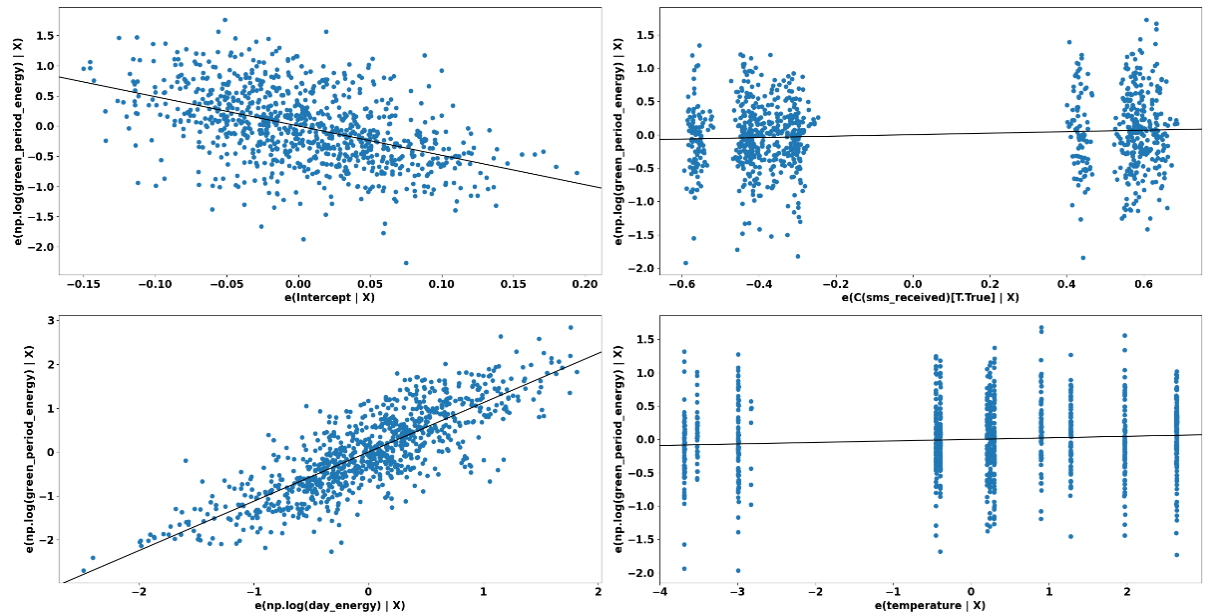
C.1 The results of OLS regression for all green alerts

OLS Regression Results

=====						
Dep. Variable:	np.log(green_period_energy)	R-squared:	0.681			
Model:	OLS	Adj. R-squared:	0.680			
Method:	Least Squares	F-statistic:	637.3			
Date:	Fri, 10 Dec 2021	Prob (F-statistic):	1.41e-221			
Time:	10:34:13	Log-Likelihood:	-718.09			
No. Observations:	899	AIC:	1444.			
Df Residuals:	895	BIC:	1463.			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----	-----	-----	-----	-----	-----	-----
Intercept	-4.8623	0.310	-15.695	0.000	-5.470	-4.254
C(sms_received)[T.True]	0.1117	0.037	3.025	0.003	0.039	0.184
np.log(day_energy)	1.1231	0.026	43.450	0.000	1.072	1.174
temperature	0.0231	0.009	2.565	0.010	0.005	0.041
=====						
Omnibus:	7.496	Durbin-Watson:	2.013			
Prob(Omnibus):	0.024	Jarque-Bera (JB):	9.533			
Skew:	-0.088	Prob(JB):	0.00851			
Kurtosis:	3.472	Cond. No.	400.			
=====						

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



C.2 The results of OLS regression for green alerts of session 1

OLS Regression Results

```

=====
Dep. Variable:      np.log(green_period_energy)      R-squared:                0.653
Model:              OLS                            Adj. R-squared:           0.651
Method:            Least Squares                   F-statistic:              381.6
Date:              Fri, 10 Dec 2021                 Prob (F-statistic):       2.53e-139
Time:              12:14:32                         Log-Likelihood:           -496.99
No. Observations: 612                               AIC:                      1002.
Df Residuals:      608                               BIC:                      1020.
Df Model:          3
Covariance Type:   nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-4.8115	0.389	-12.381	0.000	-5.575	-4.048
C(sms_received) [T.True]	0.1149	0.044	2.591	0.010	0.028	0.202
np.log(day_energy)	1.1113	0.033	33.520	0.000	1.046	1.176
temperature	0.0259	0.011	2.296	0.022	0.004	0.048

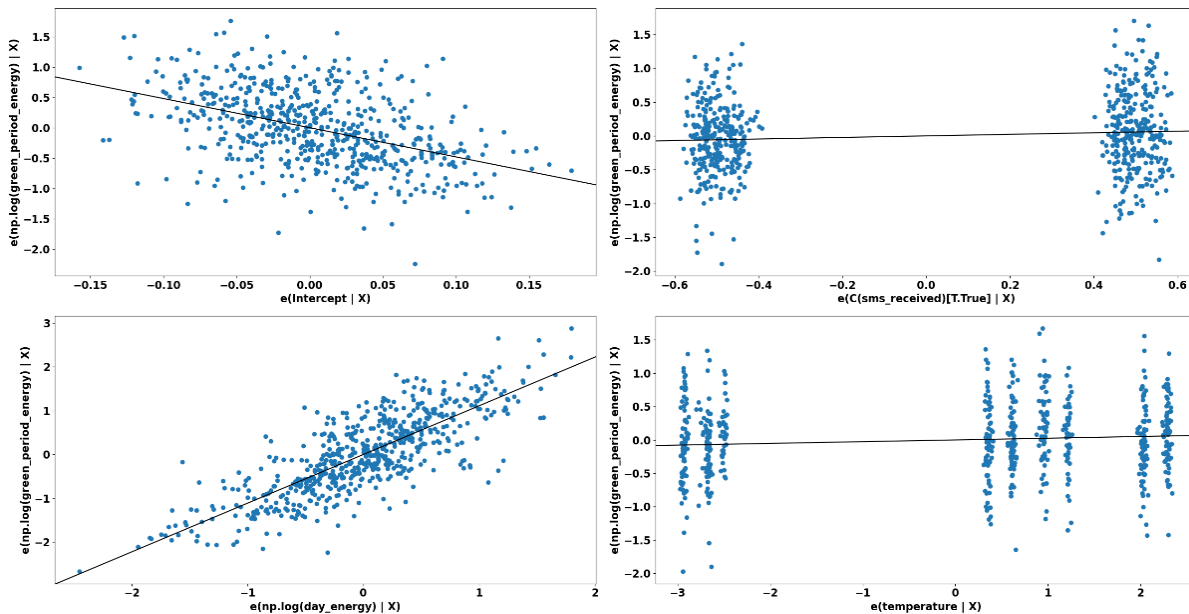
```

=====
Omnibus:          4.247      Durbin-Watson:           2.020
Prob(Omnibus):    0.120      Jarque-Bera (JB):       5.207
Skew:             -0.010     Prob(JB):                0.0740
Kurtosis:         3.451     Cond. No.                399.
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

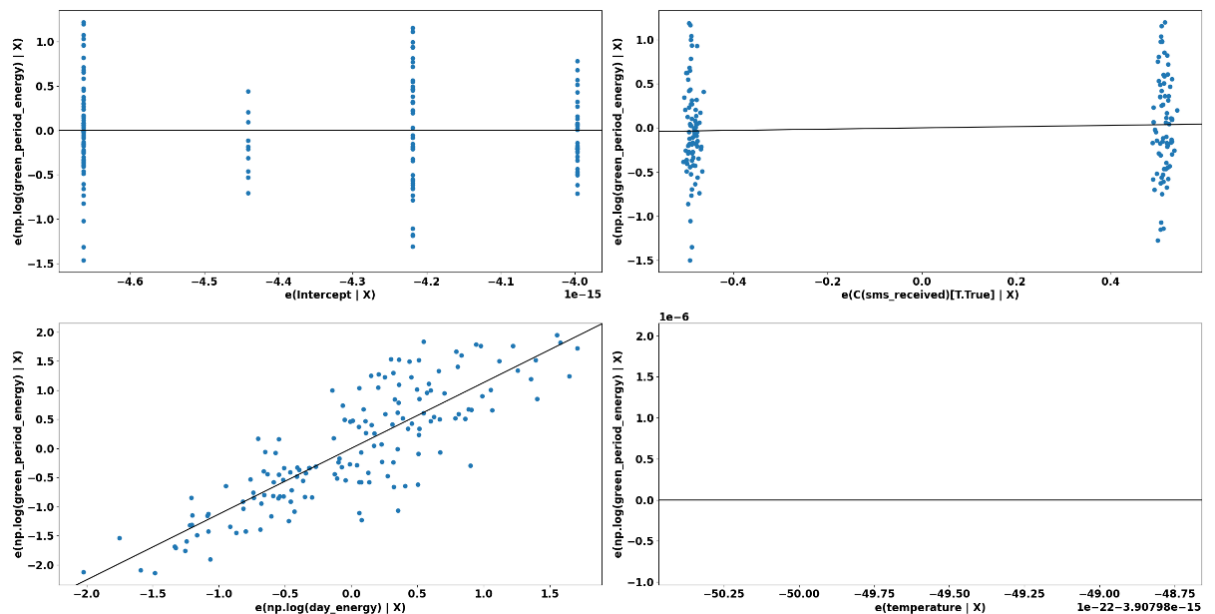


C.3 The results of OLS regression for green alerts of session 2

OLS Regression Results						
Dep. Variable:	np.log(green_period_energy)	R-squared:	0.707			
Model:	OLS	Adj. R-squared:	0.704			
Method:	Least Squares	F-statistic:	185.0			
Date:	Fri, 10 Dec 2021	Prob (F-statistic):	1.47e-41			
Time:	12:22:42	Log-Likelihood:	-125.87			
No. Observations:	156	AIC:	257.7			
Df Residuals:	153	BIC:	266.9			
Df Model:	2					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	-0.0101	0.001	-7.853	0.000	-0.013	-0.008
C(sms_received) [T.True]	0.0709	0.088	0.809	0.420	-0.102	0.244
np.log(day_energy)	1.1296	0.059	19.196	0.000	1.013	1.246
temperature	-0.2103	0.027	-7.853	0.000	-0.263	-0.157
Omnibus:	0.050	Durbin-Watson:	1.882			
Prob(Omnibus):	0.975	Jarque-Bera (JB):	0.015			
Skew:	0.020	Prob(JB):	0.993			
Kurtosis:	2.975	Cond. No.	8.61e+16			

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The smallest eigenvalue is 1.11e-29. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.



C.4 The results of OLS regression for all orange alerts

OLS Regression Results

```

=====
Dep. Variable:      np.log(orange_period_energy)  R-squared:                0.726
Model:              OLS                          Adj. R-squared:           0.726
Method:             Least Squares                F-statistic:              3035.
Date:               Fri, 10 Dec 2021             Prob (F-statistic):       0.00
Time:               10:40:31                     Log-Likelihood:           -2639.3
No. Observations:  3440                          AIC:                      5287.
Df Residuals:       3436                          BIC:                      5311.
Df Model:           3
Covariance Type:   nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-2.5978	0.102	-25.359	0.000	-2.799	-2.397
C(sms_received) [T.True]	-0.1821	0.018	-10.237	0.000	-0.217	-0.147
np.log(day_energy)	0.9840	0.010	94.256	0.000	0.964	1.004
temperature	-0.0073	0.002	-3.381	0.001	-0.012	-0.003

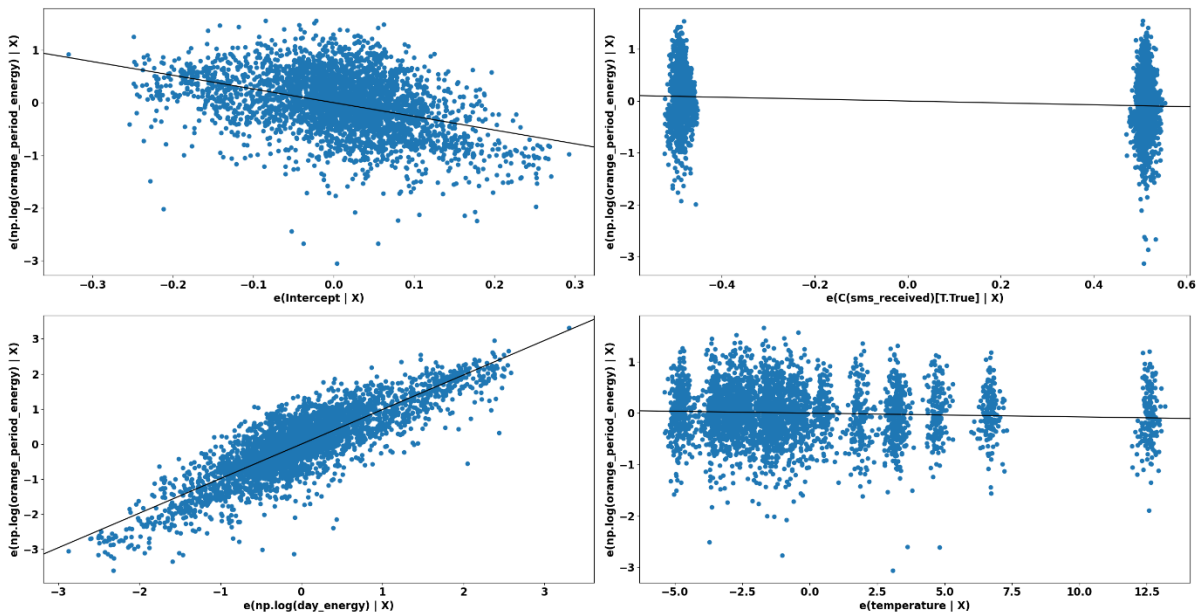
```

=====
Omnibus:           147.318  Durbin-Watson:           1.995
Prob(Omnibus):     0.000    Jarque-Bera (JB):        270.028
Skew:              -0.331   Prob(JB):                 2.31e-59
Kurtosis:          4.202    Cond. No.                  125.
=====

```

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

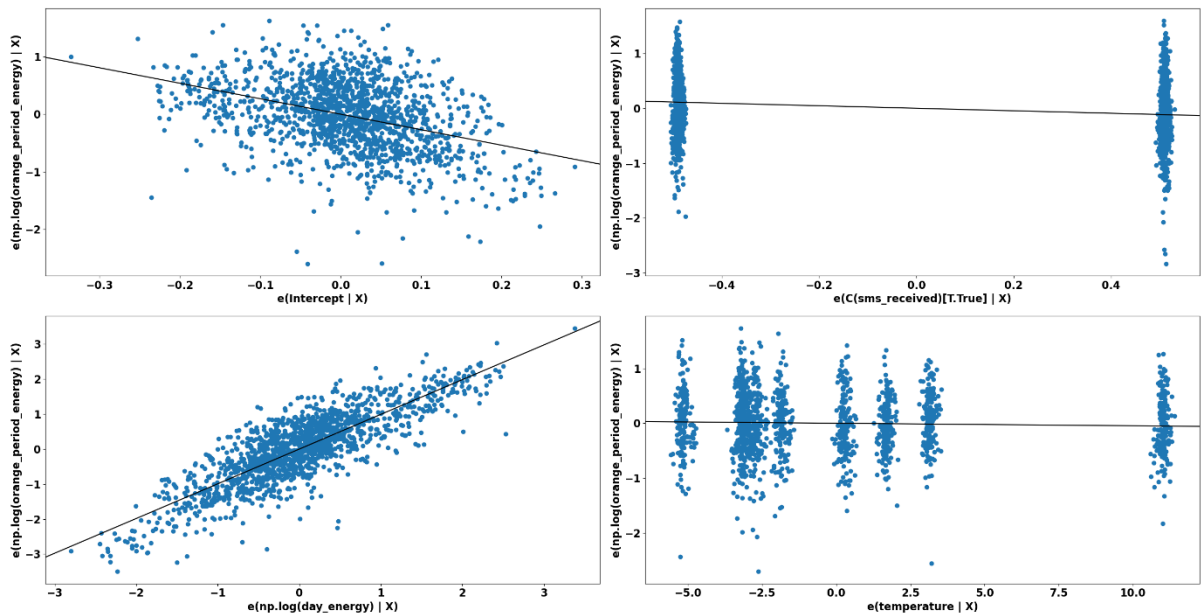


C.5 The results of OLS regression for orange alerts of session 1

OLS Regression Results						
Dep. Variable:	np.log(orange_period_energy)	R-squared:	0.704			
Model:	OLS	Adj. R-squared:	0.704			
Method:	Least Squares	F-statistic:	1195.			
Date:	Fri, 10 Dec 2021	Prob (F-statistic):	0.00			
Time:	15:29:24	Log-Likelihood:	-1248.1			
No. Observations:	1508	AIC:	2504.			
Df Residuals:	1504	BIC:	2525.			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	-2.6969	0.163	-16.538	0.000	-3.017	-2.377
C(sms_received) [T.True]	-0.2278	0.029	-7.979	0.000	-0.284	-0.172
np.log(day_energy)	0.9917	0.017	59.171	0.000	0.959	1.025
temperature	-0.0046	0.003	-1.491	0.136	-0.011	0.001
Omnibus:	48.030	Durbin-Watson:	2.034			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	80.152			
Skew:	-0.266	Prob(JB):	3.94e-18			
Kurtosis:	3.996	Cond. No.	133.			

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



C.6 The results of OLS regression for orange alerts of session 2

OLS Regression Results

```

=====
Dep. Variable:      np.log(orange_period_energy)      R-squared:      0.741
Model:              OLS                             Adj. R-squared: 0.740
Method:            Least Squares                    F-statistic:    1837.
Date:              Fri, 10 Dec 2021                 Prob (F-statistic): 0.00
Time:              15:36:17                         Log-Likelihood: -1370.6
No. Observations: 1932                             AIC:            2749.
Df Residuals:     1928                             BIC:            2771.
Df Model:         3
Covariance Type:  nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-2.4976	0.130	-19.192	0.000	-2.753	-2.242
C(sms_received) [T.True]	-0.1468	0.022	-6.547	0.000	-0.191	-0.103
np.log(day_energy)	0.9737	0.013	73.659	0.000	0.948	1.000
temperature	-0.0047	0.003	-1.341	0.180	-0.012	0.002

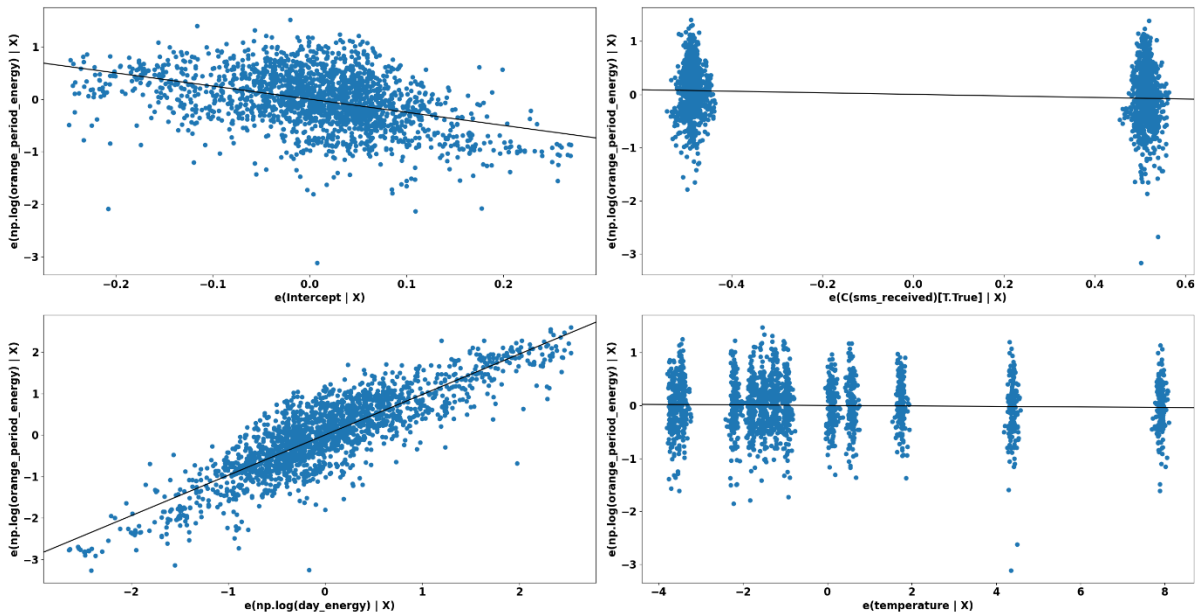
```

=====
Omnibus:          99.792      Durbin-Watson:    1.976
Prob(Omnibus):   0.000      Jarque-Bera (JB): 192.853
Skew:            -0.367      Prob(JB):         1.33e-42
Kurtosis:        4.363      Cond. No.         120.
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



ANNEX D

D.1 The Mann-Whitney U Test for green alert days of both sessions using random forest (sequence to sequence) reference load curve in place of control group

Sr. No	Alert Date	Reference Load Curve			Treated Group			Mann-Whitney U Test			$\mu_r - \mu_T$
		Obs	μ_r	σ_r	Obs	μ_T	σ_T	Stat	p-value	Null hypothesis retained	
Green Session 1											
1	27-07-2020	0.74	0.7	71	0.99	1.21	2374	0.28	Yes	-0.25	0.74
2	25-08-2020	0.79	0.56	34	0.81	0.64	143	0.6	Yes	-0.02	0.79
3	14-09-2020	0.71	0.82	88	0.79	0.83	3579	0.41	Yes	-0.08	0.71
4	17-09-2020	0.71	0.6	88	0.77	0.68	3797	0.41	Yes	-0.06	0.71
5	22-09-2020	0.7	0.57	88	0.79	0.96	3968	0.61	Yes	-0.09	0.7
Green Session 2											
6	14-06-2021	85	0.73	0.63	85	1.01	1	3299	0.16	Yes	-0.28
7	06-09-2021	-	-	-	-	-	-	-	-	-	-
Aggregated Result											
Aggregated		423	0.72	0.67	454	0.86	0.93	93040	0.21	Yes	-0.14

D.2 The Mann-Whitney U Test for orange alert days of both sessions random forest (sequence to sequence) reference load curve in place of control group

Sr. No	Alert Date	Reference Load Curve			Treated Group			Mann-Whitney U Test			$\mu_r - \mu_T$
		Obs	μ_r	σ_r	Obs	μ_T	σ_T	Stat	p-value	Null hypothesis retained	
Orange Session 1											
1	14-11-2019	91	1.63	1.99	91	1.32	2.41	5089	1	Yes	0.31
2	20-12-2019	90	1.09	0.82	90	0.92	0.74	4615	0.95	Yes	0.17

3	20-01-2020	92	1.32	1.28	92	1.3	1.57	4742	0.92	Yes	0.02
4	06-02-2020	91	1.42	1.29	91	1.11	1.31	5194	1	Yes	0.31
5	19-02-2020	91	1.27	1.17	91	1.19	1.29	4656	0.93	Yes	0.08
6	27-02-2020	91	1.18	1.16	91	1.23	1.32	4356	0.73	Yes	-0.05
7	02-03-2020	90	1.15	1.06	90	0.93	1.07	4870	0.99	Yes	0.22
8	04-03-2020	91	1.2	1.21	91	0.98	1.06	4844	0.98	Yes	0.22
9	09-03-2020	90	0.97	1	90	1.26	1.91	3883	0.32	Yes	-0.29
Orange Session 2											
10	30-11-2020	88	1.43	1.03	88	1.21	0.9	4427	0.95	Yes	0.22
11	02-12-2020	89	1.48	1.26	89	1.4	1.06	4092	0.65	Yes	0.08
12	04-12-2020	89	1.35	1.1	89	1.25	1.15	4298	0.84	Yes	0.1
13	07-12-2020	89	1.39	1.13	89	1.4	1.43	4242	0.79	Yes	-0.01
14	09-12-2020	89	1.33	1.03	89	1.48	1.32	3913	0.45	Yes	-0.15
15	18-12-2020	88	1.47	1.35	88	1.19	1.05	4503	0.97	Yes	0.28
16	04-01-2021	88	1.49	1.51	88	1.57	1.79	3954	0.6	Yes	-0.08
17	06-01-2021	88	1.62	1.58	88	1.61	2.11	4251	0.87	Yes	0.01
18	08-01-2021	88	1.61	1.74	88	1.4	1.97	4563	0.98	Yes	0.21
19	11-01-2021	88	1.54	1.88	88	1.75	1.92	3653	0.26	Yes	-0.21
20	26-01-2021	87	1.34	1.19	87	1.37	1.23	3830	0.56	Yes	-0.03
21	11-02-2021	87	1.35	1.41	87	1.23	1.44	4116	0.84	Yes	0.12
Aggregated Result											
Aggregated		187						19299			
		5	1.36	1.34	1875	1.29	1.51	17	1	Yes	0.07

ANNEX E

E.1 The split of IRISE households data with respect to each appliance

Sr. No.	Appliance Name	Households		
		Training subset	Validation subset	Test subset
01	Hot plate	908, 916, 946, 947, 967, 974, 975, 980	982	983
02	Electric Heating	909, 914, 921, 928, 934, 936, 938, 944, 945, 953, 954, 959, 967, 968, 969, 971, 974, 975, 977, 978, 982	983	990
03	Electric Oven	908, 916, 919, 922, 926, 927, 931, 935, 937, 939, 943, 946, 947, 949, 952, 955, 956, 966, 970, 973, 974, 978, 979, 980, 982	996	997
04	Electric Cooker	993, 904, 906, 909, 910, 911, 912, 914, 915, 917, 918, 921, 923, 924, 928, 932, 933, 938, 941, 945, 948, 951, 964, 968, 972, 984, 985, 986	987	902
05	Clothes Drier	901, 902, 904, 907, 908, 914, 915, 927, 928, 930, 931, 934, 936, 938, 939, 940, 943, 944, 946, 947, 952, 959, 960, 962, 965, 966, 970, 972, 978, 987, 989	992	996
06	Dishwasher	901, 903, 907, 908, 910, 914, 916, 917, 920, 922, 925, 926, 928, 930, 931, 933, 936, 937, 938, 939, 940, 941, 943, 944, 946, 947, 948, 949, 950, 952, 953, 961, 964, 966, 969, 970, 972, 974, 979, 980, 981, 983, 984, 985, 989, 991, 993	996	997
07	Microwave Oven	901, 902, 905, 906, 907, 908, 909, 910, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 925, 927, 928, 929, 931, 932, 933, 934, 937, 939, 941, 942, 943, 945, 946, 947, 949, 950, 951, 952, 956, 957, 958, 959, 960, 962, 964, 967, 969, 971, 973, 975, 976, 977, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995	996	997
08	TV	901, 902, 905, 906, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 920, 921, 922, 923, 924, 927, 928, 929, 930, 932, 933, 934, 935, 936, 937, 938, 940, 941, 942, 945, 947, 949, 950, 952, 953, 954, 955, 956, 957, 958, 959, 960, 962, 963, 964, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 992, 993, 994, 995	996	997
09	Washing Machine	900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918,	996	997

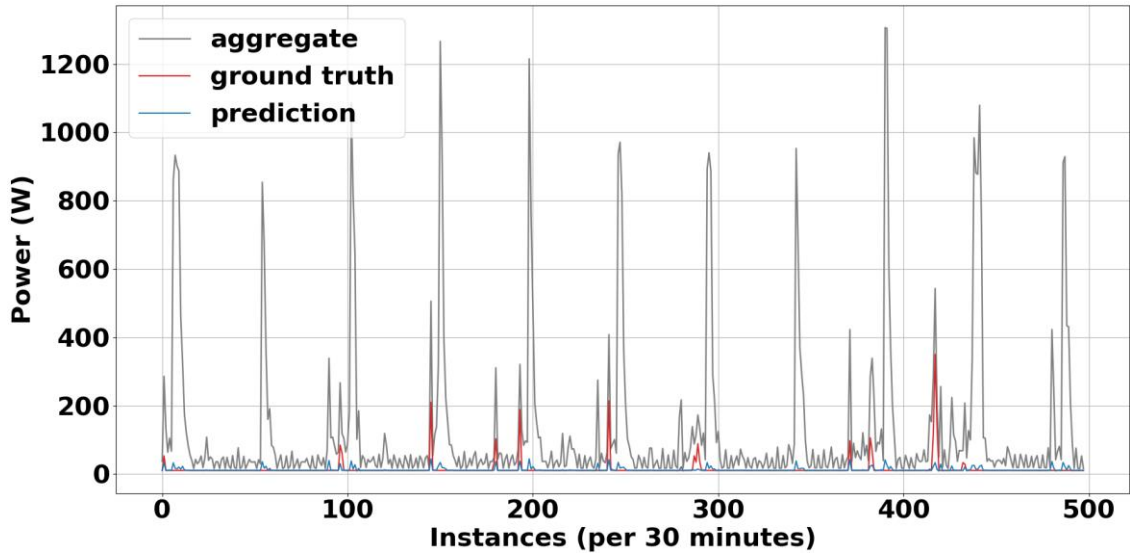
		919, 920, 921, 923, 924, 925, 926, 927, 928, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 971, 972, 973, 975, 976, 978, 979, 980, 981, 984, 985, 986, 987, 988, 989, 990, 991, 993, 994, 995		
--	--	--	--	--

E.2 The length of data subsets for training the predictive model of each appliance

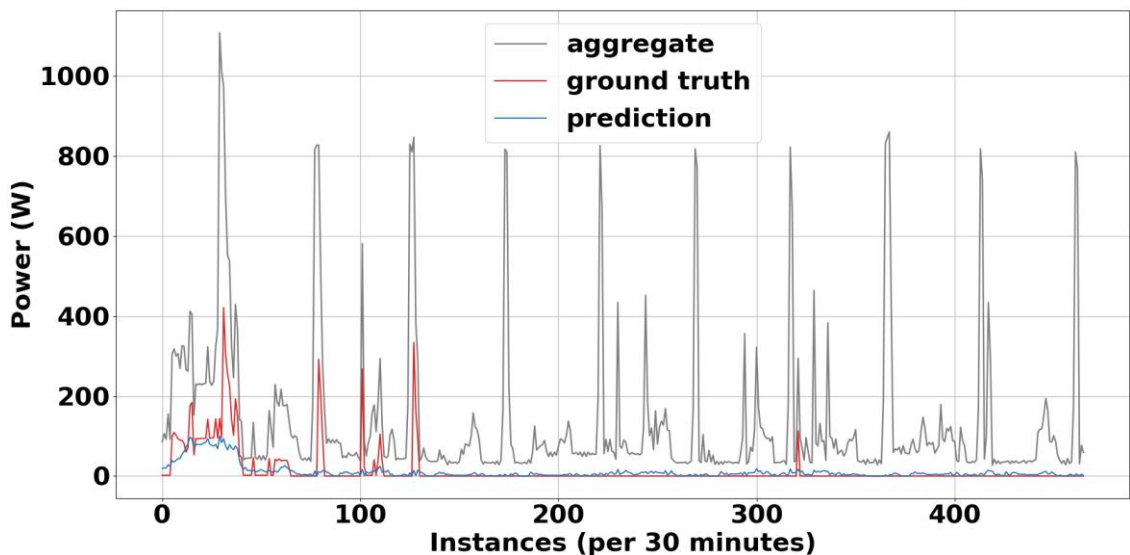
Sr. No.	Appliance	Data sampled at 10 minutes		Data sampled at 30 minutes	
		Training Loss	Validation Loss	Training Loss	Validation Loss
01	Hotplate	0.05	0.05	0.07	0.06
02	Electric Heating	0.14	0.02	0.09	0.02
03	Electric Oven	0.08	0.06	0.12	0.11
04	Electric Cooker	0.11	0.12	0.15	0.16
05	Clothes drier	0.06	0.02	0.09	0.22
06	Dishwasher	0.03	0.02	0.07	0.03
07	Microwave	0.06	0.03	0.09	0.04
08	TV	0.47	0.37	0.46	0.37
09	Washing Machine	0.13	0.12	0.10	0.09

ANNEX F

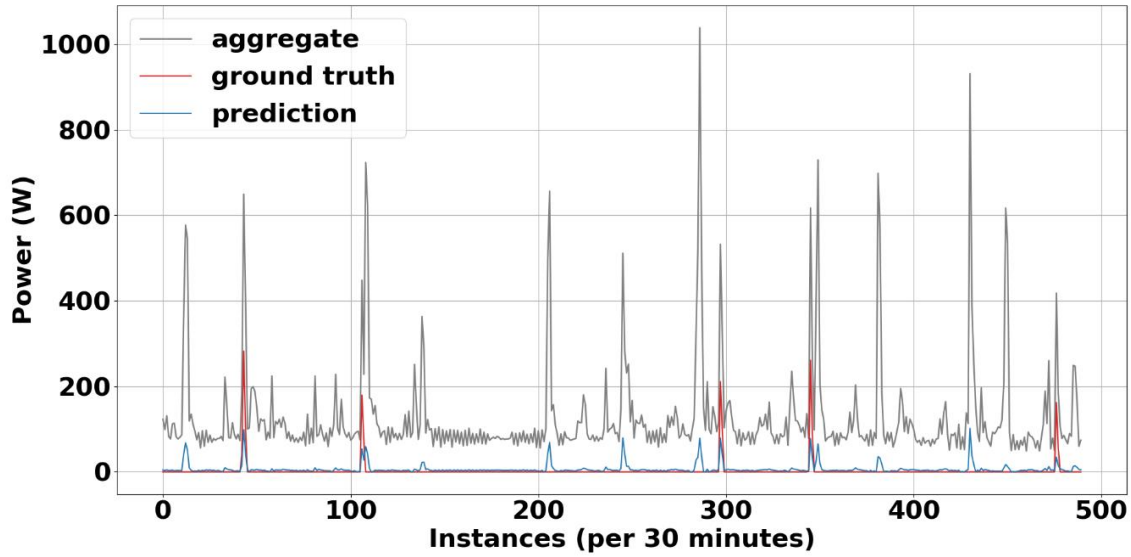
F.1 Test result for hotplate on IRISE house 993 at sampling rate of 30 minutes



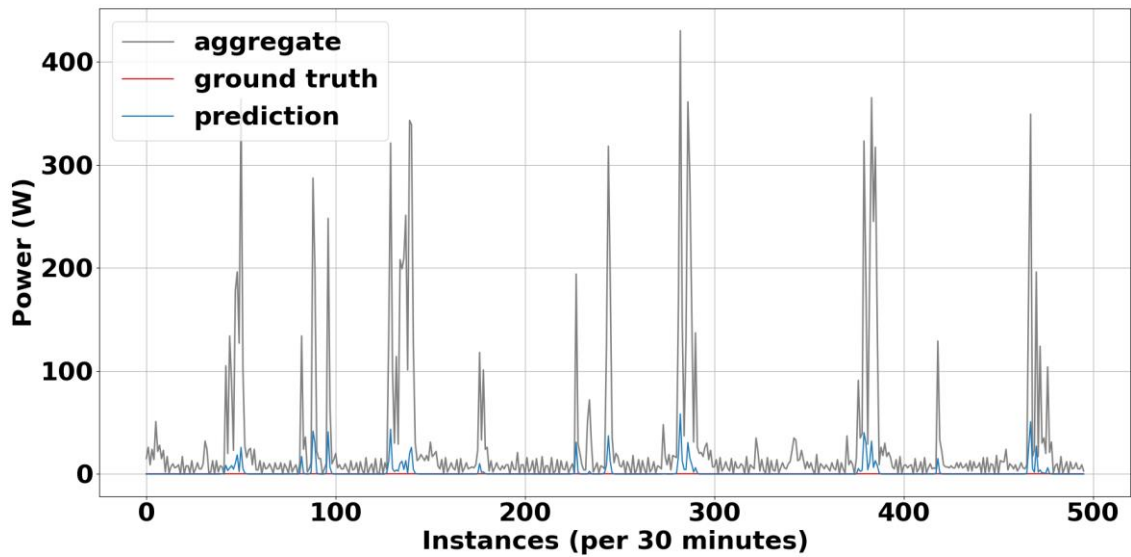
F.2 Test result for electric heating on IRISE house 990 at sampling rate of 30 minutes



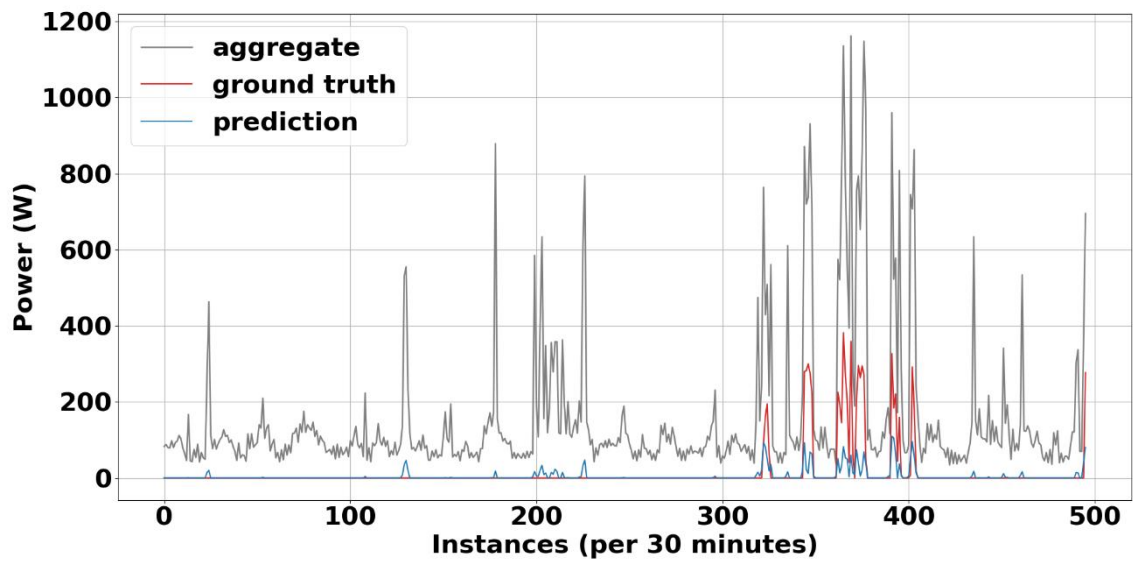
F.3 Test result for electric oven on IRISE house 997 at sampling rate of 30 minutes



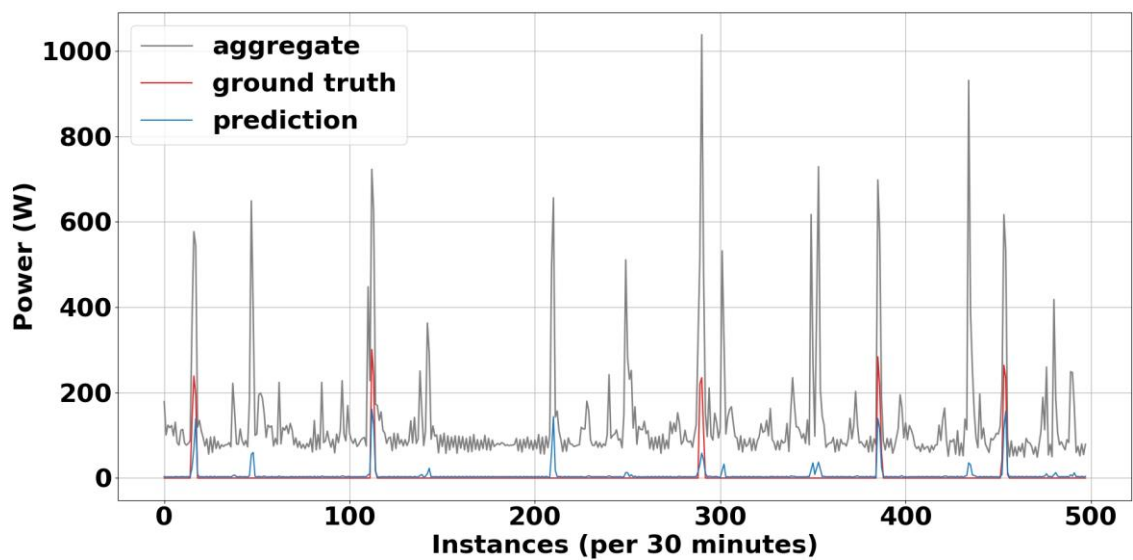
F.4 Test result for electric cooker on IRISE house 902 at sampling rate of 30 minutes



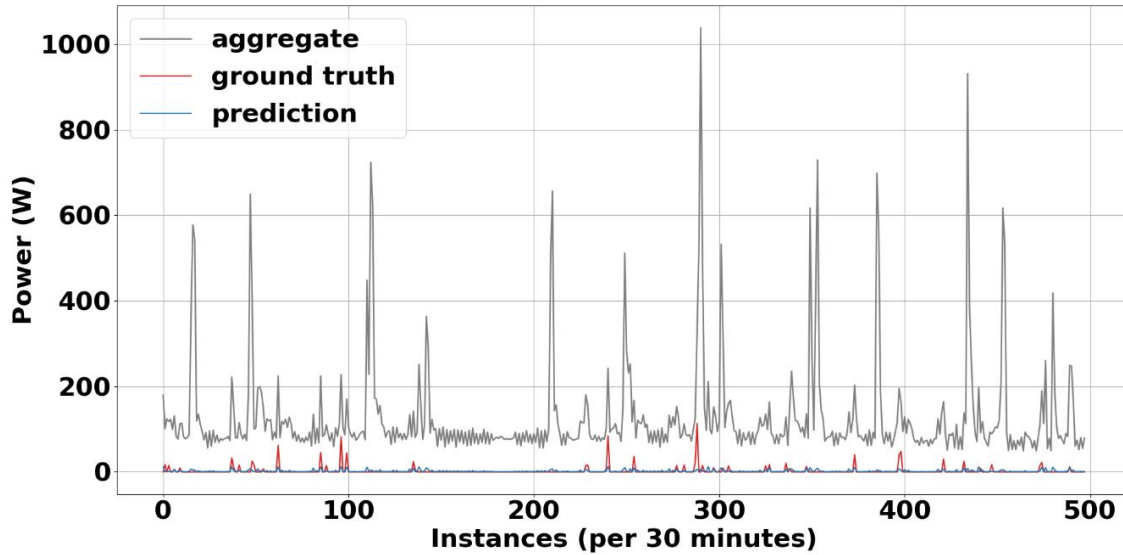
F.5 Test result for clothes drier on IRISE house 996 at sampling rate of 30 minutes



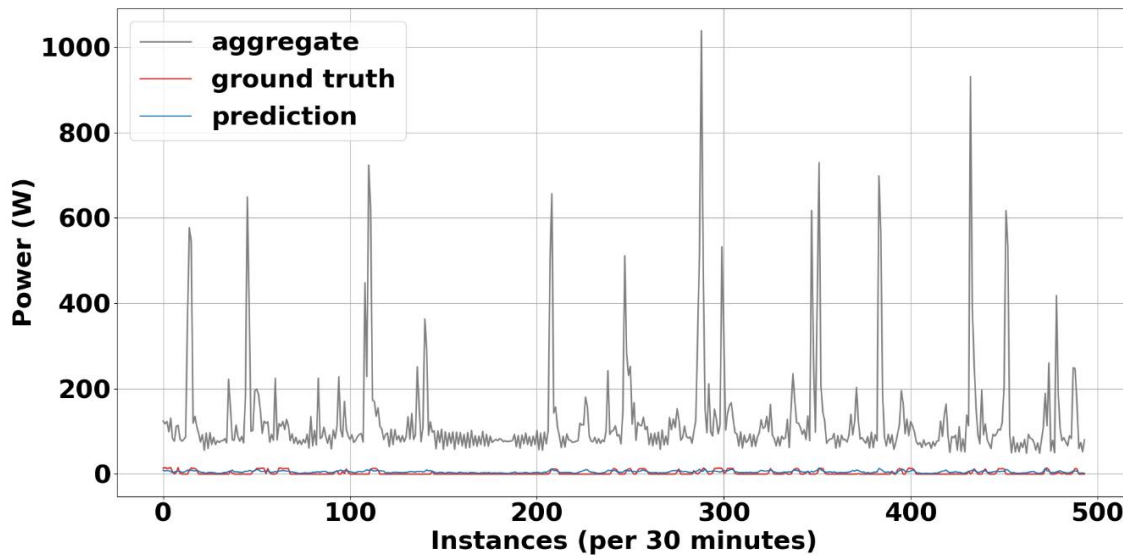
F.6 Test result for dishwasher on IRISE house 997 at sampling rate of 30 minutes



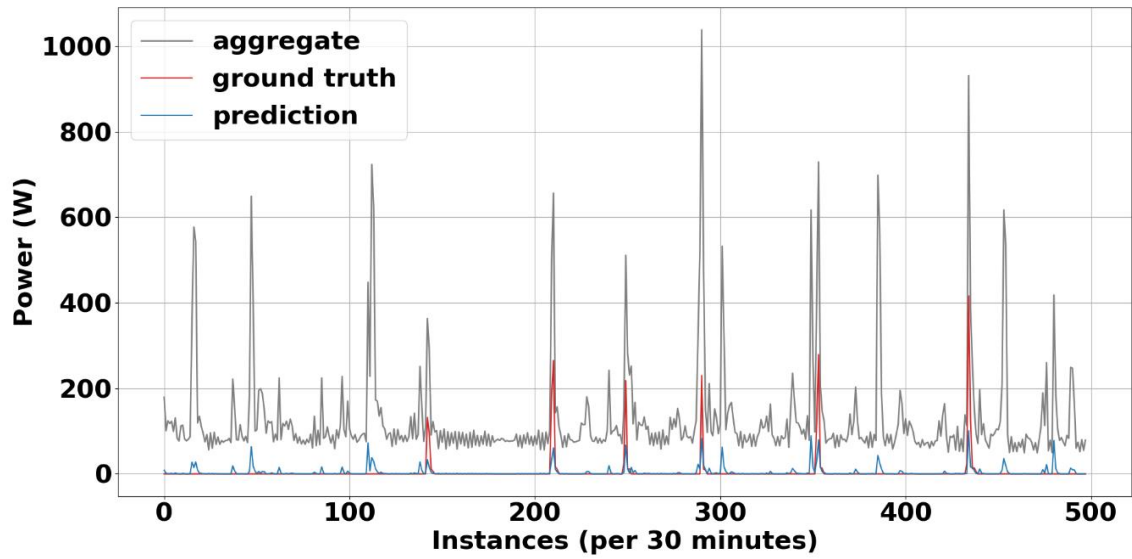
F.7 Test result for microwave oven on IRISE house 997 at sampling rate of 30 minutes



F.8 Test result for tv on IRISE house 997 at sampling rate of 30 minutes



F.9 Test result for washing machine on IRISE house 997 at sampling rate of 30 minutes



BIBLIOGRAPHY

- [1] Abreu, J. M., Câmara Pereira, F., & Ferrão, P. (2012). Using pattern recognition to identify habitual behavior in residential electricity consumption. *Energy and Buildings*, 49, 479-487. <https://doi.org/10.1016/j.enbuild.2012.02.044>
- [2] ADEME. (2014). *Energy Transition Scenarios* (p. 46). French Environment and Energy management Agency. <https://www.ademe.fr/sites/default/files/assets/documents/ademe-energy-transition-scenarios-2030-2050-english-french-7942.pdf>
- [3] ADEME. (2015). *Un mix électrique 100% renouvelable? Analyses et optimisations* (p. 166). <https://www.ademe.fr/sites/default/files/assets/documents/mix-electrique-rapport-2015.pdf>
- [4] ADEME. (2018). *Climat, Air et énergie* (p. 224). https://www.ademe.fr/sites/default/files/assets/documents/2018-climat-air-energie_chiffres-cles-010354.pdf
- [5] AL Shaqsi, A. Z., Sopian, K., & Al-Hinai, A. (2020). Review of energy storage services, applications, limitations, and benefits. *Energy Reports*, 6, 288-306. <https://doi.org/10.1016/j.egyr.2020.07.028>
- [6] Alasseri, R., Rao, T. J., & Sreekanth, K. J. (2021). Pre-implementation assessment for introducing direct load control strategies in the residential electricity sector. *International Journal of Energy and Environmental Engineering*, 12(3), 433-451. <https://doi.org/10.1007/s40095-020-00378-6>
- [7] Albadi, M. H., & El-Saadany, E. F. (2007). Demand Response in Electricity Markets: An Overview. *2007 IEEE Power Engineering Society General Meeting*, 1-5. <https://doi.org/10.1109/PES.2007.385728>
- [8] Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9), 1082-1095. <https://doi.org/10.1016/j.jpubeco.2011.03.003>

- [9] Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, 104(10), 3003-3037. <https://doi.org/10.1257/aer.104.10.3003>
- [10] Amanatidis, G. (2019). *European policies on climate and energy towards 2020, 2030 and 2050*.
- [11] Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. *The Journal of Law, Economics, and Organization*, 29(5), 992-1022. <https://doi.org/10.1093/jleo/ews020>
- [12] Bahar, H., & Sauvage, J. (2013). *Cross-Border Trade in Electricity and the Development of Renewables-Based Electric Power: Lessons from Europe*. <https://doi.org/10.1787/5k4869cdwnzr-en>
- [13] Baldassarri, D., & Abascal, M. (2017). Field Experiments Across the Social Sciences. *Annual Review of Sociology*, 43(1), 41-73. <https://doi.org/10.1146/annurev-soc-073014-112445>
- [14] Banerje, A. V., & Duflo, E. (2011). *A Radical Rethinking of the Way to Fight Global poverty* (p. 13).
- [15] Bartolome, C. A. M. de. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 1(56), 79-96.
- [16] Basu, K., Debusschere, V., Bacha, S., Maulik, U., & Bondyopadhyay, S. (2015). Nonintrusive Load Monitoring: A Temporal Multilabel Classification Approach. *IEEE Transactions on Industrial Informatics*, 11(1), 262-270. <https://doi.org/10.1109/TII.2014.2361288>
- [17] Batra, N., Kelly, J., Parson, O., Dutta, H., Knottenbelt, W., Rogers, A., Singh, A., & Srivastava, M. (2014). NILMTK: an open source toolkit for non-intrusive load monitoring. *Proceedings of the 5th international conference on Future energy systems*, 265–276. <https://doi.org/10.1145/2602044.2602051>

- [18] Bénabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3), 489-520. <https://doi.org/10.1111/1467-937X.00253>
- [19] Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5), 1652-1678. <https://doi.org/10.1257/aer.96.5.1652>
- [20] Bewick, V., Cheek, L., & Ball, J. (2003). Statistics review 8: Qualitative data – tests of association. *Critical Care*, 8(1), 46. <https://doi.org/10.1186/cc2428>
- [21] Bhotto, M. Z. A., Makonin, S., & Bajić, I. V. (2017). Load Disaggregation Based on Aided Linear Integer Programming. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 64(7), 792-796. <https://doi.org/10.1109/TCSII.2016.2603479>
- [22] Bianchi, F. M., Maiorino, E., Kampffmeyer, M. C., Rizzi, A., & Jenssen, R. (2017). An overview and comparative analysis of Recurrent Neural Networks for Short Term Load Forecasting. *arXiv:1705.04378 [cs]*. <https://doi.org/10.1007/978-3-319-70338-1>
- [23] Bivas, P. (2011). La production d'effacement : comment offrir des économies d'électricité à des millions de foyers. *Le journal de l'école de Paris du management*, n°90(4), 8-14.
- [24] Bonfigli, R., Felicetti, A., Principi, E., Fagiani, M., Squartini, S., & Piazza, F. (2018). Denoising autoencoders for Non-Intrusive Load Monitoring: Improvements and comparative evaluation. *Energy and Buildings*, 158, 1461-1474. <https://doi.org/10.1016/j.enbuild.2017.11.054>
- [25] Brandon, G., & Lewis, A. (1999). REDUCING HOUSEHOLD ENERGY CONSUMPTION: A QUALITATIVE AND QUANTITATIVE FIELD STUDY. *Journal of Environmental Psychology*, 19(1), 75-85. <https://doi.org/10.1006/jevp.1998.0105>
- [26] Browning, G., Halcli, A., & Webster, F. (1999). *Understanding Contemporary Society: Theories of the Present*. SAGE.

- [27] Buckley, P., & Llerena, D. (2018, juin). Demand response as a common pool resource game: Nudges or peak pricing. *6th World Congress of Environmental and Resource Economists*. <https://hal.archives-ouvertes.fr/hal-02088189>
- [28] Centre d'analyse stratégique. (2011). *Nudges green : new incentives for ecological behavior*. <http://archives.strategie.gouv.fr/cas/en/content/policy-brief-216-nudges-green-new-incentives-green-behavior-march-2011.html>
- [29] Chen, X.-W., & Lin, X. (2014). Big Data Deep Learning: Challenges and Perspectives. *IEEE Access*, 2, 514-525. <https://doi.org/10.1109/ACCESS.2014.2325029>
- [30] Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. <https://doi.org/10.7717/peerj-cs.623>
- [31] Climate Action Tracker. (2021). *Warming Projections Global Update*. https://climateactiontracker.org/documents/853/CAT_2021-05-04_Briefing_Global-Update_Climate-Summit-Momentum.pdf
- [32] Costa, D. L. (2010). *Energy Conservation “Nudges” and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment*. 34.
- [33] Costa, D. L., & Kahn, M. E. (2013). Energy Conservation “Nudges” and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment. *Journal of the European Economic Association*, 11(3), 680-702. <https://doi.org/10.1111/jeea.12011>
- [34] Crits-Christoph, P. (1992). The efficacy of brief dynamic psychotherapy : a meta-analysis. *The efficacy of brief dynamic psychotherapy : a meta-analysis*, 149(2), 151-158.
- [35] Darby, S. (2001). Making it Obvious: Designing Feedback into Energy Consumption. In P. Bertoldi, A. Ricci, & A. de Almeida (Éds.), *Energy Efficiency in Household*

- Appliances and Lighting* (p. 685-696). Springer. https://doi.org/10.1007/978-3-642-56531-1_73
- [36] Delinchant, B., Wurtz, F., Ploix, S., Schanen, J.-L., & Marechal, Y. (2016). GreEn-ER living lab: A green building with energy aware occupants. *2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)*, 1-8.
- [37] Desing, H., & Widmer, R. (2021). *Reducing climate risks with fast and complete energy transitions: applying the precautionary principle to the Paris agreement*. OSF Preprints. <https://doi.org/10.31219/osf.io/5wf64>
- [38] D’Incecco, M., Squartini, S., & Zhong, M. (2020). Transfer Learning for Non-Intrusive Load Monitoring. *IEEE Transactions on Smart Grid*, 11(2), 1419-1429. <https://doi.org/10.1109/TSG.2019.2938068>
- [39] Dolan, P., & Metcalfe, R. (2015). *Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation* (SSRN Scholarly Paper ID 2589269). Social Science Research Network. <https://papers.ssrn.com/abstract=2589269>
- [40] Duplex, J., Gosswiller, S., & Fagnoni, S. (2013). *A better knowledge of electricity consumption for residential customers through the Linky smart meter*. 1333-1333. <https://doi.org/10.1049/cp.2013.1188>
- [41] Dütschke, E., & Paetz, A.-G. (2013). Dynamic electricity pricing—Which programs do consumers prefer? *Energy Policy*, 59, 226-234. <https://doi.org/10.1016/j.enpol.2013.03.025>
- [42] Egarter, D., Sobe, A., & Elmenreich, W. (2013). Evolving Non-Intrusive Load Monitoring. In A. I. Esparcia-Alcázar (Éd.), *Applications of Evolutionary Computation* (p. 182-191). Springer. https://doi.org/10.1007/978-3-642-37192-9_19
- [43] Ehrhardt-Martinez, K., & Donnelly, K. A. (2010). *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities*. 140.

- [44] Eid, C., Codani, P., Chen, Y., Perez, Y., & Hakvoort, R. (2015). Aggregation of demand side flexibility in a smart grid: A review for European market design. *2015 12th International Conference on the European Energy Market (EEM)*, 1-5. <https://doi.org/10.1109/EEM.2015.7216712>
- [45] ENEDIS. (2021). *Profile Coefficients* [Data Portal]. Open Data. <https://data.enedis.fr/explore/dataset/coefficients-des-profils/information>
- [46] Energie Developpement. (2013). *Puissance moyenne des appareils électriques les plus courants.* <http://energie-developpement.blogspot.com/2011/09/consommations-classiques-des-appareils.html>
- [47] Energie Douce. (2021). *Tableau de consommation des appareils électroménagers.* <https://www.energiesdouce.com/content/12-conseils-faq-consommation-electrique-des-appareils-electromenagers>
- [48] Energie Mobile. (2021). *Tableau Estimatif des consommateurs.* <http://www.energiemobile.com/img/cms/Tableau%20de%20consommation.pdf>
- [49] Enertech. (2008). *Mesure de la consommation des usages domestiques de l'audiovisuel et de l'informatique - Project REMODECE.*
- [50] ENOPTEA. (2018). *A quoi correspondent les journées PP1 et PP2 ?* <https://www.enoptea.fr/>. <https://www.enoptea.fr/a-quoi-correspondent-les-journees-pp1-et-pp2/>
- [51] European Commission. (2018). *A Clean Planet for all A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy* (Communication COM(2018) 773 final; p. 25).
- [52] European Commission. (2020). *Kick-starting the journey towards a climate-neutral Europe by 2050; EU Climate Action Progress Report* (p. 55). https://ec.europa.eu/clima/sites/default/files/strategies/progress/docs/com_2020_777_en.pdf

- [53] Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, 59(1), 255-278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- [54] Evans, N., Eickers, S., Geene, L., Todorovic, M., & Villmow, A. (2017). *Green Nudging: A discussion and preliminary evaluation of nudging as an environmental policy instrument*. <https://doi.org/10.17169/refubium-25251>
- [55] Faruqui, A., & Sergici, S. (2010). Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of Regulatory Economics*, 38(2), 193-225. <https://doi.org/10.1007/s11149-010-9127-y>
- [56] Faustine, A., Pereira, L., Bousbiat, H., & Kulkarni, S. (2020). UNet-NILM: A Deep Neural Network for Multi-tasks Appliances State Detection and Power Estimation in NILM. *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*, 84–88. <https://doi.org/10.1145/3427771.3427859>
- [57] Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1(1), 79-104. <https://doi.org/10.1007/s12053-008-9009-7>
- [58] Giet, L. (2019). *CERRE 2019 | Smart Consumers in the Internet of Energy: Flexibility Markets & Services from Distributed Energy Resources*. 86.
- [59] Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*.
- [60] Glerup, M., Larsen, A., Leth-Petersen, S., & Togeby, M. (2010). The Effect of Feedback by Text Message (SMS) and Email on Household Electricity Consumption: Experimental Evidence. *The Energy Journal*, 31(3). <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol31-No3-6>
- [61] Gollwitzer, L., Ockwell, D., Muok, B., Ely, A., & Ahlborg, H. (2018). Rethinking the sustainability and institutional governance of electricity access and mini-grids: Electricity as a common pool resource. *Energy Research & Social Science*, 39, 152-161. <https://doi.org/10.1016/j.erss.2017.10.033>

- [62] Goulden, M., Bedwell, B., Rennick-Egglestone, S., Rodden, T., & Spence, A. (2014). Smart grids, smart users? The role of the user in demand side management. *Energy Research & Social Science*, 2, 21-29. <https://doi.org/10.1016/j.erss.2014.04.008>
- [63] Gowdy, J. M. (2008). Behavioral economics and climate change policy. *Journal of Economic Behavior & Organization*, 68(3), 632-644. <https://doi.org/10.1016/j.jebo.2008.06.011>
- [64] Graffeo, M., Ritov, I., Bonini, N., & Hadjichristidis, C. (2015). To make people save energy tell them what others do but also who they are: a preliminary study. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01287>
- [65] Haque, A. N. M. M., Nijhuis, M., Ye, G., Nguyen, H. P., Blik, F. W., & Slootweg, J. G. (2019). Integrating direct and indirect load control for congestion management in LV networks. *IEEE Transactions on Smart Grid*, 10(1), 741-751. <https://doi.org/10.1109/TSG.2017.2751743>
- [66] Hargreaves, T., Nye, M., & Burgess, J. (2010). Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy*, 38(10), 6111-6119. <https://doi.org/10.1016/j.enpol.2010.05.068>
- [67] Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12), 1870-1891. <https://doi.org/10.1109/5.192069>
- [68] Hatton, L., & Charpentier, P. (2014). *Système électrique français : estimation de l'effacement des clients résidentiels*. 14.
- [69] Haynes, L., Service, O., Goldacre, B., & Torgerson, D. (2012). Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2131581>
- [70] He, K., Stankovic, L., Liao, J., & Stankovic, V. (2018). Non-Intrusive Load Disaggregation Using Graph Signal Processing. *IEEE Transactions on Smart Grid*, 9(3), 1739-1747. <https://doi.org/10.1109/TSG.2016.2598872>
- [71] Hebrail, G., & Berard, A. (2012). *Individual household electric power consumption Data Set*. UCI Machine Learning Repository.

<http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption>

- [72] Hutcheson, G. D., & Sofroniou, N. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. SAGE.
- [73] International Energy Agency. (2018). *Power system flexibility for the energy transition, Part 1: Overview for policy makers*. 48.
- [74] Ito, K. (2014). Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing. *American Economic Review*, 104(2), 537-563. <https://doi.org/10.1257/aer.104.2.537>
- [75] Ito, K., Ida, T., & Tanaka, M. (2018). Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand. *American Economic Journal: Economic Policy*, 10(1), 240-267. <https://doi.org/10.1257/pol.20160093>
- [76] Jensen, S. Ø., Marszal-Pomianowska, A., Lollini, R., Pasut, W., Knotzer, A., Engelmann, P., Stafford, A., & Reynders, G. (2017). IEA EBC Annex 67 Energy Flexible Buildings. *Energy and Buildings*, 155, 25-34. <https://doi.org/10.1016/j.enbuild.2017.08.044>
- [77] Jorgensen, B. S., Fumei, S., & Byrne, G. (2021). Reducing Peak Energy Demand among Residents Who Are Not Billed for Their Electricity Consumption: Experimental Evaluation of Behaviour Change Interventions in a University Setting. *International Journal of Environmental Research and Public Health*, 18(16), 8406. <https://doi.org/10.3390/ijerph18168406>
- [78] Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- [79] Karjalainen, S. (2011). Consumer preferences for feedback on household electricity consumption. *Energy and Buildings*, 43(2), 458-467. <https://doi.org/10.1016/j.enbuild.2010.10.010>
- [80] Karunanithi, K., Saravanan, S., Prabakar, B. R., Kannan, S., & Thangaraj, C. (2017). Integration of Demand and Supply Side Management strategies in Generation

- Expansion Planning. *Renewable and Sustainable Energy Reviews*, 73, 966-982.
<https://doi.org/10.1016/j.rser.2017.01.017>
- [81] Kelly, D. (2016). *Disaggregation of domestic smart meter energy data*.
<https://doi.org/10.25560/49452>
- [82] Kelly, J., & Knottenbelt, W. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Scientific Data*, 2(1), 150007. <https://doi.org/10.1038/sdata.2015.7>
- [83] Kim, H., Marwah, M., Arlitt, M., Lyon, G., & Han, J. (2011). Unsupervised Disaggregation of Low Frequency Power Measurements. In *Proceedings of the 2011 SIAM International Conference on Data Mining (SDM)* (p. 747-758). Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9781611972818.64>
- [84] Kim, J., Kim, J., Thi Thu, H. L., & Kim, H. (2016). Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection. *2016 International Conference on Platform Technology and Service (PlatCon)*, 1-5. <https://doi.org/10.1109/PlatCon.2016.7456805>
- [85] Kolter, J. Z., & Johnson, M. J. (2011). *REDD: A Public Data Set for Energy Disaggregation Research*. 6.
- [86] Krishnamurti, T., Schwartz, D., Davis, A., Fischhoff, B., de Bruin, W. B., Lave, L., & Wang, J. (2012). Preparing for smart grid technologies: A behavioral decision research approach to understanding consumer expectations about smart meters. *Energy Policy*, 41, 790-797. <https://doi.org/10.1016/j.enpol.2011.11.047>
- [87] Lars Biesewig, & Annie Krautkraemer. (2018). Nudging toward a cleaner world [Blog]. *The TSEconomist*. <https://thetseconomist.com/2018/01/23/nudging-toward-a-cleaner-world/>
- [88] Le Dréau, J., Vellei, M., & Abdelouadoud, Y. (2019). *A Bottom-Up Model to Evaluate the Flexibility of French Residential Wet Appliances*. 2173-2180. <https://doi.org/10.26868/25222708.2019.210441>

- [89] Le, T.-T.-H., Kim, J., & Kim, H. (2016). Classification performance using gated recurrent unit recurrent neural network on energy disaggregation. *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, 1, 105-110. <https://doi.org/10.1109/ICMLC.2016.7860885>
- [90] Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. I., & Naccarino, J. (1996). Comparison of very short-term load forecasting techniques. *IEEE Transactions on Power Systems*, 11(2), 877-882. <https://doi.org/10.1109/59.496169>
- [91] Liu, Y., Liu, C., Shen, Y., Zhao, X., Gao, S., & Huang, X. (2021). Non-intrusive energy estimation using random forest based multi-label classification and integer linear programming. *Energy Reports*, 7, 283-291. <https://doi.org/10.1016/j.egyr.2021.08.045>
- [92] Llerena, D., Roussillon, B., Teyssier, S., Buckley, P., Delinchant, B., Ferrari, J., Laranjeira, T., & Wurtz, F. (2021). *Demand response in the workplace: A field experiment* (N° 2021-01; Working Papers). Grenoble Applied Economics Laboratory (GAEL). <https://ideas.repec.org/p/gbl/wpaper/2021-01.html>
- [93] Loewenstein, G. (1999). EXPERIMENTAL ECONOMICS FROM THE VANTAGE-POINT OF BEHAVIOURAL ECONOMICS. *THE ECONOMIC JOURNAL*, 10.
- [94] Lovie, P. (2005). Coefficient of Variation. In *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470013192.bsa107>
- [95] Makonin, S., Ellert, B., Bajić, I. V., & Popowich, F. (2016). Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014. *Scientific Data*, 3(1), 160037. <https://doi.org/10.1038/sdata.2016.37>
- [96] Manuel Baude, Aurore Colin, Jérôme Duvernoy, & Alexis Foussard. (2021). *Chiffres clés du climat, : France, Europe et Monde* (p. 92). Ministère de la transition écologique France. <https://statistiques.developpement-durable.gouv.fr/chiffres-cles-du-climat-0?rubrique=26>

- [97] McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 23(2), 143-149. <https://doi.org/10.11613/BM.2013.018>
- [98] Ministry of ecological transition and solidarity France. (2019). *French strategy for energy and climate: Multi-year programs for energy* (p. 400). <https://www.ecologie.gouv.fr/sites/default/files/PPE-Executive%20summary.pdf>
- [99] Moghaddam, Z., Ahmad, I., Habibi, D., & Masoum, M. A. S. (2019). A Coordinated Dynamic Pricing Model for Electric Vehicle Charging Stations. *IEEE Transactions on Transportation Electrification*, 5(1), 226-238. <https://doi.org/10.1109/TTE.2019.2897087>
- [100] Moghram, I., & Rahman, S. (1989). Analysis and evaluation of five short-term load forecasting techniques. *IEEE Transactions on Power Systems*, 4(4), 1484-1491. <https://doi.org/10.1109/59.41700>
- [101] Momsen, K., & Stoerk, T. (2014). From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy*, 74, 376-382. <https://doi.org/10.1016/j.enpol.2014.07.008>
- [102] Moradzadeh, A., Zeinal-Kheiri, S., Mohammadi-Ivatloo, B., Abapour, M., & Anvari-Moghaddam, A. (2020). Support Vector Machine-Assisted Improvement Residential Load Disaggregation. *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 1-6. <https://doi.org/10.1109/ICEE50131.2020.9260869>
- [103] Murray, D., Stankovic, L., & Stankovic, V. (2017). An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study. *Scientific Data*, 4(1), 160122. <https://doi.org/10.1038/sdata.2016.122>
- [104] Nguyen, T.-T., Martin, V., Malmquist, A., & Silva, C. A. S. (2017). A review on technology maturity of small scale energy storage technologies. *Renewable Energy and Environmental Sustainability*, 2, 36. <https://doi.org/10.1051/rees/2017039>
- [105] Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

- [106] Ozawa, A., Furusato, R., & Yoshida, Y. (2016). Determining the relationship between a household's lifestyle and its electricity consumption in Japan by analyzing measured electric load profiles. *Energy and Buildings*, 119, 200-210. <https://doi.org/10.1016/j.enbuild.2016.03.047>
- [107] Parson, O. (2017). *NILM Links*. <http://engdashboard.blogspot.com/2017/10/nilm-links.html>
- [108] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. <https://hal.inria.fr/hal-00650905>
- [109] Pedro Paulo Marques do Nascimento. (2016). *Applications of Deep Learning techniques on NILM*. 85.
- [110] Penha, D., & Castro, A. (2017). *Convolutional Neural Network Applied to the Identification of Residential Equipment in Nonintrusive Load Monitoring Systems*. 11-21. <https://doi.org/10.5121/csit.2017.71802>
- [111] Pinon, X. (2018, juillet 31). *6 kVA : puissance de compteur la plus répandue en France*. Kelwatt.fr. <https://www.kelwatt.fr/guide/six-kva>
- [112] Pool, E. (2015). *Unlocking energy market flexibility and demand side response*. 1-12.
- [113] Potrč, S., Čuček, L., Martin, M., & Kravanja, Z. (2021). Sustainable renewable energy supply networks optimization – The gradual transition to a renewable energy system within the European Union by 2050. *Renewable and Sustainable Energy Reviews*, 146, 111186. <https://doi.org/10.1016/j.rser.2021.111186>
- [114] Pratt, B. W., & Erickson, J. D. (2020). Defeat the Peak: Behavioral insights for electricity demand response program design. *Energy Research & Social Science*, 61, 101352. <https://doi.org/10.1016/j.erss.2019.101352>

- [115] Réseau de Transport d'Électricité France. (2017). *Bilan Electrique 2017* (p. 121).
https://assets.rte-france.com/prod/public/2020-06/bilan_electrique_2017.pdf
- [116] Réseau de Transport d'Électricité France. (2018). *Bilan Electrique 2018* (p. 175).
<https://assets.rte-france.com/prod/public/2021-03/Bilan%20electrique%202018.pdf>
- [117] Réseau de Transport d'Électricité France. (2019). *Bilan Electrique 2019* (p. 173).
- [118] Réseau de Transport d'Électricité France. (2020a). *Bilan Electrique 2020* (p. 185).
https://assets.rte-france.com/prod/public/2021-03/Bilan%20electrique%202020_o.pdf
- [119] Réseau de Transport d'Électricité France. (2020b). *Relations entre le GRD et Le Responsable d'équilibre : Dispositions Applicables Pour L'Estimation des Courbes de Charges*.
https://www.services-rte.com/files/live/sites/services-rte/files/documentsLibrary/2020-06-01_REGLES_MARE_Section_2_F_4688_fr
- [120] Schirmer, P. A., & Mporas, I. (2019). Statistical and Electrical Features Evaluation for Electrical Appliances Energy Disaggregation. *Sustainability*, 11(11), 3222.
<https://doi.org/10.3390/su11113222>
- [121] Schneider, I., & Sunstein, C. R. (2017). Behavioral considerations for effective time-varying electricity prices. *Behavioural Public Policy*, 1(2), 219-251.
<https://doi.org/10.1017/bpp.2017.2>
- [122] Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5), 429-434. <https://doi.org/10.1111/j.1467-9280.2007.01917.x>
- [123] Severin Borenstein. (2009). *To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing* (p. 38).
- [124] Shafir, E. (2013). *The Behavioral Foundations of Public Policy*. Princeton University Press.

- [125] Shroff, A., Thia Jang Ping, Elaine Tan, Do Hoang Van Khanh, Sharon Tham, Leong Wai Yan, Lim Yong Long, & Eugene Liow. (2019). *Evidence-based Policymaking in Singapore: A Policymaker's Toolkit*. <https://www.csc.gov.sg/resources/book-evidence-based-policymaking-in-singapore-a-policymaker-s-toolkit>
- [126] Sirin, S. M., & Gonul, M. S. (2016). Behavioral aspects of regulation: A discussion on switching and demand response in Turkish electricity market. *Energy Policy*, 97, 591-602. <https://doi.org/10.1016/j.enpol.2016.08.005>
- [127] Srinivasarengan, K., Goutam, Y. G., Chandra, M. G., & Kadhe, S. (2013). A Framework for Non Intrusive Load Monitoring Using Bayesian Inference. *2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 427-432. <https://doi.org/10.1109/IMIS.2013.78>
- [128] Stankovic, V., Liao, J., & Stankovic, L. (2014). A graph-based signal processing approach for low-rate energy disaggregation. *2014 IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES)*, 81-87. <https://doi.org/10.1109/CIES.2014.7011835>
- [129] TARG Bristol. (2017, novembre 13). *A three minute primer on meta-analysis*. <https://www.youtube.com/watch?v=i675gZNe3MY>
- [130] Taylor, J. W., & McSharry, P. E. (2007). Short-Term Load Forecasting Methods: An Evaluation Based on European Data. *IEEE Transactions on Power Systems*, 22(4), 2213-2219. <https://doi.org/10.1109/TPWRS.2007.907583>
- [131] Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- [132] Tin Kam Ho. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1, 278-282 vol.1. <https://doi.org/10.1109/ICDAR.1995.598994>
- [133] Tin Kam Ho. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832-844. <https://doi.org/10.1109/34.709601>

- [134] Tsai, C.-W., Yang, C.-W., Ho, W.-J., Yin, Z.-X., & Chiang, K.-C. (2018). Using autoencoder network to implement non-intrusive load monitoring of small and medium business customer. *2018 IEEE International Conference on Applied System Invention (ICASI)*, 433-436. <https://doi.org/10.1109/ICASI.2018.8394277>
- [135] United Nations. (2015). *Paris Agreement*. https://unfccc.int/sites/default/files/resource/parisagreement_publication.pdf
- [136] Urbina, D. A., & Ruiz-Villaverde, A. (2019). A Critical Review of Homo Economicus from Five Approaches. *American Journal of Economics and Sociology*, 78(1), 63-93. <https://doi.org/10.1111/ajes.12258>
- [137] Valenzuela, J., Thimmapuram, P. R., & Kim, J. (2012). Modeling and simulation of consumer response to dynamic pricing with enabled technologies. *Applied Energy*, 96, 122-132. <https://doi.org/10.1016/j.apenergy.2011.11.022>
- [138] van Bavel, R., Herrmann, B., Esposito, G., & Proestakis, A. (2013). *Applying Behavioural Science to EU Policy-Making* (JRC Working Paper N° JRC83284). Joint Research Centre (Seville site). <https://econpapers.repec.org/paper/iptiptwpa/jrc83284.htm>
- [139] Vassileva, I., Odlare, M., Wallin, F., & Dahlquist, E. (2012). The impact of consumers' feedback preferences on domestic electricity consumption. *Applied Energy*, 93, 575-582. <https://doi.org/10.1016/j.apenergy.2011.12.067>
- [140] Wang, F., Liu, L., Yu, Y., Li, G., Li, J., Shafie-khah, M., Catalão, J., Wang, F., Liu, L., Yu, Y., Li, G., Li, J., Shafie-khah, M., & Catalão, J. P. S. (2018). Impact Analysis of Customized Feedback Interventions on Residential Electricity Load Consumption Behavior for Demand Response. *Energies*, 11(4), 770. <https://doi.org/10.3390/en11040770>
- [141] Wang, X., Wang, J., Shi, D., & Khodayar, M. E. (2018). A Factorial Hidden Markov Model for Energy Disaggregation based on Human Behavior Analysis. *2018 IEEE Power Energy Society General Meeting (PESGM)*, 1-5. <https://doi.org/10.1109/PESGM.2018.8586047>

- [142] Wenninger, M., Maier, A., & Schmidt, J. (2021). DEDDIAG, a domestic electricity demand dataset of individual appliances in Germany. *Scientific Data*, 8(1), 176. <https://doi.org/10.1038/s41597-021-00963-2>
- [143] Wilhite, H. (1999). *Advances in the use of consumption feedback information in energy billing: the experiences of a Norwegian energy utility*. 14.
- [144] Wilhite, H., & Ling, R. (1995). Measured energy savings from a more informative energy bill. *Energy and Buildings*, 22(2), 145-155. [https://doi.org/10.1016/0378-7788\(94\)00912-4](https://doi.org/10.1016/0378-7788(94)00912-4)
- [145] Wolsink, M. (2012). The research agenda on social acceptance of distributed generation in smart grids: Renewable as common pool resources. *Renewable and Sustainable Energy Reviews*, 16(1), 822-835. <https://doi.org/10.1016/j.rser.2011.09.006>
- [146] Wolsink, M. (2014). *Distributed Generation Of Sustainable Energy As A Common Pool Resource: Social Acceptance In Rural Settings Of Smart (Micro-)Grid Configurations*. 17.
- [147] World Health Organisation. (2018). *COP24 - Special Report Health and Climate Change* (p. 74). <https://apps.who.int/iris/bitstream/handle/10665/276405/9789241514972-eng.pdf>
- [148] Wu, X., Gao, Y., & Jiao, D. (2019). Multi-Label Classification Based on Random Forest Algorithm for Non-Intrusive Load Monitoring System. *Processes*, 7(6), 337. <https://doi.org/10.3390/pr7060337>
- [149] Yoeli, E., Budescu, D. V., Carrico, A. R., Delmas, M. A., DeShazo, J. R., Ferraro, P. J., Forster, H. A., Kunreuther, H., Larrick, R. P., Lubell, M., Markowitz, E. M., Tonn, B., Vandenbergh, M. P., & Weber, E. U. (2017). Behavioral science tools to strengthen energy & environmental policy. *Behavioral Science & Policy*, 3(1), 68-79. <https://doi.org/10.1353/bsp.2017.0006>

- [150] Zeifman, M. (2012). Disaggregation of home energy display data using probabilistic approach. *IEEE Transactions on Consumer Electronics*, 58(1), 23-31. <https://doi.org/10.1109/TCE.2012.6170051>
- [151] Zhao, B., Stankovic, L., & Stankovic, V. (2015). Blind non-intrusive appliance load monitoring using graph-based signal processing. *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 68-72. <https://doi.org/10.1109/GlobalSIP.2015.7418158>
- [152] Zhao, B., Stankovic, L., & Stankovic, V. (2016). On a Training-Less Solution for Non-Intrusive Appliance Load Monitoring Using Graph Signal Processing. *IEEE Access*, 4, 1784-1799. <https://doi.org/10.1109/ACCESS.2016.2557460>
- [153] Zhong, M., Goddard, N., & Sutton, C. (2014). Interleaved Factorial Non-Homogeneous Hidden Markov Models for Energy Disaggregation. *arXiv:1406.7665 [stat]*. <http://arxiv.org/abs/1406.7665>
- [154] Zhuang, M., Shahidehpour, M., & Li, Z. (2018). An Overview of Non-Intrusive Load Monitoring: Approaches, Business Applications, and Challenges. *2018 International Conference on Power System Technology (POWERCON)*, 4291-4299. <https://doi.org/10.1109/POWERCON.2018.8601534>

RÉSUMÉ

La transition énergétique vers les sources d'énergie renouvelables est essentielle pour atténuer le changement climatique. Mais, l'énergie renouvelable est intermittente en nature, ce qui peut entraîner à l'avenir une congestion du réseau lors des pics de consommation. Par conséquent, la flexibilité d'énergie est inévitable en complément de la gestion de l'offre et du stockage de l'énergie. Le secteur résidentiel est le plus gros consommateur d'énergie et il est donc essentiel de mettre en œuvre la flexibilité d'énergie. En complément des approches existantes, nous présentons dans cette thèse une flexibilité "indirecte" non monétaire par la mise en œuvre d'une expérience de science comportementale avec 175 ménages français. L'ensemble des "nudges" utilisés dans cette expérience n'oblige pas les participants à mettre en œuvre la flexibilité énergétique pour des raisons de gain monétaire ou d'aversion aux pertes.

Les alertes sont générées par des algorithmes prédictifs et envoyées par SMS à un groupe de ménages dans le but d'effectuer soit un transfert de charge, soit un effacement d'énergie. Les alertes sont complétées par des suggestions sur l'utilisation (ou non) d'appareils engagés pour mettre en œuvre la flexibilité d'énergie. Après chaque alerte, le groupe reçoit un retour graphique dans lequel la courbe de charge mesurée par le compteur intelligent LINKY est superposée à une courbe de référence théorique propre à chaque ménage. En comparaison avec un groupe de contrôle, nous pouvons quantifier l'impact de ces signaux de "nudge" sur la consommation d'énergie du groupe traité. Enfin, un algorithme de désagrégation énergétique est présenté qui nous permettra de détecter les équipements utilisés lors de la flexibilité et ainsi faire une analyse de l'engagement des ménages à rechercher des appareils potentiels pour la flexibilité énergétique.