Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus Ciudad de México

A framework to create bottom-up energy models that support policy-design and decision-making of electricity end-use efficiency: A case study in residential buildings and the residential sector

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Engineering Science

Author: Marlene Ofelia Sanchez Escobar

Advisors:	Juana Julieta Noguez Monroy	Ph.D.	
	Jose Martin Molina Espinosa	Ph.D.	
	Rafael Lozano Espinosa	Ph.D.	
Thesis Committee:	Sergio Ruiz Loza	Ph.D.	President
	Jorge Carlos Valverde Rebaza	Ph.D.	Secretary
	Abel Matus Castillejos	Ph.D.	Vocal
	Juana Julieta Noguez Monroy	Ph.D.	Vocal
	José Martín Molina Espinosa	Ph.D.	Vocal

Tlalpan, Ciudad de México
6th December, 2022

“Conservation may be a sign of personal virtue, but it is not a sufficient basis for a sound, comprehensive energy policy.”

Dick Cheney

Abstract

A framework to create bottom-up energy models that support policy-design and decision-making of electricity end-use efficiency: A case study in residential buildings and the residential sector

by Marlene Ofelia SANCHEZ ESCOBAR

In this work we present a framework that guides the creation of bottom-up energy models (BUEMs) that aim to support policy design of electricity end-use efficiency. Energy models are decision-making tools for policy-makers and they are key tools to evaluate decisions. However, research reveals that bottom-up energy models are empirically created and they are not designed to guarantee support towards policy design. Likewise, the use of scenarios has not been applied as a standard technique within BUEMs. Thus, is it possible to align these models to specific policy goals and standardize the use of scenarios in them through the application of a process?

The framework proposed in this work includes phases, processes, and artifacts that conduct the modeler through the model's construction process. The processes incentive best practices for policy design for the residential sector and residential buildings. To build the proposed framework, we first research the characteristics of BUEMs and their opportunity areas. Then, we propose processes that are important for the model's design but we consider the application of best practices to overcome the problems encountered in the models. After that, we execute the framework and record all the events for future analysis. To finish, we assess the process execution quantitatively and qualitatively using the process mining technique.

The results of the framework's application are promising by outperforming other methodologies in the literature. In fact, it has been proved that model's creation time can be diminished with the application of the framework. Likewise, the framework promotes: (1) policy alignment from the start to the end of the model's development process and (2) the definition of scope boundaries. Moreover, it brings transparency to the process by the use of the proposed templates.

On the other hand, the utilization of the process mining technique to create a process model has brought advantages as well. For instance, (1) it is possible to monitor, control, and enhance the model's construction process and (2) The processes compliance can be evaluated and adaptations can be recommended with quantitative evidence.

Acknowledgements

I want to thank my advisors Julieta Noguez, Martin Molina, and Rafael Lozano for their support and commitment along the PhD program. Without your guidance this work could not be possible. Likewise, I want to thank Sergio Ruiz Loza, Jorge Carlos Valverde, and Abel Matus Castillejos for their valuable comments to this document and for their important recommendations provided. It´s been an honor to have so great Thesis Committee. Moreover, I want to thank my PhD Program Director Rogelio Bustamante and the academic and administrative staff of Tecnologico de Monterrey for their trust and their support during this endeavor.

I also want to thank my family and my partner, which gave the strength and courage to finish this work. Thank you dad for teaching me how to live a life with courage and humility. Thank you mom for showing me always your strength and the female power by example. Thank you Luisita for supporting me in hard times. Thank you Juanito for always telling me the truth even when it hurts. Thank you Samantha for helping me to find my spiritual path, and finally thank you Sergio for showing me by example how to give love and how to accept yourself for good.

Contents

List of Figures

List of Tables

I dedicate this work to my parents who taught me all they know about this world. Thank you for never given up on me.

Chapter 1

Introduction

Policy-makers (PMs) worldwide are concerned about the increasing of energy consumption and CO2 emissions in their countries, considering the negative impacts on global warming, climate change, and energy security. For instance, in 2019 the world energy outlook anticipates that the building sector (including households and services) will continue being the main contributor to global electricity demand by 2040 [1], due to a possible increment in the electricity use of air conditioners, household appliances, and electric vehicles. Given this problematic, PMs need tools to evaluate diverse strategies with the goal of improving the efficiency of electricity use [2]; and consequently reduce emissions without compromising the development of electricity services. The design of policies is an approach used by governments to promote electricity end-use efficiency [2].

1.1 Motivation

Energy-efficiency policies provide useful strategies for energy sectors without requiring an enormous investment to apply them. Yet, the design of policies rely on the creation of models, which are considered tools to support decision-making and policy selection. These last are categorized as top-down or bottom-up based on their level of data aggregation, details of technology representation, and degree of endogenous behavior. However, for the modeling of electricity end-use, research reveals that bottom-up energy models (BUEMs) provide limited support on policy design specially for residential buildings; since models have been implemented without considering documented best practices.

1.2 Proposal

In this document, we proposed a framework to improve the support that bottom-up energy models provide to policy design of electricity end-use efficiency. The framework aims to consider best practices for model creation. For instance, the framework promotes:

1. The inclusion of key energy efficiency metrics,
2. The implementation of scenario to support decision-making,
3. The utilization of a portfolio of design techniques that facilitate their selection.

The inclusion of best practices in model design is relevant to policy makers, since models are useful components to support decision-making (through scenarios) and therefore they can reduce the risk of implementing a policy that will not provide results or that will not give a return of investment in a time period. Likewise, it can be mitigated the cost of implementation of a policy that could not work out. We believe that this study propose a new way of addressing the design and quality of bottom-up models, which is relevant for model designers as well.

1.3 Hypotheses

The following hypotheses will be evaluated in the study:

1. The creation process of bottom-up energy models can be measured, monitored and managed to achieve process compliance and policy alignment,
2. The creation process of bottom-up energy models can be adapted to guide the design of energy efficiency scenarios that support decision-making.

1.4 Objectives

The specific objectives of this research are:

1. To create a mechanism to measure, monitor, and improved the creation process of bottom-up energy models,
2. To guide the design of scenarios within bottom-up energy models.

1.5 Scope and limitations of the study

In the research conducted, we focus on BUEMs that aim to be useful to design policies at the end-use level considering granular data sets and avoiding general assumptions present in top-down models. We explore the case of data-driven models created on Decision-Making Centers (i.e., DMCs), which are environments that facilitate stakeholders' decision-making processes using predictive models and diverse what-if scenarios [3]. We select DMCs by the positive impact on policy design that they provide. Furthermore, we focus exclusively on some phases of the model creation process in order to promote the support of BUEMs towards policy design of electricity end-use efficiency.

Chapter 2

Theoretical Framework

2.1 Energy Policy and Policy Design

Energy policy focus on energy supply security and affordability, competitiveness in the energy market [4], and limited impact to the environment [5]. Governments aim to select policies that improve the economic and social welfare of people and that help to build a "stronger, cleaner and fairer world" [6]. Thus, the policy life cycle shown in Figure 2.1 assists the development of energy policies in OECD[1] countries.

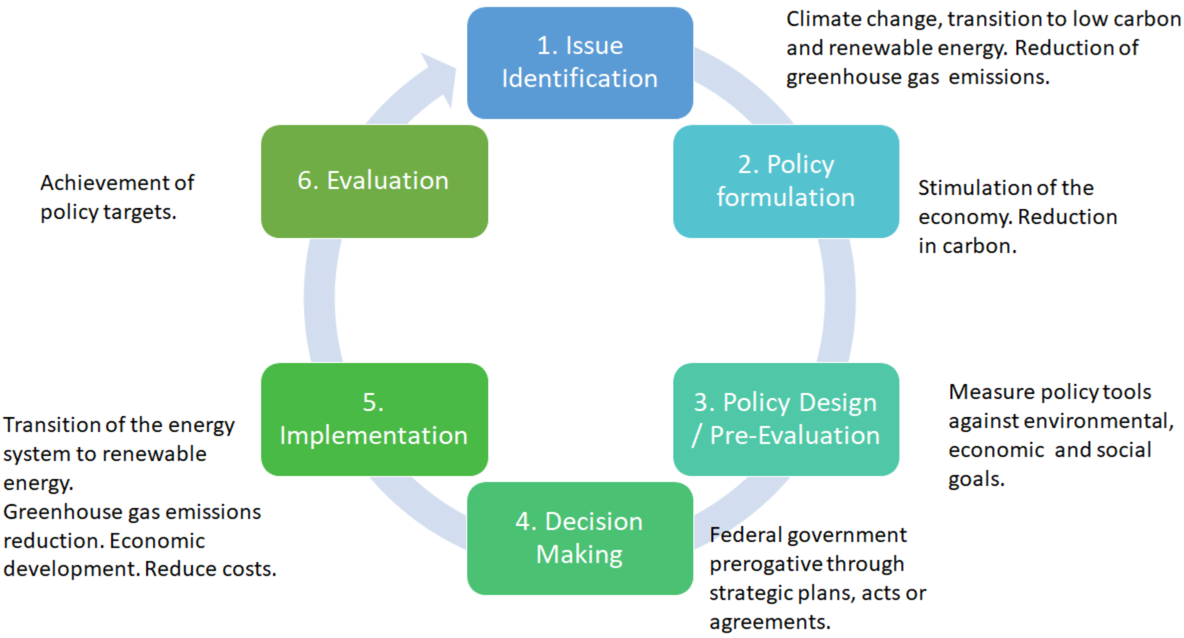


Figure 2.1: Energy Policy Life Cycle (Adapted from [6])

[1]Organization for Economic Co-operation and Development

In this research, we focus on the policy design phase, since it involves the pre-evaluation of technologies and economic approaches; considering diverse course of actions and allowing to achieve goals in a sustainable form [6].

The term policy design refers to the process of formulating the policy output [7] (i.e., policy goals, instruments, and settings in which policies are applied [8]). According to Varone et al. and Haelg et al., the process can also be focused on the selection of policy instruments[2] [10], on policy mix (i.e., combination of instruments), and on fine-grained instrument analyses [7]. The policy output creation can be performed with different levels of abstraction. aims, and means) [7]; as can be regarded in Table 2.1.

Table 2.1: Policy design - Levels of abstraction

Level of abstraction	**Policy aims**	**Policy means**
High-level	General policy goals	General policy instrument logics
Mid-level	Policy objectives	Policy instrument types
Low-level	Policy settings	Policy instrument calibrations

In the case of the promotion of energy efficiency, the literature reveals a vast list of instruments [10] to address specific policy goals (designed at the high and medium level of abstraction). For that reason, in this work, we do not focus on the creation of new instruments but on their selection using bottom-up energy models as support tool.

The selection of policy instruments considers the following four dimensions [10]: (1) degree of coercion, (2) resource intensiveness, (3) political risk, and (4) targeting. The first represents the limitations of ideas and financial resources experienced by the role of government or private market. The second contemplates the administrative operating costs to implement and evaluate specific instruments. The third refers to the possible visibility of the policy's lack of success. Finally, the fourth dimension focus on the precise selection of target recipients for policy instruments. Table 2.2 shows examples of each dimension.

Table 2.2: Dimensions for instruments selection and examples

Policy selection dimension	**Example**
Degree of coercion	A president tends to adopt instruments depending on their political affiliation (e.g., democrat or republican)
Resource intensiveness	The government creates an Energy Conservation Committee to implement diverse energy instruments
Political risk	Instruments are implemented but reverted without considering their experiences in other countries
Targeting	Electricity companies dispute certain instruments in contrast to the customer desire

[2]Policy instruments (or interventions) are approaches used to promote particular policies and reach specific goals [9]

The main concepts for the design of *electricity* end-use policies is explained in the following section.

2.2 Policy Design of Electricity End-use Efficiency

In the energy sector the study of energy efficiency and how it can be improved at the end-use level is a relevant research topic. Energy end-uses are part of the energy demand-side and include [11]: Appliances (A), Space Heating(SH), Space Cooling (SC), Lighting (L), Water Heating (WH), and Cooking (C). End-use energy efficiency focus on the efficient use of final energy in industry, services, agriculture, households, transportation, and other areas like buildings.

The policy design of energy efficiency focuses on the selection of policy instruments with the aims of achieving energy efficiency objectives [10], eliminate barriers toward efficiency, and gain energy efficiency benefits [12]. Examples of energy efficiency objectives are: reducing CO2 emissions, improve air quality and energy security [13], control energy consumption, and enhance energy efficiency levels[12]. The policy instruments shown in Table 2.3 are used in the design of energy efficiency policies [14].

On the other hand, the policy design of *electricity* end-use efficiency (which is an energy efficiency sub-field) focus on the factors that distort market and restrict the adoption of efficient technologies [2]. For instance, expected short payback period on investments, uncertainty about actual savings, the lack of trained people to invest in energy efficiency, physical barriers of the technology, attributes that affect performance, and unfavorable investment due to lower average usage of the product [2]. Electricity is relevant (as final energy) in the energy mix since its generation represents one-third of the total OECD emissions [15], and its contribution in the residential sector is considerable (greater than 40%) and continues to increase [16]. This last has made governments pay immediate attention to new forms of improving the efficiency of energy use [17].

In the next section, we define bottom-up energy models as tools to improve energy end-use efficiency, how they characterize, and finally which features are useful to support policy design.

TABLE 2.3: Energy Efficiency Policy Instruments.

Type of Instrument	**Instrument**	**Description**
Market-based	Energy Taxes	Impact the price of goods and services that generate high greenhouse emissions or the price of the emission itself [18].
	Tradable emissions permits	Limited emission permits are divided among companies that pollute to control the amount of emissions agreed by regulatory agencies [19].
	White certificates schemes	Energy suppliers commit to (1) Promote energy efficiency in final uses and (2) implement interventions to save a percentage of their distributed or supplied energy [20].
Financial incentives	Subsidies	Direct payments or tax rebates are used to motivate expenditure on energy efficiency [14].
	Access to capital measures	Grants and loans are provided to drive specific energy efficiency expenditures [14].
Regulatory Measures	Codes and Standards	Building codes and energy performance standards [14] are used to impose the compliance of minimum energy efficiency levels to products or services (e.g., building design or construction [21]).
Information and Feedback	Information	Certificates, labels, or audits are used to avoid suboptimal energy efficiency investments [14].
	Feedback	Consumption and energy information is given to consumers to promote energy conservation [22].
Non-regulatory measures	Voluntary Agreements	Adjustable agreements among firms and public authorities used to increase energy efficiency and diminish greenhouse emissions [23].

2.3 Bottom-up energy models and the modeling of end-use energy efficiency

Energy models allows the simulation of policy, technology alternatives, and their consequences in the demand and supply of energy [24]. These models are categorized as top-down and bottom-up based on their level of data aggregation. The first rely on aggregated data to perform analysis of sectors interaction; while the second use disaggregated data to analyze energy end-uses and technological choices [25]. Our research focus only on the latter and in the following sections we explain key concepts for this work.

Bottom-up energy models

Bottom-up energy models (BUEMs) are characterized by their sector coverage, geographical coverage, time horizon, methodologies, end-use energy modeling techniques, programming techniques, data time split, metrics, and residential electricity end-uses. Table 2.4 presents the complete taxonomy of bottom-up energy models.

Sector coverage: BUEMs can consider the impact of policies, technologies, or other factors in one or more sectors [26]. Examples of sectors that can be included in models are: economic sector, which comprises energy and production sectors, and demand sector that includes: households, buildings, industries, and transport sectors [27].

Geographical coverage: Models can represent at different geographical levels the electricity consumption and its efficiency (e.g., worldwide, international, sector, regional, or project level) [25].

Time Horizon: Time horizon in BUEMs is defined in time frames [30] (e.g., short, medium or long-term) and can represent the development or configuration of energy systems [25].

Methodology: The following methodological approaches are used to build BUEMs: economic, optimization, simulation, spreadsheet, back-casting, and multi-criteria [25]. A detailed description is provided in the next sections.

Economic energy models (EEMs) are oriented to evaluate how adequate and valuable is the implementation of policies to achieve specific objectives [33]. Likewise, EEMs aim to picture the economic and technical impacts of applying alternative economic strategies [34]. Common factors included in EEMs are technological innovation, efficiency improvement, and energy demand and supply. Additionally, Van Beeck [25] differentiates economic models' methodologies as econometric, macroeconomic, and economic equilibrium, however we only consider the general hierarchy (e.g., economic) in the classification. Examples of economic energy models include [34]: MIS, MEPA, MARKAL, ETSAP, MESSAGE III, EFOM-END, Conrad, MDM, Bovemberg-Goulder, Jorgenson-Wilcoxen, HERMES, MIDAS, MARKAL, PRIMES, ESCAPE, GEM-E3, LEAN, QUEST, WARM, E3 ME, DICE, RICE, PRICE, SLICE, CETA, IEA, IMAGE, FUND, PAGE, MERGE, ERIS, IIAM, ICAM, MINICAM, ERM, EPPA, SGM, GREEN, G-CUBED, Whalley-Wigle, WIAGEM, WORLSCAN, POLES, RAINS-Europe and RAINS-Asia.

Table 2.4: Taxonomy of Bottom-up Energy models.

Category	**Subcategory**	**Model Focus**
Sector Coverage [25]	Single-Sector Multi-Sector	Just one sector Interaction between sectors
Geographical Coverage [25]	Global Regional National Local Project	World economy/situation International regions All sector within a country Regions within a country Specific energy project
Time Horizon [25]	Short Medium Long-Term	Less than 5 years model 5 to 15 years model Greater than 16 years model
Methodology [25]	Economic	Representation of economic and technical effects of alternative economic strategies
	Optimization	Optimization of choices on energy investment
	Simulation	Replication of a system operation in a simplified form
	Spreadsheet	Utilization of a flexible tool to generate customized energy models
	Back-casting	Creation of views of a desired future and identification of trends to be broken to achieve the future
	Multi-criteria	Inclusion of additional criteria to the model beyond economic efficiency
	Other	Other methodology
End-use Energy Modeling Technique [28, 29]	Engineering	Calculation of energy consumption based on thermodynamics and heat transfer of all end-uses
	Data-driven statistical	Correlation of end-use features with its energy use using statistical techniques
	Data-driven AI-based	Correlation of end-use features with its energy use using artificial intelligence techniques
Programming Technique [30]	Linear Programming (LP)	Discover arrangement of activities to minimize or maximize a defined criterion
	Mixed Integer LP	Extension to LP programming which include detailed formulation of technical properties and relations in modeling of energy systems
	Dynamic	Discover optimal growth path through division of an original problem and optimization of sub-problems
	Heuristic	Manage high dimension optimization problems [31]
	Other	Other type of programming technique
Data Time Split [32]	Hourly/Minute Daily Monthly Yearly	Hourly/Minute data resolution Daily data resolution Monthly data resolution Yearly data resolution
Metrics and Tools [25]	Metrics Tools	CO_2 emissions and cost as outputs in the model Scenario utilization to show model's results
Residential Electricity end-uses [11]	A, SH, SC, L, WH, C	Detailed identification of electricity consumption, energy use and energy savings by end-use.

Note: Electricity end-uses: AL = Appliances and Lighting, SC = Space Cooling, SH = Space Heating, WH = Water Heating, A = Appliances, L = Lighting, C = Cooking.

Optimization models (OMs) are used to obtain optimal investment strategies [25], prospective of energy systems [25], and optimal energy system operation. According to Ringkjøb et al. [27] and Van Beeck [25], it is well-known the utilization of linear programming and mixed-integer linear programming techniques to implement OMs.

Simulation models (SMs) are useful to reproduce the simplified operation of a system [25]. From the energy-economic perspective, these models aim to replicate end-user behavior for technology choice, considering diverse drivers [33]. Examples of simulation models are WEM, MURE, and NEMS-RSDM. On the other hand, SMs estimate energy consumption using simulation techniques, thermodynamics principles, urban construction, and climate data [28].

Spreadsheets models are described by [25] as highly flexible models that use to contain a personalized reference model. An example of this kind of model is the Cost of Renewable Energy Spreadsheet Tool (i.e., CREST), explained in detail by Gifford and Grace in [35].

Back-casting models are oriented to create views of desirable futures [25], explore how these futures can be achieved [36], and reveal related policy implications (e.g., social, economic, political, and technology) [36]. The back-casting methodology is action-oriented since policy-makers can decide a course of action to achieve a specific target considering specific policy implications [37].

Multi-criteria models consider the inclusion of additional criteria to the model beyond economic efficiency [25]. A study that includes qualitative and quantitative analysis is an example of these kinds of BUEMs.

End-use Energy Modeling Technique: BUEMs can also be classified based on the energy modeling techniques used to build them as: enginnering and data-driven energy models [29][28].

Engineering energy models (ENEMs) employ physics (e.g., thermodynamics and heat transfer) examination of a dwelling's end-uses to calculate its energy consumption [29]. According to [38], these kinds of BUEMs allows generating a base energy estimation without historical data and validating energy-saving measures at the end-use level. On the other hand, the identification of ENEMs depends on the techniques used to create the model [29]; for instance, Firth et al. [39] label ENEMs that use the household or building archetype approach as household stock models or building stock models [38]. Examples of the first are BREHOMES, UK Domestic Carbon Model, DeCarb, and the Johnston model [39]. Finally, ENEMs can also use appliance ownership distributions or household sample data approaches for the model creation [29].

Data-driven Energy models use samples of large sizes of operational energy datasets and apply statistical or artificial intelligence techniques to create mathematical associations among energy end-uses and their use [28]. Our merged taxonomy includes the following types of DDEMs: data-driven statistical and data-driven AI-based. See Table 2.4 to review the focus of each type of model.

Programming Technique: The following types of programming techniques are used to build BUEMs: linear, mixed linear, dynamic, heuristic, and other types [25]. Table 2.4 includes the focus of each technique.

Data Time Split: The resolution of the model's input data is specified with the following types of time splits: hourly or minute, weekly, monthly, and yearly. [32]

Metrics and Tools: The following indicators are identified as necessary in the design of models than aim to support energy efficiency policy design: CO2 emissions, cost [40]. Likewise, the inclusion of scenarios as tools are also recommended to support policy making [41].

Residential Energy End-uses Finally, BUEMs taxonomy identifies the residential energy end-uses that can be implemented within models (e.g., appliances, space heating, space cooling , lighting, water heating, and cooking [11]).

In the following section, we explain how bottom-up energy models are built, in which environments, and if formal processes are used to build them.

2.4 Methodologies, processes and environments to create BUEMs

The understanding of BUEMs' creation is relevant to identify the causes behind the lack of policy support that these models present. In this study, we affirm that a model that fails to provide the expected policy support; it does not have an acceptable quality.

In [42], Zahran affirms that a product quality is ruled by the quality of the process employed to build and maintain it. Therefore, in this section we explain the processes, environments, and methodologies used to build BUEMs; so we can improve them and consequently enhance the models created.

Decision Making Centers and Processes

Decision Making Centers (DMCs) are immersive virtual environments that allow the visualization of predictive and scenario-based models, the comprehension of complex problems and the simplification of decision-making [43]. The visualization of information and its analysis is performed in seven different screens allowing story telling and collaborative decision making. DMCs apply best practices of collaborative planning processes using models [44] and allow the representation of stories with scenarios that aim to show relationships of present, past, and future events [45]. Figure 2.2 shows an example of a DMC environment.

FIGURE 2.2: Decision Making Center at Tecnologico de Monterrey

The components of DMCs include [44]: (1) decision entities that contemplate decision-makers (2) decision support components that refers to decision-support tools like models, systems or other kind of analytical artifacts [46] (3) organizational systems which include facilitators, coordinators, technicians, and procedures (4) the DMC layout which describes the size and shape of the environment rooms and display screens, and the positioning of tables and chairs, and finally (5) the DMC's technologies that contemplate the hardware and software used to support all the other DMC components. Our research focus in the models that support decision-making, specifically in the process to build them and on their implementation of scenarios.

Scenarios

Scenarios allow the evaluation of the feasibility of specific course of actions and the assessment of long-term results and effects, using cause-effect stories to forecast outcomes [47]. For instance, the change of end-use energy efficiency in residential buildings based on the application of one or more energy efficiency policies can be modeled with scenarios.

Scenario construction should consider relevant external factors, levels of uncertainty, the creation of multiple stories to identify causally related results and intrusive events, and the impact analysis of that stories on strategic planning and decision making [48]. The evaluation of scenarios should consider the totality and truthfulness of the causal relationships present in the scenario. Finally, scenarios can be evaluated using the following guidelines [48]:

1. Identification of factors, states, conditions
2. Evidence should support assumption evidences.
3. Predictions should be reasonable and based on changing and related relationships (e.g., factors, states, and conditions)
4. Evidence should support impacts among factors
5. Reasonable events, action, consequences, probabilities and impacts
6. Inclusion of an influence diagram or causal map

DMC Methodology for decision-making and model creation

The model's creation process is defined as a phase of the DMC's methodology [49], which is detailed on Figure 2.3.

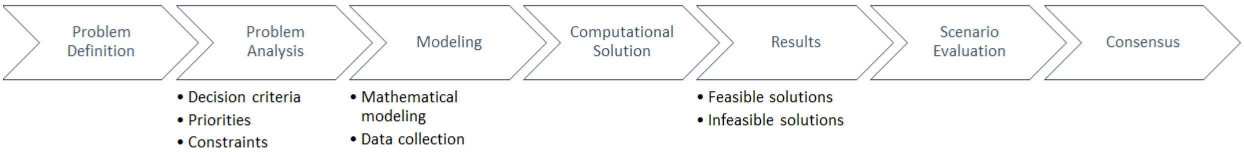


FIGURE 2.3: Decision-making center Methodology (Adapted from [49])

In the DMC's methodology, the problem definition phase contemplates the structuring of the issue to be solved specifying input and output conditions. The second stage or problem analysis allows the definition of the decision criteria, priorities, and constraints. The modeling phase, which is our principal interest, formalizes the problem using the collected data and mathematical models for future implementation (e.g. computational solution). Finally, the model's results are used to classify feasible or infeasible alternatives that will be presented, validated and analyzed in the DMC environment. The last phase represents the agreement on the selection of an alternative in the decision-making process.

The methodology is considered to have a low level of detail in the specification of tasks in comparison to other model-creation methodologies like Cross-Industry Standard Process (i.e., CRISP-DM) [46].

CRISP-DM for data-driven energy models

A useful methodology to create BUEMs is CRISP-DM, however this last is focused exclusively on the creation of data-driven energy models.

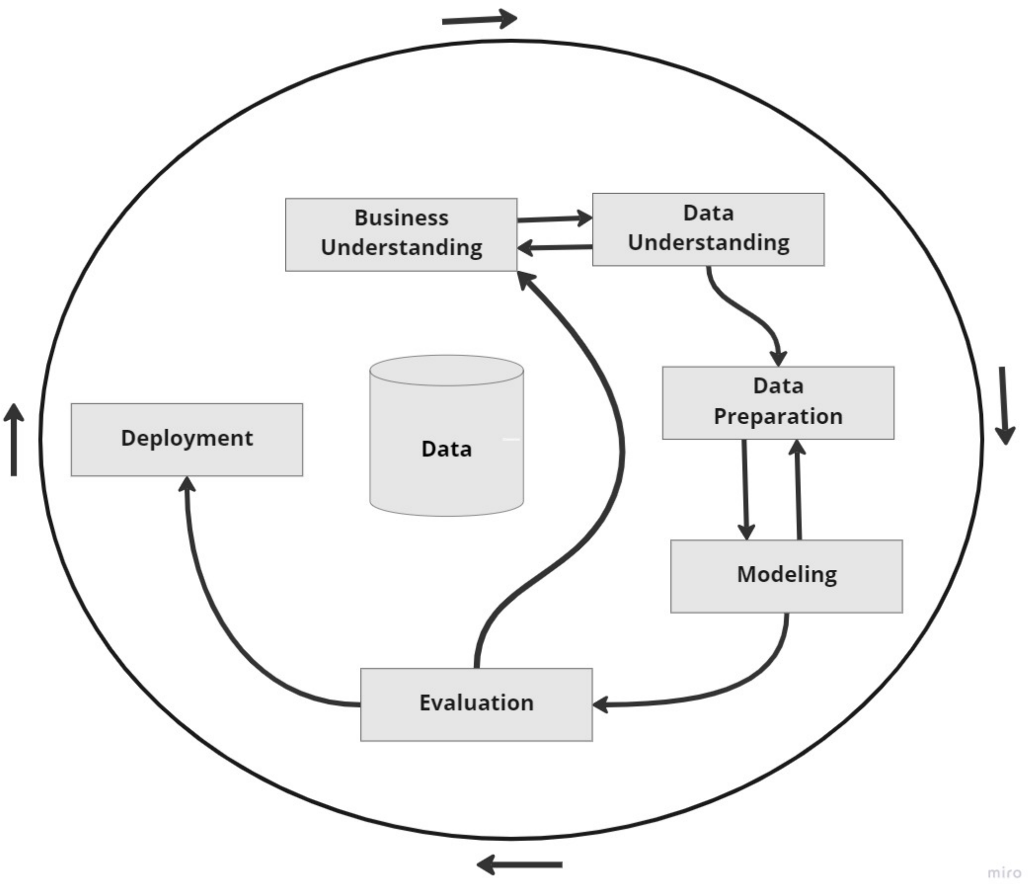


FIGURE 2.4: CRISP-DM Methodology (Adapted from [49])

The framework is divided in six phases that include generics tasks and proposed outputs. The complete framework is detailed on Table 2.5.

Table 2.5: CRISP-DM phases, tasks and outputs

Phase	**Task**	**Output**
Business Understanding	Determine Business Objectives	Background Business Objectives Business Success Criteria
	Assess Situation	Inventory Resources. Requirements, Assumptions, and Contraints Risks and Contingencies Terminology Costs and Benefits
	Determine Data Mining Goals	Data Mining Goals Data Mining Success Criteria
	Produce Project Plan	Project Plan Initial Assessment of Tools and Techniques
Data Understanding	Collect Initial Data Describe Data Explore Data Verify Data Quality	Initial Data Collection Report Data Description Report Data Exploration Report Data Quality Report
Data Preparation	Select Data Clean Data	Rationale for Inclusion/Exclusion Data Cleaning Report
	Construct Data	Derived Attributes Generated Records
	Integrate Data Format Data	Merged Data Reformatted Data
		Dataset Dataset Description
Modeling	Select Modeling Techniques	Modeling Technique Modeling Assumptions
	Generate Test Design	Test Design
	Build Model	Parameter Settings Models Model Descriptions
	Assess Model	Model Assessment Revised Parameter Settings
Evaluation	Evaluate Results	Assessment of Data Mining Results Approved Models
	Review Process	Review of Process
	Determine Next Steps	List of Possible Actions Decision
Deployment	Plan Deployment Plan Monitoring and Maintenance	Deployment Plan Monitoring and Maintenance Plan
	Produce Final Report	Final Report Final Presentation
	Review Project	Experience Documentation

Business Understanding (BU)

The BU phase contemplates the definition of model's business objectives, the assessing of the current environment to build the model, the data mining goals to be accomplished, and finally the creation of a project plan for model creation [50]. To align objectives, all stakeholders related to the problem should be included in order to define and agree a complete list of business objectives, metrics, and success criteria. All these expect to be specific and measurable [51]. The assessment of the environment involves performing a resource inventory, so it can be identified the available personnel, computing resources, and software to design the model [50]. The assessment task involves as well, the identification of requirements, terminology, risks, constraints, and assumptions for the project execution, and the perform of a cost-benefit analysis.

Likewise, the goals of the data mining modeling effort must be defined, so the model's build and deploy are aligned to the business needs and operation. Examples of possible goals are: to deploy and to build of a model that deliver business value, to create of an easy access database for model data, or to develop a knowledge base of learning models [51]. The previous goals should be implemented by defining specific objectives and tasks, which later are included in a plan than can be monitored during project execution [50]. Finally, the result of the tool evaluation should be considered during the project planning phase.

Data Understanding (DU)

The DU phase comprises the acquisition, description, exploration, and quality verification of data [50]. In this stage it has to be identified sources of data, which can be quite different (e.g., records on databases, spreadsheets, files, and paper [51]). A report of the obtained datasets, their locations, access methods, and acquisition problems should be documented. Moreover, it is relevant to describe the obtained data (i.e., format, quantity, field definition and type), thus the useful variables are used for data analysis [50].

Once the described data is available, it is possible their examination using statistical analyses to answer the data mining questions or to validate hypotheses [50]. Finally, the quality of data is validated (e.g., completeness, missing values, outliers, errors) and if quality problems are present, they are reported and possible solutions are proposed.

Data Preparation (DP)

DP phase is necessary to perform analytical modeling, since data possibly needs some kind of transformation before it can be used by modeling algorithms [51]. For instance, the selection of specific attributes or rows, the resolution of data quality issues, the merge of two or more data sources, or the creation of new attributes could be necessary for some statistical analyses. Lastly, some modeling tools require the formatting of data (e.g., removing commas, comma-delimitation, data sorting, trimming values) to be able to process datasets [50].

Modeling

The Modeling phase encompasses the selection of a one or more modeling techniques, the generation of test designs, the construction of models, and the model's evaluation [50]. Modeling techniques should be selected based on their assumptions (e.g., attributes with specific probability distribution or allowance or not of missing values) and the characteristics of the available datasets. Next, test designs are defined to validate the quality and validity of the model. Usually for this task, datasets are separated to perform the model's training and testing. Afterward, the models are built and their parameters are calibrated and documented. Finally, the model is evaluated considering their applicability to the domain knowledge, the defined success criteria, and the test design. Qualities of the models (e.g., accurancy) and rank among them are also summarized. [50].

Evaluation

The evaluation phase is oriented validate the achievement of business objectives. In this step, a final assessment of models is performed based on business success criteria. The ones that satisfy the criteria are approved. Moreover, a final evaluation of the process to build the model can be done. Future strategies regarding the project's resources and budget can also be proposed [50].

Deployment

The final stage of the framework defines the model's deployment and maintenance strategies, and recommends to report and review the results of the project and the processes.

In the following chapter, we provide a complete picture of the current research regarding bottom-up energy models that aim to support policy design of electricity end-use efficiency. We enclose the analysis considering only models for the residential sector and residential buildings, given their present and expected future participation in the global electricity consumption.

Chapter 3

State-of-the-art

The creation of BUEMs that support policy design of electricity end-use efficiency is a study field that requires to be understood and analyzed; since these kinds of models have exhibited defects in their support towards energy demand and policy assessment in the residential arena [33]. However, minimal interest in the study of models' development processes have been identified from the literature, which also limits the understanding of the problematic.

In this chapter, we present a complete study of BUEMs oriented to support policy design of electricity end-use efficiency for the residential sector and residential buildings, so we can have an understandings on how these models are built and which specific design problems needs to be addressed.

3.1 Bottom-up energy models and modeling of policy design of electricity end-use efficiency for the residential sector and residential buildings

This section examines, analyzes, summarizes, and expands the systematic review developed by the author, named "The Contribution of Bottom-Up Energy Models to Support Policy Design of Electricity End-Use Efficiency for Residential Buildings and the Residential Sector: A Systematic Review" [52]. In this work, we examine bottom-up energy models and their contribution to policy design of electricity end-use efficiency considering their features, their support towards specific policy instruments, and the techniques used to build them.

The five year period research from 2015 to 2020, reveals a limited number of publications oriented to the topic (i.e., 20 studies), out of which 70% are useful for the residential sector and only 30% are valuable for residential buildings. In the following sections, we present the analysis performed for both consumption sectors, along with the study's findings and results [52].

3.1.1 BUEMs for the Residential Sector

For the residential sector, we consider models that encompasses the electricity demand of households and their end-uses (e.g., appliances, lighting, space cooling, water heating, and cooking). In these BUEMs, it is commonly modeled: (1) the replacement and market penetration of appliances and bulbs, (2) the factors influencing energy consumption and CO2 emissions, (3) the estimation of energy efficiency improvement programs, and (4) the consumer´s behavior towards energy efficiency initiatives [52]. The complete categorization of these models is available on Table 3.3.

TABLE 3.1: Caracteristics of BUEMs - Residential Sector.

Citation	Sector Coverage	Geographic Coverage	Time Horizon	Electricity End-Uses	Data	Cost	Scenario Based	CO_2
[53]	Single-sector	Local	Short	AL	Yearly	no	no	no
[54]				A, L, SC	Hourly	yes	yes	no
[55]	Single-sector	Local	Medium	SC	Yearly	no	no	no
[56]				AL	Yearly	yes	no	no
[57]				SC	Yearly	no	no	yes
[58]				A, SC	Yearly	no	no	yes
[59]	Single-sector	Local	Long-term	A	Yearly	yes	yes	no
[60]				L	Yearly	yes	yes	no
[61]				A, SC	Yearly	no	no	yes
[62]	Single-sector	Project	Short	A, SC	Hourly	yes	yes	no
[63]				A	Hourly	no	no	no
[64]				A	Hourly	no	yes	no
[65]	Multi-sector	Local	Long-term	A, L, SC, WH	Yearly	no	yes	no
[66]	Multi-sector	National	Long-term	C	Yearly	yes	yes	yes

Note: Electricity end-uses: AL = Appliances and Lighting, SC = Space Cooling, SH = Space Heating, WH = Water Heating, A = Appliances, L = Lighting, C = Cooking.

The previous data shows that BUEMs for the residential sector are mainly single-sector oriented and rarely multi-sector. Single-sector models picture geographical regions within a country in all types of time horizons. While, multi-sector models illustrate all sectors within a country or specific regions with exclusively long-term time horizons. This last revealing a correlation between the geographical coverage and the model's time horizon in BUEMs design. Lastly, for the application of electricity end-uses, residential models show mostly the utilization of appliances, lighting, and space cooling end-uses; and barely the inclusion of water heating, cooking, and space heating.

Second, in terms of the models' input data, the analysis reveals that in general yearly data is used to create the models. Except in BUEMs with short time horizons that show a consistently use of hourly data. This last can be related not only to the scope of the model to be created, but also to the availability of input's data.

With respect to the metrics and tools provided by BUEMs, research shows that these models do not consistently provide data about cost and CO2 emissions or present data with scenarios; which can affect negatively on the decision-making and the policy design processes.

In the next subsection, we explain the techniques used to build BUEMs for the residential sector.

Analytical techniques and methods The construction of residential BUEMs relies on data-driven statistical (71%) and hybrid data-driven (AI-based + statistical) (29%) modeling techniques (See Table 3.2). In the first case, statistical techniques in combination with other methodologies and programming methods are applied to analyze energy end-uses. The economic methodology is applied in the highest proportion of models, and implemented using: discrete choice models, diffusion models, end-use models, and statistical approaches like: linear regression and probability distributions. Differently, models built with the optimization methodology include mostly linear programming techniques and statistical approaches (e.g., probability distributions, and least square methods). Lastly, non-classified methodologies are employed in the lowest proportion for BUEMs's construction and with a limited application of design techniques.

On the other hand, hybrid data-driven models are designed using non-categorized methodologies and heuristics programming techniques. For the first case, statistical approaches (e.g., regression, mediation model), analyses techniques (e.g., time series and clustering), and artificial intelligence (e.g., principal component analysis) are utilized for the model's design. In the heuristic case, a combination of analyses (e.g., bi-variate correlation, grey relational), algorithms (e.g., butterfly optimization, chicken swarm optimization), and artificial intelligence techniques (e.g., support vector machine) are implemented. Although, we can observe a great set of techniques applied for hybrid models, not all of them were used during the five year research period. In fact, data-driven AI BUEMs appears in the map in the third year of the investigation, which reveals the limited utilization of AI techniques for the creation of BUEMs.

In the next subsection, we explain the policy instruments supported by BUEMs for the residential sector.

Table 3.2: Techniques and Methodologies used in bottom-up energy models that support energy efficiency policy design—Residential Sector.

Modeling Technique	**Methodology**	**Programming Technique**	**Techniques Used**	**Citations**
Data-driven statistical	Economic	Other	Discrete choice models, Time-series analysis	[53]
			Discrete choice models	[55]
			Econometric diffusion modeling, Market share functions	[59]
			End-use model, linear regression, scenario analyses	[65]
			Material flow analysis, weibull distribution, techno-economic analysis	[60]
	Optimization	Other	Probability density functions, least square method, pearson distribution	[54]
		Linear Programming	Linear optimization	[66]
		Mixed LP	Set of sequential uninterruptible energy phases, MILP	[62]
	Other	Other	Panel data regression	[57]
			Sliding window linear regression, kernel density estimation	[64]
Data-driven statistical and data-driven AI-based (Hybrid)	Other	Other	Cluster analysis, regression analysis, two level time series, mediation model, regression analysis, principal component analysis	[56, 63]
		Heuristic	Bivariate correlation analysis, Butterfly optimization algorithm, Least square support vector machine, Grey relational analysis, Chicken swarm optimization, Support Vector Machine	[58, 61]

Policy instruments BUEMs for the residential sector support the following types of policy instruments: information and feedback (39.13%), financial (30.43%), market-based (21.74%), and in less proportion regulatory (8.70%). See Figure 3.1.

Moreover, the graph shows the most used instruments per category which include: information (21.74%), subsidies (21.74%), and energy taxes (21.74%); followed by feedback (17.39%), codes and standards (8.70%), and access to capital measures (8.70%).

We can understand this distribution by considering the current capability of utility companies to register, monitor,and report users' energy consumption. With key consumption information, final users can compare their current and their historic energy use and can make decisions oriented to decrease electricity consumption by different means. Besides, the utilization of subsidies for technology replacement is still being a dominant policy instrument, however given the related implementation costs; it reveals to be less popular than the first one. Lastly, the

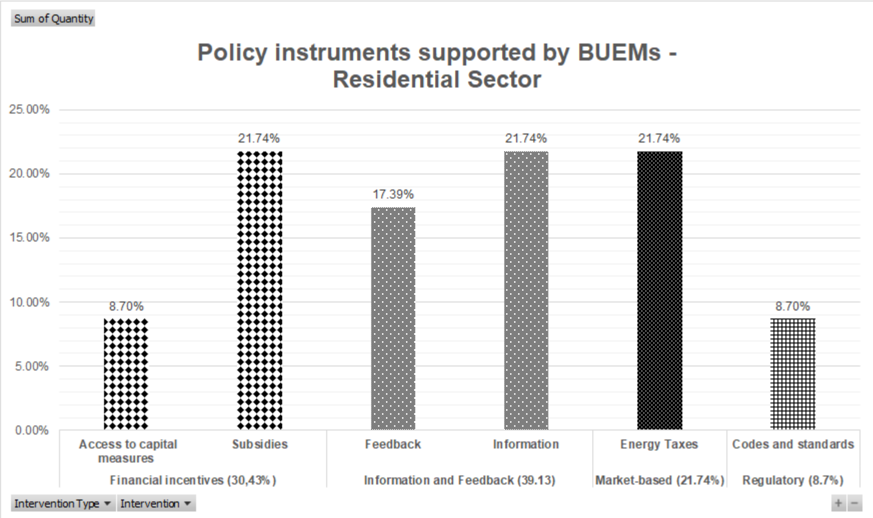


FIGURE 3.1: Policy instruments supported by BUEMS for the Residential Sector

adjust of energy prices is also used in this kind of models, since it can stimulate in a fast way the diminishing of energy consumption.

On the other hand, Table 3.3 reveals that BUEMs for the residential sector support general policy goals and specific policy instruments, however they are limited used for policy calibration. In fact, only 28% of the models can be used to design policies at the three levels of policy abstraction described on the previous chapter. This last means that in the majority of cases, these models can only support the design of general policy goals or specific policy instruments.

Besides, in terms of the capability of these models to compare different instruments. Research reveals that only 35% of them can be used to compare two instruments while only 14% can be used to compare three of them.

Finally, the results show that there is not standardization in the use of scenarios or in the utilization of key metrics like CO2 and cost. Likewise, none of the models present a mechanism for scenario configuration based on important variables for decision-makers.

TABLE 3.3: BUEMs' policy abstraction levels, type of instruments supported, scenario support, and metrics - Residential Sector.

		Policy abstraction Level			Policy instruments					Scenario and Metrics			
Citation	**Author/year**	**General Policy**	**Policy instruments**	**Policy calibration**	**Market-based**	**Info. and Feedback**	**Financial**	**Regulatory**	**Non-Regulatory**	**Scenario**	**Scenario Configuration**	**CO_2**	**Cost**
[53]	Jridi et al.	yes	yes	no	yes	no	yes	yes	no	no	no	no	no
[55]	Hara et al.	yes	yes	no	no	yes	no	no	no	no	no	no	no
[54]	Aghamohamadi and Amjady	yes	yes	yes	yes	no	yes	no	no	yes	no	no	yes
[56]	Lundgren and Schultzberg	yes	yes	no	yes	yes	no	no	no	no	no	no	yes
[57]	Meangbua et al.	yes	yes	no	no	yes	no	no	no	no	no	yes	no
[58]	Wen and Cao	yes	yes	no	no	yes	yes	no	no	no	no	yes	no
[62]	Mohseni et al.	yes	yes	yes	no	yes	no	no	no	yes	no	no	yes
[63]	Jafary and Shephard	yes	yes	no	no	yes	no	no	no	no	no	no	no
[64]	Liang et al.	yes	yes	no	no	no	yes	no	no	yes	no	no	no
[65]	Kleebrang et al.	yes	yes	no	no	yes	no	yes	no	yes	no	no	no
[66]	Pradhan et al.	yes	yes	yes	no	no	yes	no	no	yes	no	yes	yes
[61]	Wen and Cao	yes	yes	no	yes	yes	yes	no	no	no	no	yes	no
[59]	Radpour et al.	yes	yes	no	yes	no	no	no	no	yes	no	no	yes
[60]	Heidari et al.	yes	yes	yes	yes	no	yes	no	no	yes	no	no	yes

3.1.2 Residential Buildings

Fo residential buildings, we consider models that involve the electricity demand in residential buildings and their end-uses (e.g.,appliances, lighting, space cooling, and space heating). In contrast to the residential sector, BUEMs for residential buildings focus on: (1) the evaluation of: regulatory initiatives, (2) retrofit options (e.g., envelop elements, appliances, air conditioning, cool roofs, and changes on occupancy behavior) and programs, (3) the energy performance of building stocks, and (4) energy building standards. Likewise, there are scarce studies that analyzes BUEMs for residential buildings.

Common features of BUEMs for residential buildings are:

1. Dependency on surveys and statistics, simulated data, and information from literature
2. Single sector coverage
3. Local and project geographical coverage
4. Short and long-term time horizons

5. Yearly data time-split (in their majority)

6. Presence of space heating end-use

7. Absence of CO2 metrics

8. Non-regular presence of scenarios and cost metrics

See Table 3.4 for the complete categorization of BUEMs for Residential Buildings.

Table 3.4: Characteristics of bottom-up energy models—Residential Buildings.

Citation	Sector Coverage	Geographic Coverage	Time Horizon	Electricity End-Uses	Data	Cost	Scenario Based	CO_2
[67]	Single-sector	Local	Short	SH	Yearly	no	yes	no
[68]				SH, SC	Daily	no	no	no
[69]				SH, SC	Monthly	no	no	no
[70]	Single-sector	Local	Long-term	A, L, SC	Yearly	yes	yes	no
[71]				A, L, SC	Yearly	yes	no	no
[72]	Single-sector	Project	Short	SH, SC	Hourly	yes	yes	yes

Note: Electricity end-uses: AL = Appliances and Lighting, SC = Space Cooling, SH = Space Heating, WH = Water Heating, A = Appliances, L = Lighting, C = Cooking.

As can be regarded, BUEMs for residential buildings are single-sector models that represent space heating and space cooling end-uses. They have local or project coverage and short and long-term time horizons. Models with short time horizon are related to space heating and space cooling end-uses.

Finally, the authors document the exclusion of relevant metrics (i.e., CO2 emissions and costs), the inconsistency to utilize scenarios, the limited range of geographical coverage, and the absence of multi-sector models.

Analytical techniques and methods The construction of BUEMs for residential buildings rely on engineering (50%), hybrid (i.e., engineering + data-driven statistical 25%), and data-driven statistical (25%) modeling techniques. See Table 3.5.

Table 3.5: Techniques and Methodologies used in bottom-up energy models that support energy efficiency policy design—Residential Buildings.

Modeling Technique	Methodology	Programming Technique	Techniques used	Citations
Data-driven statistical	Other	Other	Propensity score matching method	[69]
Engineering	Simulation	Other	Transient thermodynamics equations, mathematical equations	[67, 71]
	Optimization	Mixed LP	Dynamic Building Model, MILP	[72]
Engineering - Data-driven statistical (Hybrid)	Simulation	Other	Occupant uncertainty modeling, Bayesian inference, INLA (Integrated Nested Laplace Approximation)	[68, 70]

Engineering models are correlated to simulation and optimization methodologies. The first reveals the utilization of mathematical and thermodynamic equations for model construction, while the second relies on dynamic building modeling and mixed linear programming techniques. On the other hand, hybrids models (i.e., engineering + data-driven statistical) present a combination of techniques. For instance, dynamic energy simulations are combined with integrated nested Laplace approximation (based on bayesian inference) to predict the energy performance of residential buildings [68]. Likewise, urban building energy modeling (building archetypes) and bayesian calibration stochastic approach are used to evaluate the relevance of occupant uncertainty modeling when predicting energy efficiency savings [70]. Finally, data-driven statistical models are scarce for residential buildings. A unique study that aim to validate the effectiveness of energy efficiency standards is proposed using the propensity score matching method [69].

Policy instruments BUEMs for residential buildings support mainly regulatory instruments (in 55.56% of the studies) and in less proportion market-based (22.22%), financial (11.11%), and information and feedback (11.11%). In particular, codes and standards, energy taxes, subsidies, and information instruments can be analyzed with these kinds of models. See Figure 3.2.

The results reveal a premature advance in the creation of models that support the policy design in residential buildings. First, the small number of studies that aim to support policy design in residential buildings confirm the finding. Second, the use of regulatory instruments shows the need to improve the technological inventory of residential buildings including the building construction technology and its maintenance. Since cities continue to expand, it seems important that BUEMs for residential buildings support these kinds of policies, however the other important instruments has been limited studied.

On the other hand, Table 3.6 shows that these models are useful to create general policies and specific policy instruments, however the policy calibration is not a focus on them. Likewise, the

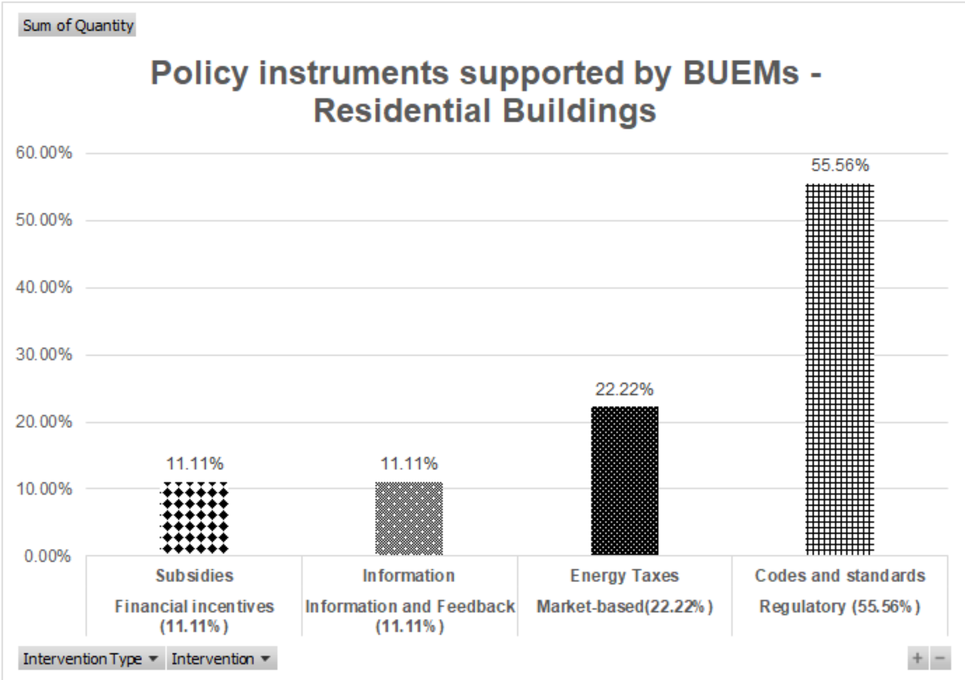


FIGURE 3.2: Proportion of policy instruments supported by BUEMs for residential buildings

capability to perform instrument comparison is possible in only two of them, which is a great opportunity area for these models in the decision-making and policy design arena.

Finally, the standardization in the use of scenarios or specific metrics is also a necessity in these kinds of BUEMs.

TABLE 3.6: BUEMs' policy abstraction levels, type of instruments supported, scenario support, and metrics - Residential Buildings.

		Policy abstraction Level			Policy instruments					Scenario and Metrics			
Citation	Author/year	General Policy	Policy instruments	Policy calibration	Market-based	Info. and Feedback	Financial	Regulatory	Non-Regulatory	Scenario	Scenario Configuration	CO_2	Cost
[67]	Marshall et al.	yes	yes	no	no	no	no	yes	no	yes	no	no	no
[68]	Braulio-Gonzalo et al.	yes	yes	no	no	no	no	yes	no	no	no	no	no
[72]	Schutz et al.	yes	yes	no	no	no	no	yes	no	yes	no	yes	yes
[70]	Cerezo Davila et al.	yes	yes	yes	yes	no	no	no	no	yes	no	no	yes
[69]	Wang et al.	yes	yes	no	no	yes	no	yes	no	no	no	no	no
[71]	Krarti et al.	yes	yes	yes	yes	no	yes	yes	no	no	no	yes	yes

3.1.3 Processes to create Data-driven BUEMs in a Decision Making Center

The modeling of data-driven BUEMs has been studied in the literature. For instance, Abbasabadi et al. [73] propose a data-driven framework for the modeling of urban energy use that localizes energy performance data considering a socio-spatial context and a bottom-up data-driven approach. The energy use data is modeled using Machine Learning techniques and the models created using this framework seems to explain with an acceptable level of certainty the variance of energy use in buildings and its transportation. The framework defines five phases to develop the models: (1) Data preparation (2) Patterns extraction (3) Prediction, (4) Analysis, and (5) Visualization. The first phase includes the location, cleaning, and processing of data.The second phase focuses on detecting and combining data to detect contextualized urban patterns. The third phase focus on calculating the performance of urban buildings by training, validating, comparing and performing predictions about the models. The fourth phase contemplates the analysis of how key attributes impact energy use. And finally, the fifth phase focuses on the creation of 2D and 3D visualizations of the models. Although the framework seems to be useful as a baseline for the creation of urban and building modeling, the methodology is not centered on decision making or policy-making processes. In this regard, no involvement of the stakeholders is reported in the first phases of the model's design. Likewise, there is no orientation in terms of how to create specific scenarios for stakeholders that facilitates decision-making and policy design.

On the other hand, a methodology for creating bottom-up modelling using simulation techniques is documented in [74]. In this work, the authors present holistic processes on how to create long-term scenarios for decarbonization for the industry sector using the FORECAST model. The scenario generation considers specific sectors (e.g., industry, services, and households), end-uses, technologies, different types of input and output data. For instance, the model can include as input: energy drivers (e.g., population, energy prices, temperature), instrument policies (e.g., taxes, CO_2-prices, standards, grants) , technology and behavior (e.g., efficiency, savings, expenses, learning, emissions, lifetime), and structure (e.g., technology distribution). Likewise, the possible results are organized in the following categories: energy demand, GHG emissions, costs, and indicators or metrics. Although this framework provides a baseline for model creation and configuration, it is only based on simulation techniques, which limits its capability to be applicable for data-driven or other kinds of techniques. Likewise, the processes are oriented to the industry sector, which restrain their applicability to the building sector specifically.

3.1.4 Research Opportunities

After reviewing the literature, we found the following research opportunities:

1. The standardization of the process and techniques used to build BUEMS should be analyzed
2. There are not documented processes for the creation of BUEMs that guarantee the support to policy design for the residential sector and residential buildings
3. The study of BUEMs metrics for the residential sector and residential buildings should be analyzed. Not only its inclusion within models but also the form in which they are included, and where they can be presented to decision-makers.
4. The utilization of data-driven techniques for BUEMs creation should be revised in future research.
5. The study of the level of policy support that BUEMs provide to decision-makers should be studied by scholars, since the concept of policy calibration is not mentioned in the model design field.
6. The design and configuration of scenarios in BUEMs is a subject that should be revised in the literature, as a mechanism to support policy design and decision-making.

In this study, we address the standardization of processes to create BUEMs through the definition of a framework. This last considers generic steps and also energy policy and energy efficiency best practices that were discovered in this chapter. Specifically, we aim to avoid the following problems visible in current BUEMs: 1) the lack of application of policy design at different levels of abstraction (e.g., high, medium, low levels) 2) The absence of a detailed method to document processes, inputs, and outputs (e.g.,), and 3) the lack of inclusion of scenarios in a consistent way on the energy models. In the next section, we document the framework and the best practices included in them.

Chapter 4

A framework to support the creation of bottom-up energy models that support electricity end-use efficiency

The literature analysis described in Chapter 3 has revealed that there are no standard processes for the creation of BUEMs that aim to support policy design of electricity end-use efficiency. This last has derived in the creation of models that do not include key information for decision making and that are not useful for policy design.

In this chapter, we present a framework that eases the creation process of these kinds of bottom-up energy models; specifically models for the residential sector and residential buildings. The framework aim to promote: 1) the use of standard processes for model design 2) the inclusion of key information useful for decision-making, and 3) the use of tools that facilitate policy design and decision-making, like the creation and configuration of scenarios. The following proposal aim to be useful for policy-designers or model-designers that want to apply proven solutions to improve the design of bottom-up energy models.

The framework includes standard phases and processes with detailed activities and artifacts (e.g., formats with best practices and technical sheets). Likewise, it is divided in six phases. The first two are Business Understanding and Data Definition and Understanding, which are devoted to the comprehension of the problem, the data, and the policy goals. The next two are Policy Alignment and Scenario Definition and Design, which are focused on the policy alignment of models and scenarios. The last two are Model Construction and Model Evaluation, which are oriented to the selection of modeling techniques, the creation and testing of the model, and to the review and assessment of the model and the creation process. See Figures 4.1 and 4.2.

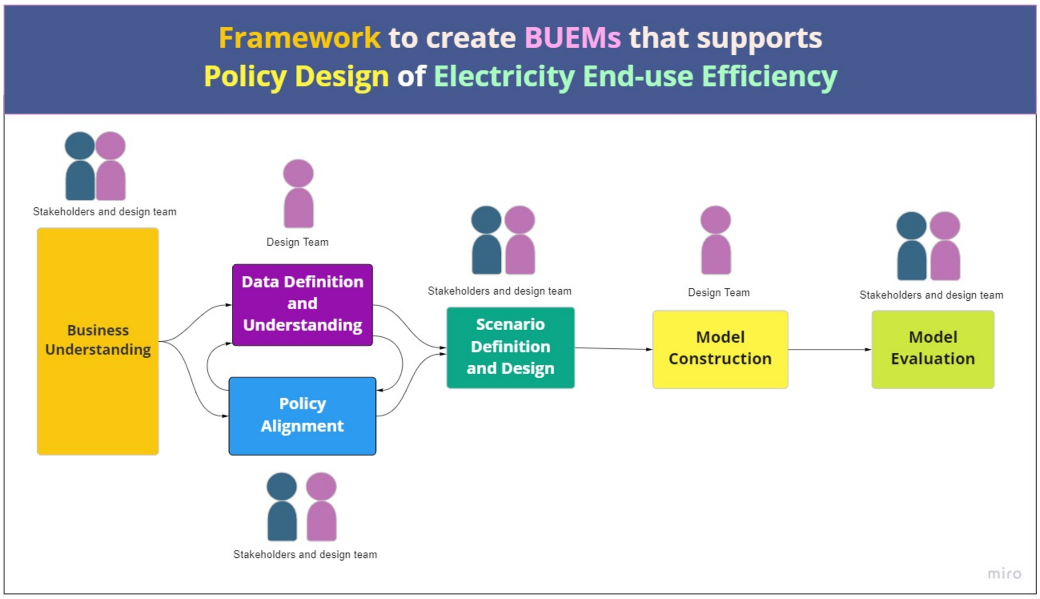


FIGURE 4.1: Proposed Framework

1.Business Understanding	2. Policy Alignment	3.Data Definition and Understanding	4.Scenario Definition and Design	5.Model Construction	6.Model Evaluation
1.1 Problem understanding 1.2 Goal definition 1.3 Business Requirements and Constraints 1.4 Strategies to achieve objectives 1.5 Plan to achieve objectives	2.1 Policy level, aims, and means 2.2 Instruments to implement 2.3 Model´s requirements definition and goal alignment	3.1 Data definition guidelines 3.2 Data acquisition 3.3 Data processing	4.1 Scenario definition 4.2 Scenario design	5.1 Create/update portfolio of design techniques 5.2 Modeling techniques selection 5.3 Model´s build and testing	6.1 Model´s results assessment 6.2 Review process 6.3 Determine next steps

miro

FIGURE 4.2: Processes included in the framework

These processes aim to guarantee the alignment of bottom-up energy models to specific policy objectives and to the stakeholder's requirements. Our processes were inspired by the CRISP-DM and Decision Making Center processes explained in Chapter 2 of this dissertation, and on the findings of the state of art, which reveals a lot of opportunity areas in the creation of bottom-up energy models.

The following sections provide a detailed explanation of the phases, processes and activities included in the framework.

4.1 Business Understanding

These processes aim to collect and understand the main problem(s) to be solved by the model, the stakeholder´s requirements, and the policy goals to be supported by the model. The phase includes the following processes: Problem Understanding, Goal Definition, Business Requirements and Constraints, Strategies to achieve objectives, and the Plan to achieve objectives. This is the first effort to align the policy goals to the model's design.

P.1.1. Problem Understanding The understanding of the problem includes the description of the problematic and its background. The inputs of the process include: literature about the problem (books, articles, or reports) and the expected output is the information contained in the "Problem Understanding - Template" that is visible on Figure 4.3.

The detailed description of the problem includes an explanation of the current situation and the roles involved. The background can contain the historic summary of it (what, when, how, and who); the definition of important concepts related to it, their consequences and possible causes, the internal and external factors that could affect it, and finally a summary of previous actions taken in the past to solve the problem and how they have worked out or failed.

P.1.2. Goal Definition The second part involves the definition of the business and policy goals considering the previous information. Business goals should represent the desired results expected by an organization, a department or a specific stakeholder. These goals can be financial, organizational, time-related, or related to the model's design. On the other hand, policy goals should indicate the expected results of the application of specific policy instruments. Policy goals can be financial or energy-related. In this step is important to identify the metrics to evaluate the accomplishment of each goal. For instance, decision-makers can be interested on metrics related to the performance of the instrument's application (e.g., energy consumption savings, energy intensity), or in financial metrics to evaluate how much is costing the policy implementation. The proposed template to document this information is available on Figure 4.4.

P.1.3 Business Requirements and Constraints This section involves the documentation of business requirements and constraints, which are important. In this phase it is performed the business requirements elicitation related to the model and the identification of possible constraints for its creation. This last allows not only define the initial scope but also the factors that could affect its construction. The requirements are prioritized and selected based on their relevance and their alignment with the business and policy objectives (defined previously).

Process Number: P.1.1. **Name: Problem Understanding**	
Inputs: Literature about the problem **Outputs:** This template´s information	
Detailed description of the problem	
Detailed Description of the Problem	Roles involved (How?)
Background	
Historic Summary	Definition of important concepts
Consequences and possible causes	Internal and External Factors
Previous actions to solve the problem	

FIGURE 4.3: Problem Understanding - Template

On the other hand, it is important to identify the model's restrictions given that these can impact the model's construction process. Examples of constraints can be limited budget or bounded resources. The template to fill the information of this step can be regarded on Figure 4.5.

P.1.4. Strategies to achieve objectives A strategy contemplates activities, methods, and resources to accomplish an objective. The framework recommend to identify the three of them for each of the goals. The strategies to achieve objectives can be divided in two kinds: 1) business and 2) policy strategies. Business strategies can be varied. For instance, if an organization has limited resources to implement a set of goals; a possible strategy could be to control and evaluate constantly the resource utilization. On the other hand, the policy strategies are bounded by

Process Number: P.1.2.
Name: Goal Definition

Inputs: Problem Understanding Template
Outputs: This template´s information

Business Goals

ID	Goal	Type of goal	Metric(s)	Requested By

Policy Goals

ID	Goal	Type of goal	Metric(s)	Requested by

Figure 4.4: Goal Definition - Template

the list of policy instruments presented in Chapter 2 of this study. For instance, if the policy goal aims to decrease energy consumption in the next 5 years in residential buildings at end-use level; then it is possible to apply one or a combination of the policy instruments (e.g., financial, regulatory, market-based, or non-regulatory) explained in the background chapter of this study. The template to document the strategies can be regarded in Figure 4.6.

P.1.5. Plan to achieve objectives To finish the phase, a plan to design policies is created using the activities, resources, and the methods identified in the previous process. In this moment, it can be also identified issues with resources and the impacts of constraints. Use the template of Figure 4.7 to document this process.

The following two phases (e.g., policy alignment and data definition and understanding) depend on each other, and usually, stakeholders execute them in parallel.

Process Number: P.1.3.
Name: Business Requirements and Constraints

Inputs: Goal Definition Template
Outputs: This template´s information

Business Requirements

ID	Requirement	Aligned with Goal	Stakeholder	Priority

Constraints

ID	Constraint	Type of constraint

Figure 4.5: Business Requirements and Constraints - Template

Process Number: P.1.4.
Name: Strategies to achieve objectives

Inputs: Business requirements and constraints Template
Outputs: This template´s information

Business Strategies

ID	Strategy	Activities	Methods	Resources

Policy Strategies

ID	Strategy	Activities	Methods	Resources

Figure 4.6: Strategies to achieve objectives - Template

Process Number: P.1.5. **Name: Plan to achieve objectives**	
Inputs: Strategies to achieve objectives Template **Outputs:** This template´s information	
Plan description	**Plan Schedule**
Issues	**Impacts of constraints**
Identified risks	

Figure 4.7: Plan to achieve objectives - Template

4.2 Policy alignment

The policy alignment phase aims to define in detail, how the policy design is going to be performed with the support of a bottom-up energy model. First, the policy level(s), aims, and means should to be defined to understand the level of granularity that expects to be managed by the model. Later (and depending on the model's granularity), it is important to detail energy policy instruments that are going to be implemented (how, when, by who) and finally the objectives that need to be achieved by the model. We explain in the following sections how to perform these processes.

P.2.1 Policy level, aims, and means In this step, the policy level is identified based on the stakeholder´s requirements and the data available policy design. For instance, if the stakeholders require to design low level policies but only high level information [1] is available, the possibility to perform the model is low. On the contrary, if high level policies are required but only low level information is available, the feasibility should be assessed. In the case of the design of electricity end-use efficiency policies for the residential sector and residential buildings; information at the level of end-uses and if it´s possible detailed timestamps (e.g., monthly, weekly or daily) can be present. In this process, the possible level(s) of abstraction must be defined (e.g., high, medium and low) with a justification or rational of the level selection. Also, the policy´s aims and means and the guidelines for instrument selection should be described in terms of the dimensions of policy selection, which include: degree of coercion, resource intensiveness, political risk and target audience. Figure 4.8 shows the template to document this section.

P.2.2. Instruments to implement In this process, the instruments to achieve the policy goals are identified and pre-selected. The successful criteria is defined for each of them (e.g., expected inputs and outputs (including metrics), implementation mechanisms(s), required resources, calibrations, and calculations). In this regard, inputs, outputs, and calculations are considered with different values (e.g., calibrations), so the decision makers can evaluate the instruments' implementation from diverse points of views. The format to document this step is available on Figure 4.9.

This step defines part of the model´s scope since these policy instruments are expected to be evaluated in the bottom-up energy model.

P.2.3. Model's requirements definition and goal alignment Finally, to close the phase, it is necessary to define the requirements that the model should meet and their alignment

[1] The Data Definition and understanding phase describes the possible levels of information that can be managed.

Process Number: P.2.1. **Name: Policy level, aims and means**	
Inputs: Business Understanding phase Templates **Outputs:** This template´s information	
Policy Level	
Policy aims	
Policy means	
Justification for instrument selection	

FIGURE 4.8: Policy level, aims, and means - Template

to the goals and requirements previously defined in the Business Understanding phase. In contrast to the stakeholder´s requirements, the model´s requirements should be defined using the characteristics of BUEMs explained on the Background Chapter. For instance, the model´s sector and geographical coverage, the data time split, time horizon, end-uses, and metrics should be defined based on the available data (documented in the Data Definition and Understanding phase) and the stakeholder´s needs. Likewise, each model should define their expected input and output variables. 4.10. Lastly, the methodology, and different techniques used to build the model are defined in the Design phase given their dependency with an available portfolio of techniques and the type of analyses requested by stakeholders.

<table>
<tr><td colspan="2">Process Number: P.2.2.
Name: Instruments to implement</td></tr>
<tr><td colspan="2">Inputs: Policy level, aims, and means Templates
Outputs: This template´s information</td></tr>
<tr><td rowspan="6">Policy instrument:

Specific instrument:

Type:</td><td>Successful Criteria</td></tr>
<tr><td>Inputs and outputs</td></tr>
<tr><td>Implementation Mechanisms</td></tr>
<tr><td>Required resources</td></tr>
<tr><td>Calibrations</td></tr>
<tr><td>Calculations</td></tr>
</table>

FIGURE 4.9: Instruments to implement - Template

Process Number: P.2.3.
Name: Model´s requirements definition and goal alignment

Inputs: Business Understanding phase and policy level Templates
Outputs: This template´s information

Stakeholder´s requirements

ID	Requirement	Business/Policy Objective ID

Inputs	Outputs

Model´s requirements

Sector Coverage	
Geographical Coverage	
Time Horizon	
Data Time Split	
End-uses	
Metrics	

Figure 4.10: Model´s requirements Definition and goal alignment - Template

4.3 Data Definition and Understanding

This phase aim to localize, analyze, understand, and document data sources and the data obtained from them. These processes are executed in parallel with the policy alignment phase's processes, since they are dependent of each other. For instance, the data search process is dependent on the specification of the stakeholders information's needs. Likewise, the model´s policy level, can be impacted by the availability of certain data, as has been mentioned in the policy alignment phase. Therefore, a close coordination in the execution of these processes is crucial to avoid delays and long dependency time. The processes included in this phase are: 1) Data definition guidelines 2) Data acquisition and 3) Data processing which includes the data quality, processing, and analyses.

P.3.1. Data definition guidelines The guideline definition process intends to standardize the information that should be documented about data sources. This last includes information about the data sources themselves, and also about the data available within. Figure 4.11 shows the proposed template to document the required data.

The documentation of the data sources include the following information: name and description of data sources, owner, location. Likewise the updating policy, type of data source, number of records, data time split, initial record date, and final record date should be specified. The updating policy refers to the periodicity in which the information should be updated, how much information should be updated, and which information should be updated. The type of data source can be: structured, semi-structured or non-structured data. The data time split can include the following resolutions: hourly/minute, daily, monthly, and yearly; as has been explained on theoretical framework. The record date identifies the time frame of the obtained information. Data time split are not specified in the CRISP-DM processes.

On the other hand, the metadata of each data source should be documented using the template shown in Figure 4.11. The recommended information to document is the following: Type of data artifact (e.g., file, table, JSON, etc.), Name of the artifact, and description. Likewise, the field information of each artifact must include: the field name, the field type (e.g, text, number, etc.), value range, and required data. Finally, if a related metric is identified it must be documented per field.

P.3.2. Data acquisition The procurement of data is a crucial process for the design of any model and to perform energy analyses. However, for energy models the kind of information that should be gathered is varied and depends on the model that wants to be constructed. Although other frameworks like CRISP-DM focus in a general way on how to obtain data and knowledge[50], they can't provide guidance on the gathering of specific kind of information and

Process Number: P.3.1. and P.3.2.
Name: Data definition guidelines and Data acquisition

Inputs: Business Understanding phase and Policy phase Templates
Outputs: This template´s information

List of Data Sources

Data Sources Documentation

ID	
Data source Name	
Description	
Owner	
Location	
Updating Policy (Which, how, how much info)	
Type of data source (structured/semi/non-structured)	
Number of records	
Data Time Split	hourly/minute/daily/monthly/yearly
Initial/final record dates	

Metadata Documentation

Type of data (file/JSON/table/other)	
Name of artifact	

Field name	Field type	Value range	Required	Business Metric

FIGURE 4.11: Data Definition guidelines - Template

the means to obtain it. In our framework we recommend the utilization of matures taxonomies and acquisition approaches, like the one created by C. Wang et al. and that can be visualized on Table 4.1. The format to document the obtained data is available on Figure 4.11.

Table 4.1: Data acquiring approaches and possible data sources for energy modeling

Type of data	Type of- approach	Sub-approach	Possible data Sources
Non-Geometric data (e.g., occupancy and appliance patterns, and HVAC-Heating, ventilation, and air conditioning)	Archetypes	Deterministic	Templates of simulation tools, standard codes and guides, and open project data
		Probabilistic	Standard codes and guides, and open project data
	Non-Archetypes	Occupancy patterns	Stochastic models, big data extraction, spatial-temporal approaches, urban factors methods
		Appliance patterns	Stochastic models,combined occupancy and environment approaches, machine learning
		HVAC Systems	Machine learning
Geometric data (e.g., building footprints and buildings heights, Windows -to-wall ratios, number of stories, and terrain data)	Direct 3D modeling		Existing databases, LiDAR, Oblique photogrammetry
	Specific Approaches	Buildings footprints	Existing databases, OpenStreetMap, Image Recognition
		Building heights	Existing databases, OpenStreetMap,nDSM, Shadow management
		WWRs	Standard and Codes, Experts´opinions, image recognition
		Number of stories	Existing databases, estimation
		Terrain	DTM, CityGML
Weather data (e.g., outdoor temperature, solar radiation, humidity, wind velocity and direction)			Typical Meteorological Year (TMY), real weather data, and future weather data.

The following steps describe the proposed process to acquire data for energy models.

First, a search of the possible data sources is performed using the approaches just described. Second, a justification of the selection of each data source should be explained. Third, the interfaces to specific data sources or artifacts should be specified. Later the processes to add, remove, and change of a data source should be defined and documented. This is a crucial step to guarantee data availability in the expected times. Finally, the impact of changes in any data sources should assessed and documented, so the possible implications of changes are identified and considered during the operation of the model.

P.3.3. Data processing The data processing phase aim to support the different treatments that can be performed to the data obtained in the previous phase. The processes that includes this phase are: 1) Data exploration, 2) Data Quality 3) Data selection, 4) Data Cleaning, 5) Data construction, 6) Data integration and 6) Data Formatting. These phases are already present in the CRISP-DM framework and are detailed explained on the CRISP-DM Guide.

4.4 Scenario Definition and Design

The Scenario design is a step that aim to continue aligning the support that the model should provide to decision-makers. In this step the following activities are executed: 1) Scenario definition and 2) scenario design.

P.4.1. Scenario definition The scenario definition process is new in any energy models' creation framework. It was included to facilitate policy design and decision-making of stakeholders and to guarantee the alignment of models to policy goals. In the background chapter, we have explained the advantages of these tools and some recommendations to implement them, however none of the methodologies studied have given the proper importance to this step or has proposed it as a method to support decision-making. We propose a process to define scenarios inspired on the research of model-based energy scenarios performed by Cao et al. [75] and on formal approaches to develop scenarios for environmental decision-making [76]. The proposed activities to perform this process are the following: 1) Specify the name of the scenario 2) Define the characteristics of the scenario (e.g., time horizon, geographical and sector coverage). These should be aligned with the model´s requirements defined in the second phase of this framework 3) Determine the way scenarios are going to provide policy recommendations. In this step is important to describe which instruments are going to be included in the scenario (e.g., none, one or more) and the justification of this decision 4) Identify key variables (e.g., endogenous and exogenous) and driving forces of the scenario (e.g., climate variations, practices, regulations, policies or socio-economic patterns) 5) Describe scenario´s narrative and outputs 6) Document the assumptions of input and output data, calculations, and processes and how they can affect the results provided by the scenario and finally 7) Identify uncertainties (e.g., possible future changes). In this activity, is important to describe possible uncertainties on variables or processes, so they can be considered in future iterations of scenarios' construction.

P.4.2. Scenario design In the scenario design phase: 1) the causal relationships of variables and external conditions are identified using tools like: cause-effect diagrams or statistical techniques, so they can be modeled in subsequent phases 2) Identifying critical uncertainties is also key for scenario design since it is possible that designers are not able to predict the true value of certain variables. According to [76], some causes of uncertainty can be: lack of knowledge, data errors, mistakes with model structures and parameters, incorrect assumptions, among others. 3) The documentation and acquisition of scenario´s datasets is key to perform any design or construction of models. This activity should be aligned with the data acquisitions processes described on the previous phases of this framework 4) Finally, it should be validated the data resolutions and scales since they can impact the model´s results. This activity should be aligned with the data quality sub-activity mentioned on the Data understanding phase.

Process Number: P.4.1. **Name: Scenario Definition**	
Inputs: Previous templates **Outputs:** This template´s information	
Scenario Number:	**Scenario Name:**
Inputs:	
Features	

Time Horizon	
Geographical coverage	
Sector coverage	

Policy recommendations that the model provides.

Key variables	**Scenario Narrative**
Exogenous variables: Endogenous variables:	
Outputs	
Assumptions	

FIGURE 4.12: Scenario Definition - Template

Process Number: P.4.2. **Name: Scenario Design**
Causal relationships The following diagram shows the analyzed causal relationships of variables:
Critical Uncertainties
Scenario Datasets

Figure 4.13: Scenario Design - Template

4.5 Model Construction

The model construction phase aims to build a model that is complete, usable, and aligned to the model´s objectives using a set of design techniques.

P.5.1. Create/update portfolio of design techniques As was evidenced in the background of this research, the techniques and methodologies used to model a system are diverse and depend on the kind of model that want to be constructed. In this research we provide a taxonomy of modeling techniques and methods 2.4 for the residential sector and residential buildings, however it can be expanded and improved to be a future reference for organizations. The following Template 4.15 can be used for documenting a portfolio of design techniques.

Process Number: P.5.1. and P.5.2. **Name: Create/update portfolio of design techniques and techniques selection**
Inputs: Used/new techniques **Outputs:** This template´s information
Design Techniques

ID	
Technique Name	
Type of technique (modeling/methodology/programming)	
Description	**Applications**
	References
Selection/Rejection Justification	
Used tools:	

Figure 4.14: Portfolio of design techniques - Template

P.5.2. Modeling techniques selection One of the disadvantages of frameworks like CRISP-DM is their strictly focus on the selection of data-driven modeling techniques; however as we have shown in Chapter 3, diverse kinds of methods are used to model BUEMs, including engineering and hybrid techniques (e.g., data-driven + engineering). In the state of the art presented on Chapter 3, we identify the kind of techniques used for the design of Bottom-up energy models than aim to support policy design of end-use energy efficiency. Tables 3.2 and 3.5 aim to be reference points that helps stakeholders decide which techniques to use. However, the reader should contemplate specific constraints before taking a final decision like: business or organizational policies, budgets or future maintenance costs.

P.5.3. Model's build and testing The construction of the model should be aligned to the model´s goals and the defined scenarios. Likewise, the required data should be available. The following activities are proposed to build the model: 1) Define a prototype to construct (inputs, outputs parameters, default settings, and calculations) 2) Create a test design for the prototype (include different calibrations) 3) Create the model´s prototype 4) Test each scenario adjusting inputs and calibrations [2] 5) Iterate phase until the final model is completed. For this process, we plan to construct the model in short stages, so designers can encounter in initial stages any problem with its design or execution.

[2]CRISP-DM recommend that models should be evaluated considering its accuracy and generality, however according to [76] scenarios can be evaluated using sensibility analyses

<table>
<tr><td colspan="2">Process Number: P.5.3.
Name: Model´s build and testing</td></tr>
<tr><td colspan="2">Inputs: Previous templates
Outputs: This template´s information</td></tr>
<tr><td colspan="2">Prototype definition</td></tr>
<tr><td colspan="2">Inputs:
Parameters:
Default settings:

Outputs:
Calculations:</td></tr>
<tr><td colspan="2">Prototype and calibrations</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td><u>Scenario Name:</u>

Calibrations:</td><td><<Model execution>></td></tr>
<tr><td colspan="2">Graphical Representation</td></tr>
<tr><td colspan="2">Prototype testing</td></tr>
<tr><td colspan="2"></td></tr>
</table>

FIGURE 4.15: Model's build and testing - Template

4.6 Evaluation and Review

This phase aims to validate that the model and scenarios meet the stakeholder´s business objectives and requirements.

P.6.1. Model's results assessment In this activity the model´s results are presented to the stakeholders and they evaluate them, based on their expertise.

Process Number: P.6.1., P.6.2, and P.6.3.
Name: Evaluation and Review

Inputs: Previous templates
Outputs: This template´s information

Model´s results assessment

To evaluate the model´s results, we evaluate it compliance towards the previous defined goals:

Goal	Description	Compliance Level (Achieved/Non Achieved)	Comments

Review Process

Strengths	Weaknesses
Opportunities	**Threats**

Determine next steps

Figure 4.16: Evaluation and Review - Template

P.6.2. Review process After performing the evaluation decision-makers and designers can review the processes to create the model, find opportunity areas, and best practices,

P.6.3. Determine next steps In this final step, a list of possible activities to overcome the problems are presented (in order of importance) and decision makers define which actions will be taken to overcome them.

In the next section, we present a case study to validate the usefulness of presented framework and how can it be applied to real life projects.

Chapter 5

Case Study: A Model to support electricity end-use policies for stand-by appliances in a Decision-Making Center

The following case study aims to show how the framework described in the previous chapter can guide the design of BUEMs that intent to support policy design and how its utilization can avoid the problems described in the state of the art of this research, or existing problems in models developed in a Decision Making Center.

The model to be constructed was defined based on the analysis needs of Tecnologico de Monterrey´s Decision Making Center. The center has collaborated with the Mexican government in diverse energy modeling projects in the past[1], therefore some of these models can be improved and expanded. The model aims to predict the effects of the implementation of energy efficiency policies in the residential sector and residential buildings.

In the following sections, we explain how the new model was defined, designed, and built using the proposed framework. We include the detailed outputs of the process and how the activities were executed to obtain the expected result. Finally, the creation process has been measured and recorded activity by activity, so it can be compared with the development of other similar models and validate its utility.

[1]The Mexican Electric Sector Model [45] was integrated with the consumption sector model, the petrochemical model, the exploration model, the supply sector model, the environmental and economic impacts model, and the refining model.

5.1 Business Understanding Phase

The model selected for this case study shows the effects of regulatory instruments in the residential sector and residential buildings. Specifically, the application of energy efficiency norms in stand-by devices. We choose the stand-by end-use because it has been identified as an important end-use in the Mexican residential sector in the last years. The Mexican Commission of Efficient Use of Energy (CONUEE) reports that this end-use represent 5.6% of electricity consumption in the residential sector and in fact, it has become an independent end-use by itself. Figure 5.1 shows the description of the problem, the roles involved, and the background research about the problem.

<table>
<tr><td colspan="2">Process Number: P.1.1.
Name: Problem Understanding</td></tr>
<tr><td colspan="2">Inputs: Literature about the problem
Outputs: This template´s information</td></tr>
<tr><td>Detailed description of the problem</td><td>Roles involved</td></tr>
<tr><td>Electricity consumption of standby devices has increased in the residential sector and residential buildings in previous years</td><td>Final users: People that consumes end-use energy in the form of electricity
CONUEE: Commission that monitors energy efficiency in different sectors of the country.
DMC: Decision Making center at Tecnológico de Monterrey</td></tr>
<tr><td colspan="2">Background</td></tr>
<tr><td>Historic Summary</td><td>Definition of important concepts</td></tr>
<tr><td>In the past, CONUEE has implemented different kinds of policies to promote electricity end-use efficiency in the residential sector; however the impact of those are not completely understood</td><td>End-use energy consumption: Energy consumed by end-uses (e.g., appliances, lighting, etc,...)
Energy efficiency instruments: Mechanisms to address specific policy goals.</td></tr>
<tr><td>Consequences and possible causes</td><td>Internal and External Factors</td></tr>
<tr><td>Consequences:
• More electricity needs to be generated to procure the demand
Possible causes:
• Use of more electric devices
• Changes in temperature encourage the use of more electricity to achieve comfortable temperatures.
• Increase in population/households</td><td>External factors:
• Electricity price
• Government Regulations
• Geography - Climate (Seasonality)
• Device penetration

Internal factors:
• Number of people in the household
• Number of devices in the households
• Energy efficiency levels of devices
• Lifespan of devices
• Thermal envelope of the household
• Habits of people in the household</td></tr>
<tr><td colspan="2">Previous actions to solve the problem

CONUEE has implemented the following normatives to decrease the end-use of electricity in the residential sector.
• 2013 Application of Energy efficiency norms for stand-by energy in devices</td></tr>
</table>

FIGURE 5.1: Case Study-Problem Understanding

The main literature source for the background research has been technical reports [77] [45] of past projects and reports published by the government [78]. In this information, we have been able to identify the possible causes, consequences, internal and external factors of the problematic; with a high level of abstraction and without worrying about implementation details at this level. We also learned in this step, that in 2013 a normative to stand-by energy devices has been issued by the government [78]; and that it was possible to predict the energy savings of that program.

On the other hand, regarding the business goals, Tecnologico de Monterrey´s Decision Making Center, is mainly concerned with the tools to be used for the model´s development; since they rely on their DMC Environment and they want to make sure that these tools are 100% compatible with their DMC´s technologies. See Figure 5.2. Additionally, CONUEE is interested on diminishing the energy consumption in households through the introduction of more efficient stand-by devices to the Mexican market [78], which has been a common policy that have given results in the previous years.

Process Number: P.1.2.
Name: Goal Definition

Inputs: Problem Understanding Template
Outputs: This template´s information

Business Goals

ID	Goal	Type of goal	Metric(s)	Requested by
BG1	Any tool or system should be aligned with a portfolio of tools and techniques defined by the organization.	Organizational	Metric: Alignment percentage	DMC

Policy Goals

ID	Goal	Type of goal	Metric(s)	Requested by
PG1	Decrement/Control energy consumption in households of stand-by devices	energy-related	* Energy consumption *energy savings	CONUEE
PG2	Final users should use more efficient stand-by devices	energy-related	* Energy consumption *energy savings	CONUEE

Figure 5.2: Case Study- Goal Definition

In the same way, the inclusion of specific scenarios in the model was defined as a key requirement with a high priority. However, there were identified also some constraints for the policy design, for instance: 1) budget constraints and 2) lack of availability of information that could jeopardize the model´s construction; as can be regarded in Figure 5.3.

Finally, to achieve the defined objectives, stakeholders have proposed the following strategies: 1) the search and evaluation of tools to build scenarios through the utilization of prototypes (with limited scope), and 2) to incentive the creation of efficient stand-by devices through the application of codes and standards. See Figure 5.4.

The plan proposed by the DMC include not only the activities to build the model but also the required actions to implement, monitor, and evaluate the selected policy. See Figure 5.5.

Process Number: P.1.3.
Name: Business Requirements and Constraints

Inputs: Goal Definition Template
Outputs: This template´s information

Business Requirements

ID	Requirement	Aligned with Goal	Stakeholder	Priority
BR1	The model should allow the evaluation of energy saving scenarios in the residential sector and residential buildings	PG1	CONUEE	High

Constraints

ID	Constraint	Type of constraint
1	Budget of implementation is: $500,000 Mexican pesos	Financial
2	Availability of information	Information

FIGURE 5.3: Case Study- Business Requirements

Process Number: P.1.4.
Name: Strategies to achieve objectives

Inputs: Business requirements and constraints Template
Outputs: This template´s information

Business Strategies

ID	Strategy	Activities	Methods	Resources
S1	Evaluate tools that allow the implementation of scenarios	*Search tools in the organization´s techniques´ portfolio *Compare tools features *Select candidate techniques		*People: Experts in design tools and techniques *Budget

Policy Strategies

ID	Strategy	Activities	Methods	Resources
S2	Make efficient stand-by products	*Meet with producers *Agree an efficiency standards *Introduce products to the market *Measure electricity consumption *Evaluate the plan	*Meetings *Models	People: Experts on stand-by devices, Managers Budget Model´s designers

FIGURE 5.4: Case Study - Strategies to Achieve Objectives

Process Number: P.1.5. **Name: Plan to achieve objectives**	
Inputs: Strategies to achieve objectives Template **Outputs:** This template´s information	
Plan description	**Plan Schedule**
Important phases: 1. Business understanding (BU) 2. Policy design (PD) 3. Model creation activities a. Data definition and understanding (DDU) b. Policy Alignment (PA) c. Scenario definition (SD) d. Model construction (MC) e. Model evaluation (ME) 4. Policy implementation (PI)	BU - 2 to 3 weeks PD - 2 to 4 weeks DDU - 2 weeks PA - 1 week SD - 1 week MC - 2 or 3 weeks ME - 1 week PI - < 12 months
Issues	**Impacts of constraints**
• DDU phase can be delayed due to lack of information availability	• Delay the next phases
Identified risks	
• Availability of manufacturers • Constraint budget	

Figure 5.5: Case Study - Plan to Achieve Objectives

5.2 Policy Alignment Phase

The previous phase give us an general understanding of the problem to be solved and which strategies can be applied. In fact, we learned in the previous phase that: 1) policy instruments should be applied at a national level given the specified scope (e.g., residential sector and residential buildings) and 2) instruments should be applicable to specific end-uses. In Mexico if a policy instrument of that magnitude needs to be applied, it should be issued by a national energy commission like CONUEE and then applied nationwide. Since stakeholders wanted to validate the implementation of policy instruments and possibly its calibration, the policy level to be implemented in the new model can be medium or low. See Figure 5.6.

Process Number: P.2.1. **Name: Policy level, aims and means**	
Inputs: Business Understanding phase Templates **Outputs:** This template´s information	
Policy Level	*Medium/Low* level policies are used since policy objectives are defined and policy instruments are used to address them.
Policy aims	The defined policies aim to achieve the goals specified in the “Goal Definition” template
Policy means	Policy instruments are used as mechanisms to implement energy policies
Justification for instrument selection	For Instrument selection, we consider the following guidelines: 1) **Degree of coercion**: The Instruments used and accepted by Mexican regulators are recommended for implementation. 2) **Resource intensiveness**: Instruments that are affordable to implement and easy to manage by administrative offices are recommended. 3) **Political risk**: Instruments that has been validated and tested in the past are recommended as first options for implementation 4) **Target**: Instruments that address the concerns and necessities of different stakeholders are favored for implementation.

Figure 5.6: Case Study - Policy level, aims and means

In previous years, the Mexican government has favored the use of regulatory instruments for the residential sector given their low risks of implementation, their moderated costs, and their impacts on different stakeholders (e.g., final users, the government, and device manufacturers). Since they want to continue with same policy orientation, the selected instrument considers the implementation of codes and standards. The instrument's implementation is detailed on Figure 5.7.

The most important output of these kinds of instruments is the normative itself, but also its validation, which normally is performed with a model. Figure 5.8 shows the agreed requirements

<table>
<tr><td colspan="2">Process Number: P.2.2.
Name: Instruments to implement</td></tr>
<tr><td colspan="2">Inputs: Policy level, aims, and means Templates
Outputs: This template´s information</td></tr>
<tr><td rowspan="6">Policy instrument: Codes and standards

Specific instrument: Energy performance standard of stand-by devices

Type: Regulatory measures</td><td>Successful Criteria</td></tr>
<tr><td>Inputs and outputs

Inputs: Electricity consumption of current stand-by devices, Number of stand-by devices per household, expected efficiencies of current stand-by devices, regulations on stand-by devices, adoption rate of new efficient stand-by devices.

Outputs: Electric Efficiency Mexican Norm and its implementation validation.</td></tr>
<tr><td>Implementation Mechanisms
• Meetings
• Definition of technical norms
• Energy efficiency improvements of devices</td></tr>
<tr><td>Required resources
• People: Managers, policy-makers, designers, manufacturers
• Tools: Models</td></tr>
<tr><td>Calibrations
• Efficiency of electric devices</td></tr>
<tr><td>Calculations
• Electricity consumption per end-use
• Electricity savings per end-use</td></tr>
</table>

Figure 5.7: Case Study - Instruments to Implement

for this model. Now, since the application of regulatory policies takes time, because it depends on different variables like: the technical capabilities of device manufacturers, the lifespan of current devices in the market, the introduction to the market of the efficient devices, and finally the device acquisition by final users. The model should reflect that reality in the scenarios. In general, stakeholders require that this BUEM calculate and predict the national consumption of electricity of specific stand-by devices (e.g., microwave ovens) in a period of 5 to 8 years considering the application of specific codes and standards to stand-by devices.

Process Number: P.2.3.
Name: Model´s requirements definition and goal alignment

Inputs: Business Understanding phase and policy level Templates
Outputs: This template´s information

Stakeholder´s requirements

ID	Requirement	Business/Policy Objective ID
1	The model should show electricity consumption scenarios	PG1
2	The model should show the application of codes and standards in stand-by devices	BR1

Inputs	Outputs
• Adoption rate of new efficient devices	• Prediction of Electricity consumption in the next 5 years or more • Prediction of electricity savings in the next 5 years or more

Model´s requirements

Sector Coverage	Residential and buildings sectors
Geographical Coverage	National (sector within a country)
Time Horizon	Short (< 5 years) or Medium (>=5 years)
Data Time Split	Yearly/monthly data resolution
End-uses	stand-by devices
Metrics	electricity consumption per end-use

FIGURE 5.8: Case Study - Models Requirements

5.3 Data Definition and Understanding Phase

The following data sources are considered key in the design of the expected model (See Figure 5.9):

1. Surveys performed by Mexican government from the year 2014 to the year 2020. However in the year 2019 no survey was conducted by the government, so we assume that the number of microwaves remain constant in the years 2020 and 2019.
2. Norms issued by the Mexican government in regard to the maximum energy consumption of microwaves in stand-by mode.
3. Metrics of the current efficiency of microwaves in the market and its lifespan, which have to be obtained from the state of the art, since no official report in Mexico provides that information.
4. Energy efficiency national reports that contains energy efficiency metrics of the last 20 years in the residential sector and residential buildings (i.e., average number of microwaves in households and microwaves saturation rates).
5. Recently reports about the estimated energy saving of specific norms and standards (specifically for the year 2021), so we can perform a validation of the model.

The energy consumption and energy savings metrics requested by the stakeholders are reported on GWh/year (Giga-watts-hour per year), which is the standard metric used by the government.

The data acquisition is documented per data source and it's available in Appendix A of this document. The data processing phase is executed but not formally documented in this moment, however it's an opportunity area for the future.

5.4 Scenario Definition and Design Phase

For this phase, stakeholders propose two scenarios for modeling: 1) A Business as Usual scenario (BAU) without the application of policies and 2) A Technology Potential scenario, where the replacement of non-efficient for efficient devices (with average efficient consumption) is visible. Results for the application of these kind of policies are visible between 5 and 10 years after the norm was issued so, we find it useful to show a 8 year period prediction (i.e., 2021-2028) in all the scenarios. Figure 5.10 and 5.11 show the specification of these scenarios.

On the other hand, since some devices can achieve more or less efficiency depending on the kind of microwave device that is modeled, we decided to create a third scenario that shows

Process Number: P.3.1.
Name: Data definition guidelines

Inputs: Business Understanding phase and Policy phase Templates
Outputs: This template´s information

List of Data Sources (Obtained on October 12th, 2022)

1) **Number of mexican households with stand-by devices(e.g., microwave ovens) 2014-2020 - INEGI**

   2014 - Encuesta Nacional en los Hogares
   https://www.inegi.org.mx/programas/enh/2014/#Microdatos
   2015 - Encuesta Nacional en los Hogares
   https://www.inegi.org.mx/programas/enh/2015/#Microdatos
   2016 - Encuesta Nacional en los Hogares
   https://www.inegi.org.mx/programas/enh/2016/#Microdatos
   2017 - Encuesta Nacional en los Hogares
   https://www.inegi.org.mx/programas/enh/2017/#Tabulados
   2018 - Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). 2018 Nueva serie
   https://www.inegi.org.mx/programas/enigh/nc/2018/#Microdatos
   2020 - Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). 2020 Nueva serie
   https://www.inegi.org.mx/programas/enigh/nc/2020/#Microdatos

2) **Regulations of stand-by devices**
   https://www.dof.gob.mx/nota_detalle.php?codigo=5330530&fecha=23/01/2014#gsc.tab=0

3) **Current efficiency of stand-by devices**
   https://www.researchgate.net/publication/260585939_Reduce_the_Standby_Power_Consumption_of_a_Microwave_Oven

4) **Average number of stand-by devices per household (CONUEE) - Specifically microwaves)**
   https://www.cepal.org/es/publicaciones/43612-informe-nacional-monitoreo-la-eficiencia-energetica-mexico-2018

5) **Average Life-span of microwave ovens**
   https://www.researchgate.net/publication/330998256_Analysis_of_the_impact_of_energy_efficiency_labelling_and_potential_changes_on_electricity_demand_reduction_of_white_goods_using_a_stock_model_The_case_of_Switzerland

6) **Saturation rates**
   https://www.cepal.org/es/publicaciones/43612-informe-nacional-monitoreo-la-eficiencia-energetica-mexico-2018

7) **Estimated savings of norms and standards 2021 (For model validation)**
   https://www.gob.mx/cms/uploads/attachment/file/739347/BALANCE_NOM-ENER_al_2021_ver_final_05072022__1_.pdf

Figure 5.9: Case Study - Data Definition Guidelines

the replacement of devices for the most efficient ones on the market. The scenario is called Technology Potential scenario with high efficiency devices and its definition is available on Figure 5.12. For these scenarios the calibration mechanism is the variable that represent the efficiency of the new devices introduced to the market.

Process Number: P.4.1.
Name: Scenario Definition

Inputs: Previous templates
Outputs: This template´s information

Scenario Number: 1	**Scenario Name:** Electricity consumption of stand-by devices (BAU-Business as Usual)

Inputs: Stand-by electricity consumption of devices = baseline (NO governmental regulation)

Features

Time Horizon	<= 8 years (2014-2028)
Geographical coverage	National (sector within a country)
Sector coverage	Residential sector and residential buildings

Policy recommendations that the model provides.

- The scenario will provide a perspective of electricity consumption of stand-by devices if no policy action is taken

Key variables	**Scenario Narrative**
Exogenous variables: • Desired electricity consumption of stand-by devices Endogenous variables: • Stand-by electricity consumption of devices • Saturation of stand-by devices	1. The model considers the baseline electricity consumption of devices as desired electricity consumption (before any policy) 2. The scenario calculated the electricity consumption and energy saving of the specified time horizon. 3. The model shows the scenario to the user

Outputs

- National stand-by energy consumption (Gwh/year)
- National Electricity savings (Gwh/year)

Assumptions

- The average appliance saturation rate is considered constant per year

Figure 5.10: Case Study - Scenario Definition(1)

Process Number: P.4.1.
Name: Scenario Definition

Inputs: Previous templates
Outputs: This template´s information

Scenario Number: 2	**Scenario Name:** Electricity consumption of stand-by devices (TEP-Technology Potential Scenario) - Devices with average efficiency

Inputs: Stand-by electricity consumption of efficient devices = (WITH governmental regulation: norms and standards)

Features

Time Horizon	<= 8 years (2014-2028)
Geographical coverage	National (sector within a country)
Sector coverage	Residential sector and residential buildings

Policy recommendations that the model provides.
This scenario will provide a view of the application of energy efficiency policies and norms to stand-by devices in the residential sector and residential buildings.

Key variables	**Scenario Narrative**
Exogenous variables: • Desired electricity consumption of stand-by devices Endogenous variables: • Stand-by electricity consumption of devices • Saturation of stand-by devices	1) The user provides the energy performance (the norm/standard defined by the government) 2) The scenario calculates the electricity consumption and energy saving in the specified time horizon. 3) The model shows the scenario

Outputs
- National stand-by energy consumption (Gwh/year)
- National Electricity savings (Gwh/year)

Assumptions
- The average appliance saturation rate is considered constant per year

Figure 5.11: Case Study - Scenario Definition(2)

Process Number: P.4.1.
Name: Scenario Definition

Inputs: Previous templates
Outputs: This template´s information

Scenario Number: 3	**Scenario Name:** Electricity consumption of stand-by devices (TEP-Technology Potential Scenario) - Devices with high efficiency

Inputs: Stand-by electricity consumption of high efficient devices = (WITH governmental regulation: norms and standards)

Features

Time Horizon	<= 8 years (2014-2028)
Geographical coverage	National (sector within a country)
Sector coverage	Residential sector and residential buildings

Policy recommendations that the model provides.
This scenario will provide a view of the application of energy efficiency policies and norms to stand-by devices in the residential sector and residential buildings.

Key variables	**Scenario Narrative**
Exogenous variables: • Desired electricity consumption of stand-by devices Endogenous variables: • Stand-by electricity consumption of devices • Saturation of stand-by devices	4) The user provides the energy performance (the norm/standard defined by the government) 5) The scenario calculates the electricity consumption and energy saving in the specified time horizon. 6) The model shows the scenario

Outputs

- National stand-by energy consumption (Gwh/year)
- National Electricity savings (Gwh/year)

Assumptions

- The average appliance saturation rate is considered constant per year

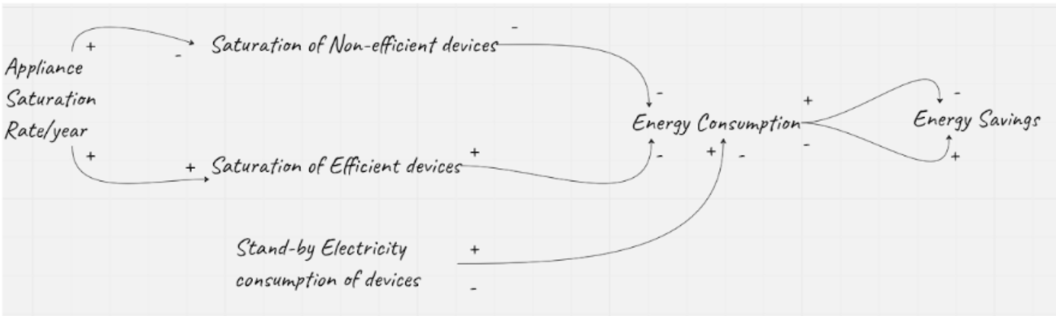
FIGURE 5.12: Case Study - Scenario Definition(3)

The scenario design is based on the following ideas: 1) there are efficient and non-efficient devices in the market (in some proportion) 2) There is a regular appliance saturation rate per year 3) The saturation rate is directly proportional with the number of efficient devices and inversely proportional with the number of non-efficient devices 4) The energy consumption depends on the number of efficient and non-efficient device and on the electricity consumption of each of them. Figure 5.13 shows the complete causal relationship diagram of the model´s variables, such as the uncertainties and the required datasets. With these inputs, we began the implementation of the model.

Process Number: P.4.2.
Name: Scenario Design

Causal relationships

The following diagram shows the analyzed causal relationships of variables:



Critical Uncertainties

- The appliance saturation rate was calculated using the average saturation metrics provided by the government, since no available data source provides that information.

Scenario Datasets

- The datasets described on the Data acquisition phase are necessary for the modeling of the scenario
- The metric related to the lifespan of microwave ovens was obtained from the literature

FIGURE 5.13: Case Study - Scenario Design

5.5 Model Construction

For the model construction, stakeholders allow the evaluation of new and experimental techniques, so we assessed two possible choices. The first choice was time series analysis and the second was machine learning techniques (supervised and unsupervised methods). However, the second choice was rejected immediately given the yearly resolution of the data; normally AI modelling techniques are useful in granular analyses with hour, minute, or day time resolution (See Figure 5.14).

Process Number: P.5.1. and P.5.2.
Name: Create/update portfolio of design techniques and techniques selection

Inputs: Used/new techniques
Outputs: This template´s information

Design Techniques

ID	2
Technique Name	Machine Learning (supervised and unsupervised models)
Type of technique (modeling/methodology/programming)	Data-driven AI-based
Description	**Applications**
• Useful to estimate end-use energy • Historical data is necessary	• Applicable in modeling heat energy prediction, operational energy modeling, hourly energy use prediction.
	References
	[73]
Selection/Rejection Justification	
We reject these techniques for this model since their applications are not aligned with our model´s goals and aims. Likewise, the data is not in hour, minute or day resolution.	

Figure 5.14: Case Study - Rejected design technique

Conversely, the time series technique seems like a more promising approach due to its applications on the prediction of energy trends, which is exactly what we needed to do in the model. We select the double exponential smoothing method (i.e., Holt's Method[2]) since the data exhibits a linear trend and also because it was the method that better represents the data.

The technique's documentation and the selection rationale can be regarded in Figure 5.15.

[2]The Holt's method calculates the expected level of the time series and the expected increasing or decreasing rate of the trend [79]

Process Number: P.5.1. and P.5.2.
Name: Create/update portfolio of design techniques and techniques selection

Inputs: Used/new techniques
Outputs: This template´s information

Design Techniques

ID	1
Technique Name	Time series analysis (Double Exponential Smoothing - Holt's Method)
Type of technique (modeling/methodology/programming)	Data-driven statistical
Description	**Applications**
• Useful to estimate end-use energy • Historical data is necessary	• Prediction of Energy trends **References** [52]

Selection/Rejection Justification

The focus of our model is to validate the results of the application of end-use policies for the residential sector and residential buildings. Specifically, we focus on the regulation of specific appliances (e.g., microwave ovens). Therefore with these data-driven techniques, we aim to estimate and predict the future end-use electricity consumption and energy savings of microwave ovens in stand-by mode.

Used tools: R programming language

Figure 5.15: Case Study - Accepted design technique

Next, the model´s prototype was designed with the inputs, outputs, parameters, setting, calibrations, and calculations described on Figure 5.16.

Process Number: P.5.3. **Name: Model´s build and testing**
Inputs: Previous templates **Outputs:** This template´s information
Prototype definition
Inputs: Stand-by electricity consumption of devices (W) **Parameters:** with policy/without policy **Default settings:** Saturation of devices/microwaves , saturation percentage (devices >=10 years, devices < 10 years), households, saturation rate per year **Outputs:** Total electricity consumption (GWh), Total electricity savings (GWh) **Calculations:** Saturation of efficient and non-efficient devices, energy consumption, and energy savings
Prototype and calibrations
• Electricity consumption and savings ○ No Policy scenario (Business as Usual) ○ Normal Saving scenario ○ Best Saving Scenario • Electricity consumption of efficient devices

Figure 5.16: Case Study - Model´s build

The definitions of the variables used in the model are the following:

- Households: Number of households in the country
- Stand-by electricity consumption: The average consumption of energy in stand-by mode of a device.
- Saturation of microwaves ovens/devices: The number of microwave ovens/devices in the Mexican market
- Saturation percentage: Proportion of devices (i.e., microwaves) that have certain age (e.g., less than 10 years or more o equal to ten years)
- Saturation rate per year: Proportion of new efficient devices that enter the market in a year

Likewise, the calculations performed by the model can be viewed in Figure 5.17. For simplicity, we present only the variables and the basic computations, however the real code was implemented with the R programming language. The code can be regarded in the following repository: https://github.com/tsetsuna/Thesis/blob/main/TSModel-Thesis.

```
Saturation Of Devices =  Households X  percentage Of Device Saturation

Saturation Efficient Devices With Policy =
Saturation Devices  X  Saturation Percentage efficient devices

Saturation Non Efficient Devices With Policy =
 Saturation Of Devices X Saturation Percentage non-efficient devices

Saturation Efficient Devices No Policy  =
 Saturation Efficient Devices X ( 1 - Saturation rate per year)

Saturation Non Efficient Devices No Policy  =
 Saturation Efficient Devices X ( 1 + Saturation rate per year)

###############   NO POLICY SCENARIO ###############################
Stand_by electricity consumption_efficient= 2.5
Stand_by electricity consumption_Non_efficient = 4
Hours_in_a_year = 8760

ConsumptionGwH NoPolicy =
((Saturation Efficient Devices No Policy x Stand_by electricity consumption_efficient)
+ (Saturation Non Efficient Devices No Policy x Stand_by electricity
consumption_Non_efficient))/10^9 X Hours_in_a_year

###############   NORMAL POLICY SCENARIO ###############################
Stand_by electricity consumption_efficient= 2
Stand_by electricity consumption_Non_efficient = 3
Hours_in_a_year = 8760

ConsumptionGwH PolicyNormal =
((Saturation Efficient Devices Policy x Stand_by electricity consumption_efficient)
+ (Saturation Non Efficient Devices Policy x Stand_by electricity
consumption_Non_efficient))/10^9 X Hours_in_a_year

###############   BEST POLICY SCENARIO ###############################
Stand_by electricity consumption_efficient= 1
Stand_by electricity consumption_Non_efficient = 3
Hours_in_a_year = 8760

ConsumptionGwH PolicyBest =
((Saturation Efficient Devices Policy x Stand_by electricity consumption_efficient)
+ (Saturation Non Efficient Devices Policy x Stand_by electricity
consumption_Non_efficient))/10^9 X Hours_in_a_year


###############   ELECTRICITY SAVINGS ################################

SavingsGWh_PolicyNormal = ConsumptionGwH NoPolicy  - ConsumptionGwH PolicyNormal

SavingsGWh_PolicyBest = ConsumptionGwH NoPolicy  - ConsumptionGwH PolicyBest
```

FIGURE 5.17: Case Study - Model´s Calculation Code

The Business as Usual scenario considers no changes in technology, so no savings are expected. With these conditions it is 95% probable that electricity consumption of microwave ovens on stand-by mode from 2021 to 2028, ranges between 418-557 GWh per year.

The executed scenario can be regarded in Figure 5.18.

We calibrate the model by changing the parameters in the exponential smoothing algorithm (i.e., beta = TRUE which means that we are modeling data with trend). When we made this change the SSE (Sum of Squared errors) decrease considerably.

Additionally, we test the electricity savings on the Technology Potential Scenario on devices with average efficiency. The scenario revealed that it´s 95% probable that energy savings ranges

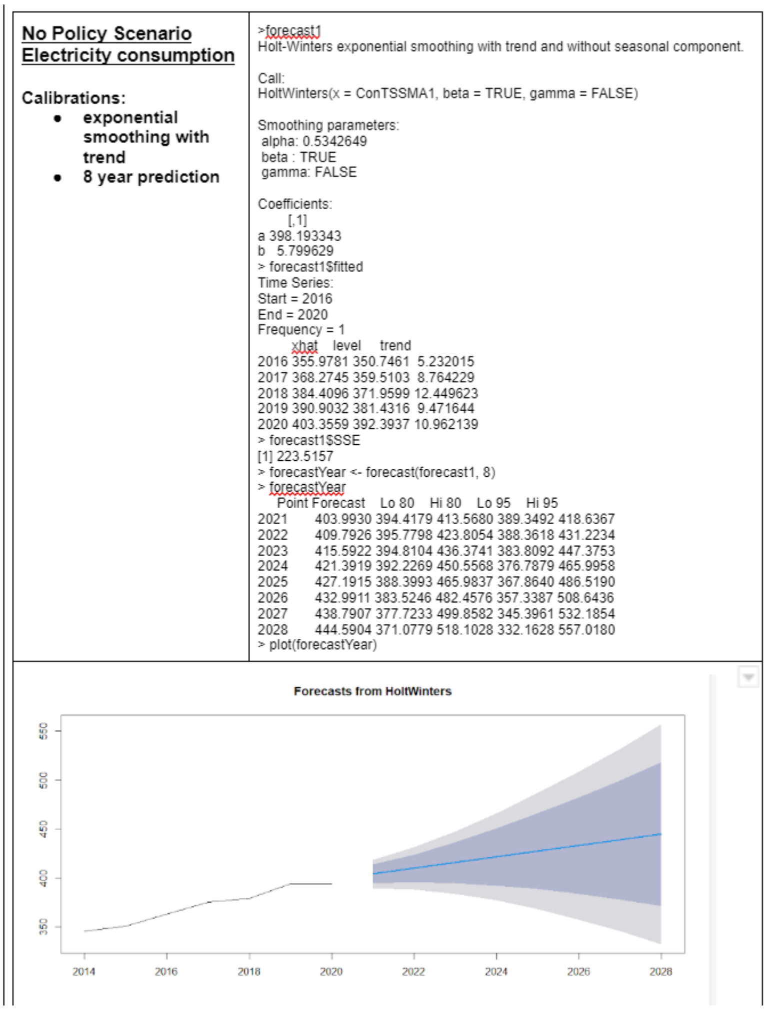


FIGURE 5.18: Case Study - No Policy Scenario - Template

between 87 and 116 GWh per year. See Figure 5.19.

Conversely the same scenario but with high efficient devices shows energy savings ranges between 206 to 274 GWh per year in the same 8 year period. See Figure 5.20.

Finally, the prototype was tested using real data of the Mexican government that was released in 2021. In the mentioned report, the government provide energy saving information about the implemented norms, which reveals that the high efficient device scenario is more aligned to reality that the other two. See Figure 5.21.

In our understanding, this last could be explained as follows. It´s possible that manufacturers create more efficient devices compared to the issued norm. If that was the case, it would be necessary to review the established norms, and update them. However, a specific research can be done on the subject to understand the causes and find logical responses to the case.

Normal Saving Scenario

Calibrations:

- **exponential smoothing with trend**
- **8 year prediction**

```
>forecast4
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = ConTSSMA4, beta = TRUE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.5342649
 beta : TRUE
 gamma: FALSE

Coefficients:
       [,1]
a 83.473687
b  1.215782
> forecast4$fitted
Time Series:
Start = 2016
End = 2020
Frequency = 1
        xhat    level    trend
2016 74.62406 73.52727 1.096793
2017 77.20178 75.36452 1.837255
2018 80.58418 77.97435 2.609827
2019 81.94545 79.95990 1.985551
2020 84.55591 82.25791 2.298005
> forecast4$SSE
[1] 9.822439
> forecastYear4 <- forecast(forecast4, 8)
> forecastYear4
     Point Forecast    Lo 80     Hi 80    Lo 95     Hi 95
2021       84.68947 82.68224  86.69670 81.61968  87.75926
2022       85.90525 82.96773  88.84278 81.41270  90.39781
2023       87.12103 82.76452  91.47755 80.45832  93.78375
2024       88.33682 82.22293  94.45070 78.98644  97.68719
2025       89.55260 81.42056  97.68464 77.11571 101.98948
2026       90.76838 80.39867 101.13809 74.90928 106.62748
2027       91.98416 79.18253 104.78580 72.40575 111.56258
2028       93.19995 77.78945 108.61044 69.63163 116.76826
```

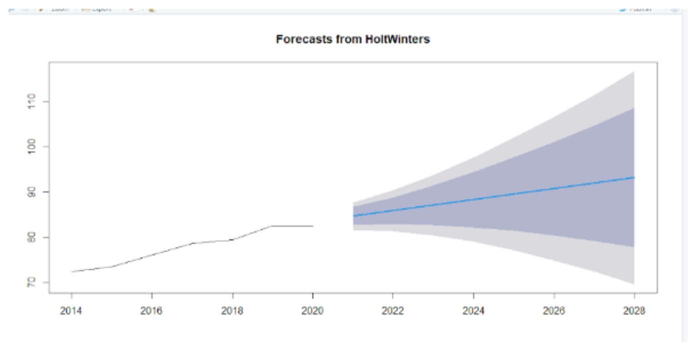


FIGURE 5.19: Case Study - Normal Saving Scenario

Best Saving Scenario

Calibrations:
- **exponential smoothing with trend**
- **8 year prediction**

```
> forecast5
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = ConTSSMA5, beta = TRUE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.5342649
 beta : TRUE
 gamma: FALSE

Coefficients:
       [,1]
a 196.362260
b   2.859988
> forecast5$fitted
Time Series:
Start = 2016
End = 2020
Frequency = 1
        xhat    level    trend
2016 175.5445 172.9645 2.580079
2017 181.6083 177.2864 4.321930
2018 189.5650 183.4257 6.139319
2019 192.7673 188.0965 4.670780
2020 198.9081 193.5023 5.405792
> forecast5$SSE
[1] 54.35458
> forecastYear5 <- forecast(forecast5, 8)
> forecastYear5
     Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2021       199.2222 194.5005 203.9440 192.0009 206.4436
2022       202.0822 195.1720 208.9924 191.5140 212.6505
2023       204.9422 194.6940 215.1904 189.2689 220.6155
2024       207.8022 193.4200 222.1844 185.8065 229.7979
2025       210.6622 191.5325 229.7919 181.4059 239.9185
2026       213.5222 189.1286 237.9157 176.2155 250.8289
2027       216.3822 186.2678 246.4966 170.3262 262.4381
2028       219.2422 182.9908 255.4936 163.8004 274.6839
```

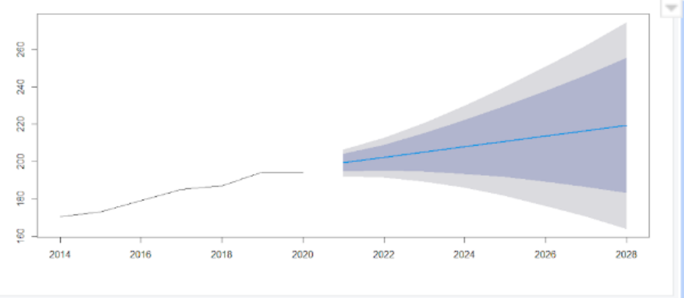


Figure 5.20: Case Study - Best Saving Scenario

Prototype testing

The prototype testing is performed by making a comparison of forecasted savings vs real metrics obtained from the government's recent reports.

```
> forecastYear5
     Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2021       199.2222 194.5005 203.9440 192.0009 206.4436
2022       202.0822 195.1720 208.9924 191.5140 212.6505
2023       204.9422 194.6940 215.1904 189.2689 220.6155
2024       207.8022 193.4200 222.1844 185.8065 229.7979
2025       210.6622 191.5325 229.7919 181.4059 239.9185
2026       213.5222 189.1286 237.9157 176.2155 250.8289
2027       216.3822 186.2678 246.4966 170.3262 262.4381
2028       219.2422 182.9908 255.4936 163.8004 274.6839
```

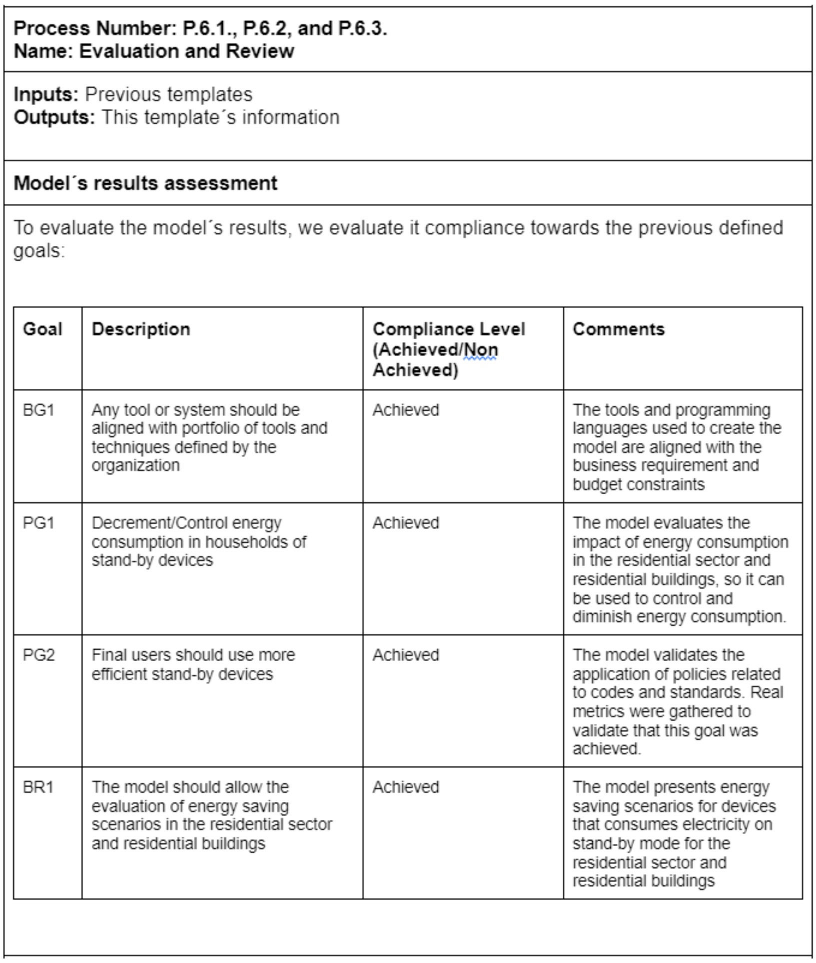
In this forecast, the model calculates the electricity savings range, which oscillates between 192 and 206 GWh with 95% probability. The real metrics reveal that on average microwave ovens norms save around 200.5GWh per year. This metric validates the results of the model.

NORMA	**Energy savings**
NOM-032-ENER-2013- All stand-by devices (2021)	**802.0**
Estimated savings of microwaves ovens (25% de standy-devices)	**802 *.25 = 200.5 GWh**

Figure 5.21: Case Study - Prototype Testing

5.6 Model Evaluation

In the final phase, the model is evaluated using the template visible in Figure 5.22. For this project, they indicate that all goals are met and provide comments to justify the assessment.



Process Number: P.6.1., P.6.2, and P.6.3.
Name: Evaluation and Review

Inputs: Previous templates
Outputs: This template´s information

Model´s results assessment

To evaluate the model´s results, we evaluate it compliance towards the previous defined goals:

Goal	Description	Compliance Level (Achieved/Non Achieved)	Comments
BG1	Any tool or system should be aligned with portfolio of tools and techniques defined by the organization	Achieved	The tools and programming languages used to create the model are aligned with the business requirement and budget constraints
PG1	Decrement/Control energy consumption in households of stand-by devices	Achieved	The model evaluates the impact of energy consumption in the residential sector and residential buildings, so it can be used to control and diminish energy consumption.
PG2	Final users should use more efficient stand-by devices	Achieved	The model validates the application of policies related to codes and standards. Real metrics were gathered to validate that this goal was achieved.
BR1	The model should allow the evaluation of energy saving scenarios in the residential sector and residential buildings	Achieved	The model presents energy saving scenarios for devices that consumes electricity on stand-by mode for the residential sector and residential buildings

Figure 5.22: Case Study - Model´s results assessment

The identification of strengths, weaknesses, opportunities and threats is visible in Figure 5.23.

Based on the experience of this project, it was recommended the creation of an organizational structure that support the process execution and the procurement of a budget for that objective.

Review Process

Strengths	**Weaknesses**
1. Provide organized steps with specific activities to perform and with specific data that should be recollected. 2. The process can be measured and improved using specific metrics 3. The documentation of the process is reinforced through the utilization of the templates	1. It takes time to fill out the templates of the framework 2. The information should be updated constantly since it is an iterative process.
Opportunities	**Threats**
1. The processes can be documented with flow chart diagrams 2. The diagrams can include specific roles that perform each activity	1. The framework requires that management, quality and technical teams work together. In order to achieve compliance. If the executed processes are not compliant with the framework's processes, the final model could be misaligned with the expected policy goals.

Determine next steps

1. Create an organizational structure that support the processes execution
2. Prepare the required budget

Figure 5.23: Case Study - Process Review and Next steps

In the next chapter, we analyze the processes execution of this framework, so we can evaluate the advantages and disadvantages of its application in a Decision Making Center.

Chapter 6

Results and Discussion

In the previous chapter, it was confirmed that the proposed framework is useful for creating BUEMs and to document the model's creation process. However, how can we know if this process is better than the de facto methodologies used in Decision-making centers?. In the following section, we analyze the results of the framework's execution and we compare them with the execution of de facto methodologies.

6.1 Review of framework's execution procedure

During each phase of the framework, the time of activities was registered using the free version of the Process Dashboard [1] application. Likewise, we use the Process Mining technique to extract and analyze the process execution. If fact, the authors has documented in [46], how this technique has been useful in different fields to discover processes, validate conformance, and to diminish process execution time.

The experimental design was performed similar as in authors' previous research [46]:

1. The process to analyze (e.g., the framework´s execution) is defined and their activities are registered in a Time Log.
2. The questions to be answered are specified.
3. An event Log is created using the Time Log registered in the Process Dashboard tool.
4. The Log is analyzed using a Process Mining Tool (e.g., Disco Application [2]).

[1] Process Dashboard is a support tool to execute the PSP (Personal Software Process). More information about the tool can be found in: https://www.processdash.com/

[2] Disco is application that uses event logs to analyze and discover process models [80]

5. The results are analyzed and reported.

For this analysis, we execute the process two times. First, for the control group (i.e., a process executed without the framework or de facto methodology) and later, for a process executed with the framework.

6.2 Control Group results

The data of the control group, that use the de facto methodology, was obtained from a project executed in 2019 in a Decision Making Center. We performed interviews to five modelers and one manager using the format A.8. The gathered information included activities from four CRISP-DM processes: data understanding, data preparation, modeling and evaluation), but we didn't have access to information from the business understanding and deployment phases [3]. As can be regarded in the format, we collected start and end dates of activities, average worked hours per task per day, and the number of resources per activity; so we can create a realistic event log.

After that, the researchers needed to adjust the control group data sources, so they could be comparable with the framework's processes. It was necessary to map the activities of the original de facto methodology (i.e., control group) with the activities of the new proposed framework. The result of these work is visible on Table 6.6.

Table 6.1: Activities to compare among frameworks

Activities (Proposed Framework)	**Activities (De Facto)**
Data acquisition	Collect Initial data, Explore Data, Select data
Data definition guidelines	Describe Data
Data Processing	Verify Data Quality, Clean Data, Construct Data, Integrate Data, Format Data
Create, update portfolio of design techniques	Not Applicable
Model´s build and testing	Generate Test Design, Build Model
Modeling techniques selection	Select Modeling Technique
Determine next steps	Determine next steps
Model´s results assessment	Assess Model
Review process	Review Process
Scenario definition	Build Model
Scenario design	Build Model

The event log revealed that the process was executed in 151 days and 8 hours and 248 events were executed. The global statistics of the process are available on Table 6.2 and the process execution is available on Figure 6.1.

Table 6.2: De Facto Process - Global Statistics

Metric	**value**
Events	248
Activities	8
Median case duration	151 days and 8 hours
Start	January 15th, 2019 9:00
End	June 15th, 2022 19:00

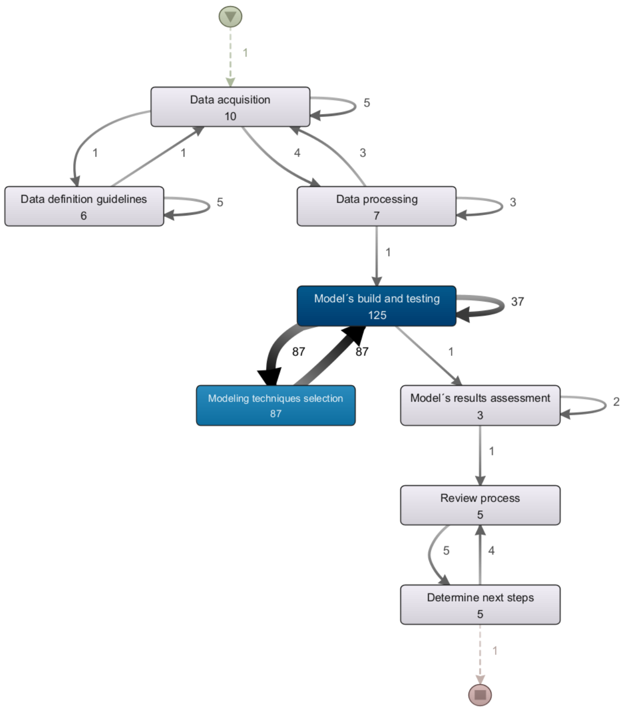


Figure 6.1: Execution Map - Control Group processes

It is evidenced in the diagram that the execution frequency of two activities is extremely high. For instance: (1) the model´s build and testing phase is executed in 50.4% of the cases and (2) the Modeling Technique Selection in the (35%). The remaining activities (i.e., Data acquisition, Data processing, Data definition guidelines, Review process, Determine next steps, and the model´s results assessment) are executed jointly in just the 14.6% of the cases. These facts reveal possible delays and bottleneck problems with these activities. In fact, it seems that there were problems with the selection of a feasible modeling technique for the development process.

6.3 Framework's execution results

The framework's processes were executed in 17 days and 21 hours, following the proposed steps and filling the proposed formats. In total there were executed 19 tasks recorded in 50 events.

See Table 6.3.

TABLE 6.3: Framework Process - Global Statistics

Metric	**value**
Events	50
Activities	19
Median case duration	17 day and 21 hours
Start	October 17th, 2022 19:21
End	November 4th, 2022 15:32

On the other hand, the frequencies of execution per activity are listed in Table 6.4. The activities that have more repetitions are the Model's build and testing (14%), followed by the Model's requirements and goal alignment (10%), and the Data acquisition (8%). This last make sense, since we are analyzing a new model. It is expected that the data acquisition phase has a high frequency, specially if the information is not provided (like in this case). Conversely, the Model´s build and testing has a iterative nature which explains the encountered repetitions.

TABLE 6.4: Frequencies of execution per activity

Activity	**Relative Frequency**
Model´s build and testing	14%
Model´s requirements and goal alignment	10%
Data acquisition	8%
Scenario definition	6%
Problem Understanding	6%
Business Requirements and Constraints	6%
Strategies to achieve objectives	6%
Instruments to implement	6%
Data definition guidelines	6%
Create, update portfolio of design techniques	6%
Goal Definition	4%
Policy level, aims and means	4%
Model's results assessment	4%
Review process	4%
Plan to achieve objectives	2%
Data processing	2%
Scenario design Modeling technique selection	2% 2%
Determine next steps	2%

The process execution is visible on Figure 6.2. The picture reveals that processes are iterative. Specifically three main execution cycles are identified. The first involves the activities of the Business Understanding phase. The second include tasks from the policy alignment phase, and the third contemplates activities from the data definition and understanding, scenario definition and design; and the model construction phases. The activities that have a strong dependency with others are the ones that have higher weights in the arrows. The greatest dependency is visible between the Business Requirements and constraints activity and the definition of strategies to achieve objectives. It is interesting to analyze one of the cycles that involves the data definition guidelines activity, and the model´s requirements definition and goal alignment; given that the model in theory should be impacted by previous design decisions; however according to the diagram, the model and the data definition can also affect the goals and requirements of the model.

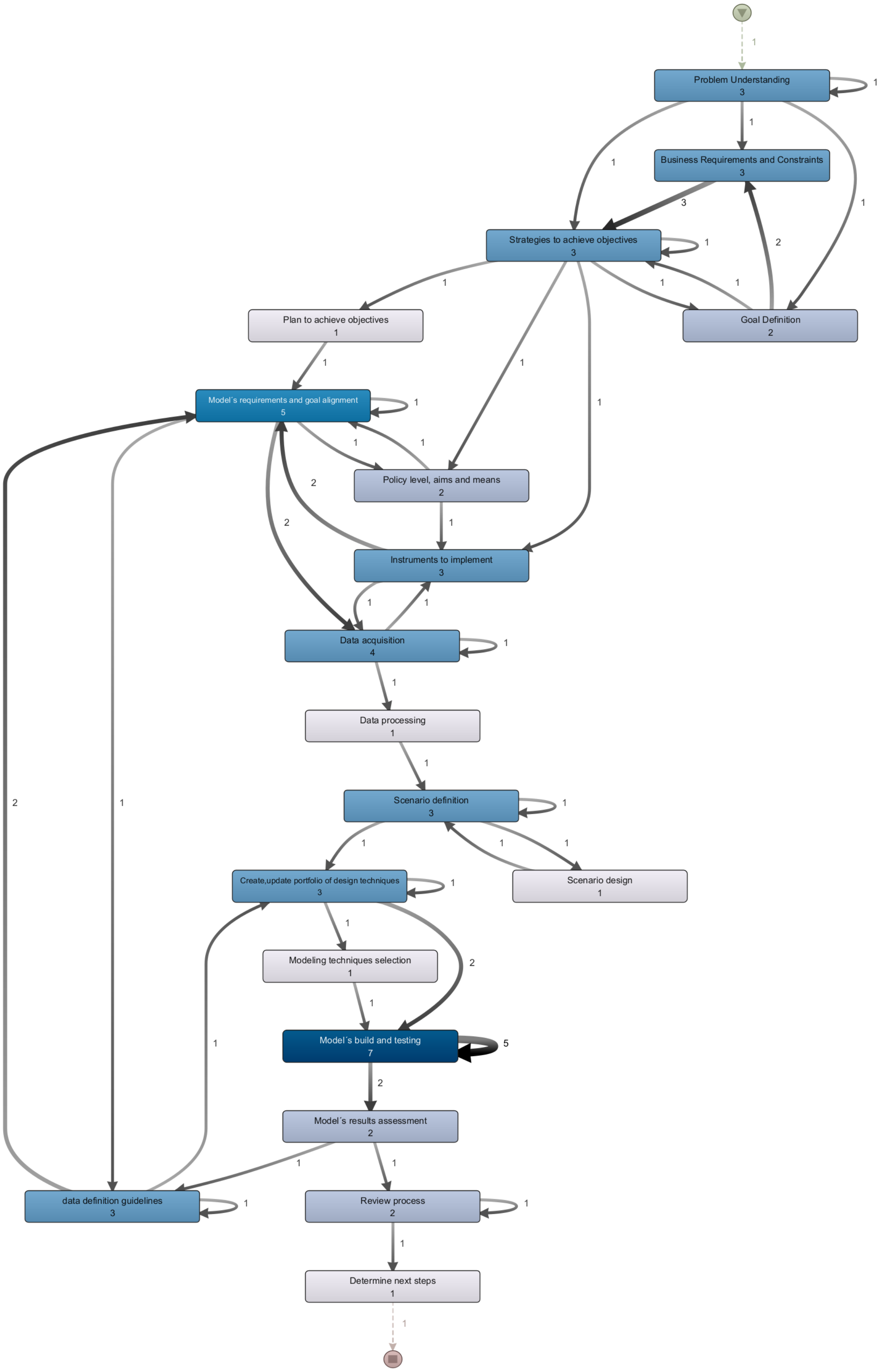


FIGURE 6.2: Execution Map - Framework processes

In addition, in the Performance view (See Figure 6.3) it is visible that the Data acquisition (5.3hrs), the Model´s build and testing (5hrs), the Problem Understanding (6 hrs), and the Model´s requirements and alignment (4.9 hrs) phases are the longest activities in the process execution. We attribute this results to the difficulty to find information from different sources, to the limited expertise of the model´s designers, and to the quantity of literature that was reviewed to understand the problem. The total duration of activities are available on Table 6.5. Lastly, there are two visible thick arrows in the diagram, that represent delays among activities. These lines should be monitored in order to avoid bottlenecks among activities during the process execution.

Table 6.5: Duration per activity

Activity	**Total duration (hrs)**
Problem Understanding	6.0
Data acquisition	5.3
Model´s build and testing	5
Model´s requirements and goal alignment	4.9
Instruments to implement	2.9
Scenario definition	2.7
Scenario design	2
Model's results assessment	1.96
Business Requirements and Constraints	1.83
Data definition guidelines	1.83
Review process	1.58
Create, update portfolio of design techniques	1.45
Modeling technique selection	1.4
Policy level, aims and means	1.23
Goal Definition	1.15
Plan to achieve objectives	1.08
Data processing	1
Strategies to achieve objectives	.98
Determine next steps	.7

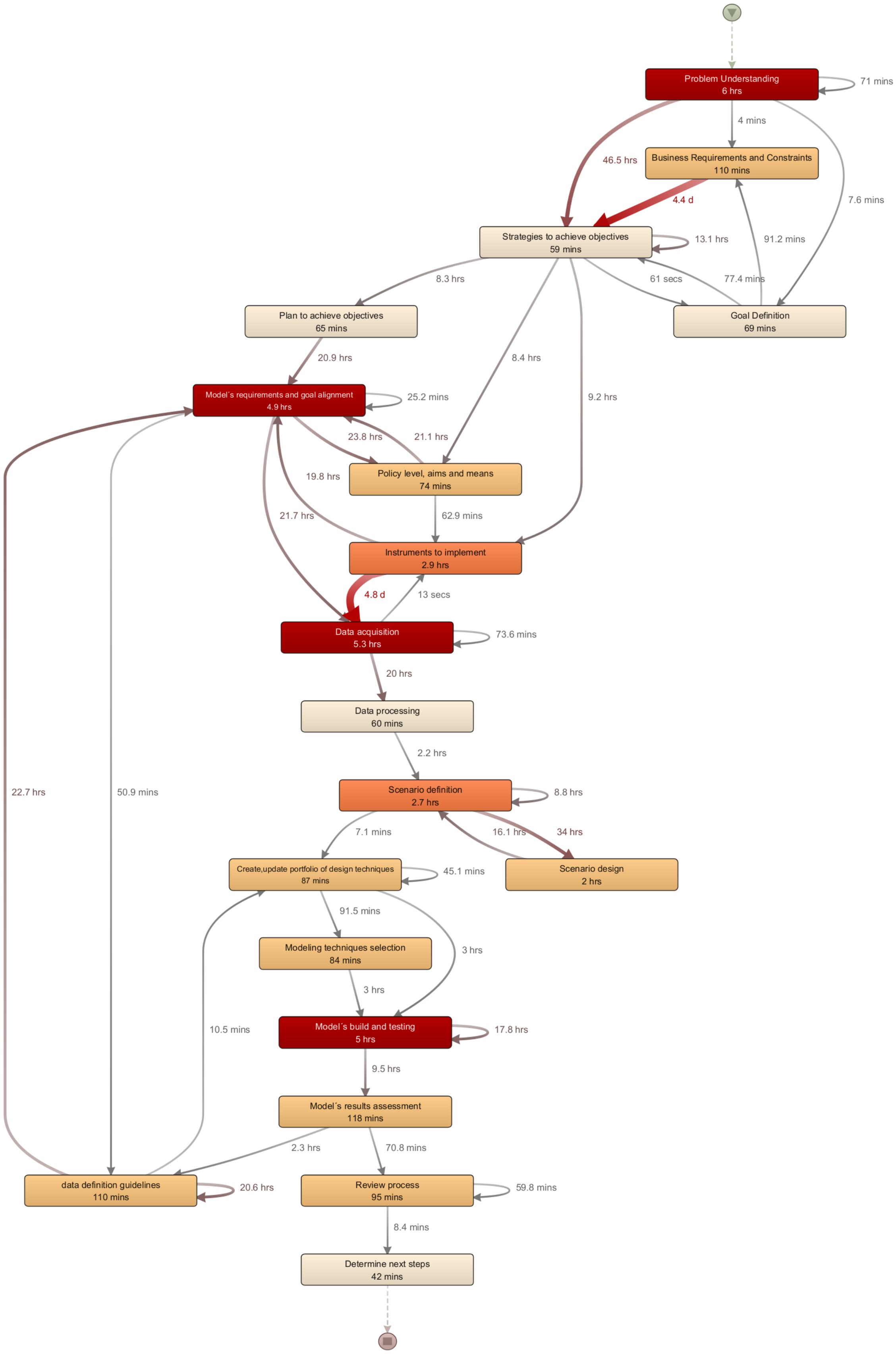


FIGURE 6.3: Performance Map - Framework processes

6.4 Comparison with other methodology

In this section, we compare the execution of the proposed framework with the execution of the de facto methodology used in Decision Making Centers [3] [45]. This methodology has been studied by the authors in previous research but it has never been compared quantitatively and qualitatively to other methods. With this comparison, we aim to validate the contribution of our framework to the creation of BUEMs that support policy design in the residential sector and residential buildings.

6.4.1 Model's comparison

The model comparison was performed with the process execution data from the new proposed framework and from the de facto methodology. We need to take a proportion of 7% of the control group data, in order to be able to compare the processes. The data used for the comparison is available on Table 6.6.

Table 6.6: Duration of Activities

Activity	**Total Duration (hrs) - (before proposed framework) 7% proportion**	**Total Duration (hrs) - with proposed framework (hrs)**
Data acquisition	1.8	3.2
Data definition guidelines	3.4	1.83
Data processing	1.9	1
Model´s build and testing	12	5
Modeling techniques selection	2	1.4
Determine next steps	1.4	0.7
Model´s results assessment	1.68	1.96
Review process	1.4	1.58
Scenario definition	5.8	2
Scenario design	3.4	2

The graphical representation of this comparison can be regarded on Figure 6.4.

The figure shows how the activity duration diminishes by using the proposed framework. Likewise, we can see that the magnitude of outliers is more measured, since the process execution is more controlled. The next significance test aims to demonstrate the time efficiency obtained by the utilization of the proposed framework.

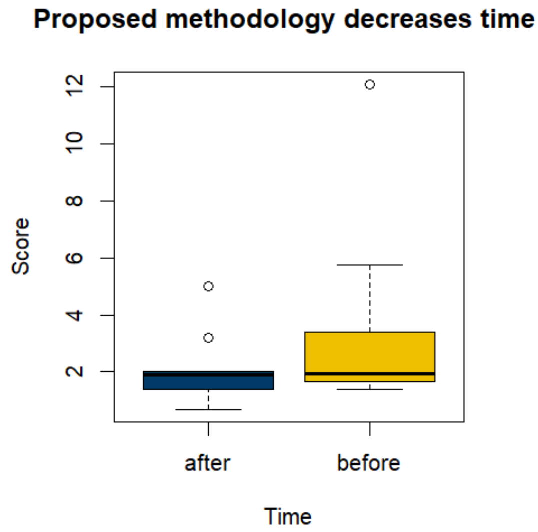


FIGURE 6.4: Comparison of activities duration per framework

6.4.2 Significance Test

We use a hypothesis test of paired samples to explore if there is a statistically significant difference in time duration because of the use of the proposed framework. The null hypothesis (H0 : mwith = mwithout) and alternative hypothesis (H1 : mwith mwithout) are defined in the following R code:

$t.test(formula = dat4\$score\ dat4\$time,$

$+ alternative = "less",$

$+ mu = 0,$

$+ paired = TRUE,$

$+ var.equal = FALSE,$

$+ conf.level = 0.95)$

$Pairedt - test$

$data : dat4\$scorebydat4\$time$

$t = -1.8484, df = 9, p - value = 0.0488$

$alternativehypothesis : truemeandifferenceislessthan0$

$95percentconfidenceinterval :$

$-Inf - 0.01166068$

$sampleestimates :$

$meandifference$

-1.41

Since the p-value of the test is less than .05, then we have enough evidence to reject the null hypothesis and we can affirm that on average the use of the proposed framework is more time efficient than the use of other de facto methodologies.

The results of the significance test are shown in Table 6.7.

TABLE 6.7: Hypotheses test of the paired samples of reported times with and without the proposed framework

Phase	**Avg Time Without Framework**	**Avg Time With Framework**	Difference in Times (h)	Hypotheses Test of Paired Samples (p-value)
All the phases in the framework	34.78	20.67	14.11	0.0488

6.4.3 Discussion

Hypothesis 1: The creation process of bottom-up energy models can be measured, monitored and improved to achieve process compliance and policy alignment.

In this research, we use diverse tools to test our hypothesis. First we use the process Dashboard application to measure all the activities executed during the model´s construction and later we use the Process Mining technique (specifically the Disco Application) to validate the process execution and its compliance. Likewise, in this chapter, we make a significance test to validate the time efficiency obtained by the utilization of the proposed framework. So far, the possibility to measure and monitor processes has brought the following advantages:

- The activities can be compared, analyzed, and changed if necessary based on specific metrics.
- The process itself promotes policy alignment through the framework processes and templates.
- The processes can be improved based on detailed quantitative feedback.
- Time efficiency can be measured per activity or phase to apply specific technical or management strategies.

Hypothesis 2: The creation process of bottom-up energy models can be adapted to guide the design of scenarios.

The definition and design of scenarios is one of the key phases of the proposed framework, not only because it help to bound the scope of the model, but also because it represents the first effort to construct a functional prototype of the model. So far the execution of the new Scenario design phase has exhibited the following advantages:

- It is possible to execute the processes and activities even if the stakeholder is not an expert on the subject.
- The detailed documentation of the scenario definition and design allows the exploration of new scenarios variables, and environments.
- The scenarios can be modified and even the processes to create them, based on the necessities of the business and the stakeholders.

Additionally, in the previous study case, we validate the utility of the scenario design phase with the Evaluation and Review Template (with a binary -yes or no - metric), however a more detailed assessment is necessary to assess specific features of the scenario for future iterations of the framework.

To conclude it is important to recognize that in the state of the art, there are technical methodologies to create models, but none of them focus on the decision-makers point of view. Our framework aims to interrupt that dynamic by offering a tool that not only provides guidelines but also that promotes the communication and the agreement on important decisions among people. The framework can be useful for energy policy-makers, energy model designers, energy efficiency experts, and managers that want to create models with a mature methodology.

Chapter 7

Conclusions and Future Work

The following research has presented a framework to facilitate the creation of bottom-up energy models (BUEMs) that aim to support policy design of electricity end-use efficiency for the residential sector and residential buildings. The framework includes six phases: business understanding, policy alignment, data definition and understanding, scenario definition and design, model construction and model evaluation. The model promotes best practices like:

1. The use of a mature process to build BUEMS which can be analyzed and improved,
2. the detailed specification of the different goals to be achieved,
3. the application of techniques that allow the construction of energy efficiency scenarios, and
4. the inclusion of key metrics to support policy design, and the documentation of the model's construction for future analysis and maintenance.

We obtained these best practices from previous research work [3] [52], where the opportunity areas of BUEM's design were revealed.

The framework has been applied, with favorable results, in the creation of a model that supports the design of end-use electricity policies for stand-by devices in the Mexican residential sector. During the case study execution, the framework allowed:

1. the model's alignment to business, stakeholders, and policy goals,
2. the documentation of the model and it's creation processes,
3. the application of best practices for the creation of BUEMs than aim to support policy design of electricity end-use efficiency, and

4. the support of different kinds of models (e.g., bottom-up and top-down).

On the other hand, the process mining technique has been applied to the framework's processes to identify the activities that are executed, their frequency and time performance. In fact, our research hypothesis was successfully proved through a quantitative comparison of real energy models, using this technique, which actually has revealed the advantages of the framework utilization.

The contributions of this research to the literature are:

1. A framework that promote best practices for the development of BUEMs that aim to support policy design,
2. a complete taxonomy to create bottom-up energy models that can be used to compare and evaluate different models,
3. a process model that allow the comparison and improvement of processes to create BUEMs, and
4. a framework that can be expanded and applied to other energy sectors and to other kinds of models (e.g., top-down models).

Lastly, it has been mentioned all the advantages of the framework, however there are important aspects that should be considered for future research: First, it must be studied new ways to gather event logs of the processes execution. This subject is crucial in order to calculate with precision the real efficiency of an executed process. If there is no maturity in the data gathering of processes, the information could no reveal truthful data. Second, it is necessary to include additional best practices for model creation to the framework (e.g., key metrics and model evaluation techniques), and quality assurance processes. Third, it should be implemented a mechanism to document errors (e.g., an error log) during the process execution, so their appearance can be analyzed and avoided in future iterations of the framework's execution. Finally, the processes could be improved with the addition of execution roles and flow chart diagrams, which were not possible to include given time restrictions. The execution roles can be documented using swimlanes in a process flow diagram, so the responsibilities can be clearly specified.

Appendix A

DataSources Documentation

ID	1
Data source Name	Number of mexican households with stand-by devices (e.g., microwave ovens) 2014-2020
Description	Publication time: 2014-2020
Owner	INEGI
Location	2014 - Encuesta Nacional en los Hogares https://www.inegi.org.mx/programas/enh/2014/#Microdatos 2015 - Encuesta Nacional en los Hogares https://www.inegi.org.mx/programas/enh/2015/#Microdatos 2016 - Encuesta Nacional en los Hogares https://www.inegi.org.mx/programas/enh/2016/#Microdatos 2017 - Encuesta Nacional en los Hogares https://www.inegi.org.mx/programas/enh/2017/#Tabulados 2018 - Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). 2018 Nueva serie https://www.inegi.org.mx/programas/enigh/nc/2018/#Microdatos 2020 - Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). 2020 Nueva serie https://www.inegi.org.mx/programas/enigh/nc/2020/#Microdatos
Updating Policy (Which, how, how much info)	Update yearly the metric (if the survey is available)
Type of data source (structured/semi/non-structured)	Structured (Table)
Number of records	(1) 44740 (2) 59535 (3) 59080 (4) 57520 (5) 74648 (6) 89007
Data Time Split	Yearly
Initial/final record dates	2014-2020

<u>**Metadata Documentation**</u>

Type of data (file/JSON/table/other)	electronic document
Name of artifact	MexicanHouseholdsWithStand-by devices

Field name	Field Type	Value range	Required	Business Metric
Folio_Hogar	numeric	100,000	yes	N/A
Horno de Microondas	numeric	1-5	no	N/A

Figure A.1: Case Study- Data Source 1

ID	2
Data source Name	Regulations of stand-by devices
Description	Publication time: 23/01/2014
Owner	Secretaría de Gobernación
Location	https://www.dof.gob.mx/nota_detalle.php?codigo=5330530&fecha=23/01/2014#gsc.tab=0
Updating Policy (Which, how, how much info)	Not applicable. Monitor if an policy is updated in "Diario Oficial de la Federación"
Type of data source (structured/semi/non-structured)	unstructured (document)
Number of records	1
Data Time Split	Not applicable
Initial/final record dates	2014

Metadata Documentation

Type of data (file/JSON/table/other)	Document
Name of artifact	PotenciaElectricaMáxima-StandupDevices

Field name	Field type	Value range	Required	Business Metric
Appliance	Character	char (50)	yes	N/A
Stand-by power	Float	0-10.00	yes	W(Watts)

Tabla 4. Potencia eléctrica máxima en modo de espera de equipos para hornos de microondas

Horno de microondas	Potencia eléctrica máxima en modo de espera (W)
Convencionales	2.50
Combinados	5.00
Empotrables	5.00

FIGURE A.2: Case Study- Data Source 2

ID	3
Data source Name	Current efficiency of stand-by devices
Description	Publication time: 23/01/2014
Owner	International Journal of Emerging Electric Power Systems
Location	https://www.researchgate.net/publication/260585939_Reduce_the_Standby_Power_Consumption_of_a_Microwave_Oven
Updating Policy (Which, how, how much info)	Updated technical characteristics of devices should be reviewed with manufacturers
Type of data source (structured/semi/non-structured)	unstructured (document)
Number of records	1
Data Time Split	Not applicable
Initial/final record dates	2013

Metadata Documentation

Type of data (file/JSON/table/other)	table
Name of artifact	EficienciaActual-stand-by devices

Field name	Field type	Value range	Required	Business Metric
Appliance	Character	char(50)	yes	N/A
Stand-by power	Number (decimal)	0-10.00	yes	W(Watts)

Appliance	Stand-by power
Horno de Microondas	1 - 3

Figure A.3: Case Study- Data Source 3

ID	4
Data source Name	Average number of stand-by devices per household (CONUEE) - Specifically microwaves)
Description	Publication time: 2018
Owner	CEPAL, CONUEE, ADEME, Cooperación alemana
Location	https://www.cepal.org/es/publicaciones/43612-informe-nacional-monitoreo-la-eficiencia-energetica-mexico-2018
Updating Policy (Which, how, how much info)	Monitor metrics published by CEPAL or CONUEE
Type of data source (structured/semi/non-structured)	unstructured (document)
Number of records	1
Data Time Split	Not applicable
Initial/final record dates	2018

Metadata Documentation

Type of data (file/JSON/table/other)	table
Name of artifact	AverageofStand-by devicesPerHousehold

Field name	Field type	Value range	Required	Business Metric
Appliance	Character	char(50)	yes	N/A
Stand-by power	Number (decimal)	0-10.00	yes	W(Watts)

Appliance	Microwaves per household
Horno de Microondas	1

Figure A.4: Case Study- Data Source 4

ID	5
Data source Name	Average Life-span of microwave ovens
Description	Publication: 2019
Owner	Applied Energy
Location	Researchgate https://www.researchgate.net/publication/330998256_Analysis_of_the_impact_of_energy_efficiency_labelling_and_potential_changes_on_electricity_demand_reduction_of_white_goods_using_a_stock_model_The_case_of_Switzerland
Updating Policy (Which, how, how much info)	Monitor metrics published by the literature
Type of data source (structured/semi/non-structured)	unstructured (document)
Number of records	1
Data Time Split	yearly
Initial/final record dates	2019

Metadata Documentation

Type of data (file/JSON/table/other)	Table
Name of artifact	Average Life-span of microwave ovens

Field name	Field type	Value range	Required	Business Metric
Life-span (years)	Number (decimal)	0-100	yes	years

Appliance	Average lifespan
Oven	2.5 - 13.8

Figure A.5: Case Study- Data Source 5

ID	6
Data source Name	Saturation rates
Description	Average saturation rate of ovens, percentage of saturation of efficient and non-efficient microwave ovens.
Owner	CEPAL, CONUEE, ADEME, Cooperación alemana
Location	https://www.cepal.org/es/publicaciones/43612-informe-nacional-monitoreo-la-eficiencia-energetica-mexico-2018
Updating Policy (Which, how, how much info)	Monitor metrics published by CEPAL or CONUEE
Type of data source (structured/semi/non-structured)	unstructured (document)
Number of records	3 and 2
Data Time Split	yearly
Initial/final record dates	2015

Metadata Documentation

Type of data (file/JSON/table/other)	Table
Name of artifact	Average saturation

Field name	Field type	Value range	Required	Business Metric
Year	decimal	number(4)	yes	N/A
Average Saturation of microwave ovens	Float	0-100	yes	saturation rate
Percentage of saturation of microwave ovens (<= 10 years)	Float	0-100	yes	saturation rate
Percentage of saturation of efficient microwave ovens(> 10 years)	Float	0-100	yes	saturation rate

Year	Saturation rate
1996	13.3%
2006	38.4%
2015	45.5%
Average	1.6% / year

Year	Percentage of saturation (Microwaves <=10 years)	Percentage of saturation (Microwaves >10 years)older than 5 years
2015	79.2%	20.8%

Figure A.6: Case Study- Data Source 6

ID	7
Data source Name	SavingsByNormApplication
Description	Estimated savings of norms and standards 2021
Owner	SENER (Secretaría de Energía) CONUEE (Comisión Nacional de Uso Eficiencia de Energía)
Location	https://www.gob.mx/cms/uploads/attachment/file/739347/BALANCE_NOM-ENER_al_2021_ver_final_05072022__1_.pdf
Updating Policy (Which, how, how much info)	Yearly (At the end of the year)
Type of data source (structured/semi/non-structured)	unstructured (Document)
Number of records	1
Data Time Split	yearly
Initial/final record dates	2021

Metadata Documentation

Type of data (file/JSON/table/other)	Table
Name of artifact	Consumo de electricidad en el sector residencial

Field name	Field type	Value range	Required	Business Metric
Norm	Character	char (50)	yes	N/A
Energy Consumption	Float	0-2000	yes	GWh

NORMA	Energy consumption
NOM-032-ENER-2013- Energía en espera	802.0

FIGURE A.7: Case Study- Data Source 7

Project Name					
Start date		**End date**			
Phase: Data Understanding					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Collect initial data					
Describe data					
Explore data					
Verify data quality					
Phase: Data Preparation					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Select data					
Clean data					
Construct data					
Integrate data					
Format data					
Phase: Modeling					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Select modeling technique					
Generate test design					
Build model					
Assess model					
Phase: Evaluation					
Activity	**Start date**	**End date**	**Hours/day/rol**	**Resource**	**Order**
Evaluate results					
Review Process					
Determine Next Steps					

FIGURE A.8: Case Study- Data Gathering Format (Control Group)

Appendix B

Published peer-reviewed works

energies MDPI

Review

The Contribution of Bottom-Up Energy Models to Support Policy Design of Electricity End-Use Efficiency for Residential Buildings and the Residential Sector: A Systematic Review

Marlene Ofelia Sanchez-Escobar [1,*], Julieta Noguez [1], Jose Martin Molina-Espinosa [1], Rafael Lozano-Espinosa [1] and Genoveva Vargas-Solar [2]

1 Tecnologico de Monterrey, School of Engineering and Science, Ave. Eugenio Garza Sada 2501, Monterrey 64849, NL, Mexico; jnoguez@tec.mx (J.N.); jose.molina@tec.mx (J.M.M.-E.); razolano@tec.mx (R.L.-E.)
2 LIRIS UMR5205, CNRS, 69621 Villeurbanne, France; genoveva.vargas-solar@liris.cnrs.fr
* Correspondence: A00704709@itesm.mx; Tel.: +521-5544995733

Abstract: Bottom-up energy models are considered essential tools to support policy design of electricity end-use efficiency. However, in the literature, no study analyzes their contribution to support policy design of electricity end-use efficiency, the modeling techniques used to build them, and the policy instruments supported by them. This systematic review fills that gap by identifying the current capability of bottom-up energy models to support specific policy instruments. In the research, we review 192 publications from January 2015 to June 2020 to finally select 20 for further examination. The articles are analyzed quantitatively in terms of techniques, model characteristics, and applied policies. The findings of the study reveal that: (1) bottom-up energy models contribute to the support of policy design of electricity end-use efficiency with the application of specific best practices (2) bottom-up energy models do not provide a portfolio of analytical methods which constraint their capability to support policy design (3) bottom-up energy models for residential buildings have limited policy support and (4) bottom-up energy models' design reveals a lack of inclusion of key energy efficiency metrics to support decision-making. This study's findings can help researchers and energy modelers address these limitations and create new models following best practices.

Keywords: energy modelling; electricity efficiency; energy policy; residential buildings; households; data-driven approach

check for updates

Citation: Sanchez-Escobar, M.O.; Noguez, J.; Molina-Espinosa, J.M.; Lozano-Espinosa R.; Vargas-Solar, G. The Contribution of Bottom-Up Energy Models to Support Policy Design of Electricity End-Use Efficiency for Residential Buildings and the Residential Sector: A Systematic Review. *Energies* **2021**, *14*, 6466. https://doi.org/10.3390/en14206466

Academic Editors: Carlos Henggeler Antunes and Javier Reneses

Received: 16 June 2021
Accepted: 24 August 2021
Published: 10 October 2021

1. Introduction

The 2019 world energy outlook anticipates that the building sector (including households and services) will continue being the main contributor to global electricity demand by 2040 [1]. Alone households expect a 60% increase in electricity consumption in developing economies. The outlook associates the possible increment in the buildings' electricity demand by utilizing more air conditioners, household appliances, and electric vehicles. On the other hand, since the increase in energy use is related to the increment of carbon dioxide (e.g., CO_2) emissions, it is advised in the literature to enhance the efficiency of electricity use [2]. Thus, it is possible to reduce emissions without compromising the development of electricity services. To this end, governments design policies that promote electricity end-use efficiency [2] with the support of bottom-up energy models [3].

Policy design refers to the selection of policy tools to achieve energy efficiency objectives [4], eliminate barriers toward efficiency, and gain energy efficiency benefits [5]. The design of policies for electricity end-use efficiency focus on the factors that distort market and restrict the adoption of efficient technologies [2]. For instance, expected short payback period on investments, uncertainty about actual savings, the lack of trained people to invest in energy efficiency, physical barriers of the technology, attributes that affect

Energies **2021**, *14*, 6466. https://doi.org/10.3390/en14206466 https://www.mdpi.com/journal/energies

Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 1084-1090 (2021)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering
ASTESJ
ISSN: 2415-6698

Supporting the Management of Predictive Analytics Projects in a Decision-Making Center using Process Mining

Marlene Ofelia Sanchez-Escobar*, Julieta Noguez, Jose Martin Molina-Espinosa, Rafael Lozano-Espinosa

School of Engineering and Science, Tecnologico de Monterrey, Mexico City, 14380, Mexico

ARTICLE INFO

Article history:
Received: 11 January, 2021
Accepted: 12 April, 2021
Online: 28 April, 2021

Keywords:
Process design
Process modeling
Project Management
Data Mining
Best Practices
Decision making

ABSTRACT

A Decision-Making Centers (DMCs) Environment facilitates stakeholders' decision-making processes using predictive models and diverse what-if scenarios. An essential element of this environment is the management of Decision Support Components (e.g., models or systems) that need to be created with mature methodologies and good delivery time. However, there has been a gap in the understanding of project management best practices in DMC environments and in the application of methodologies to ease project execution. In the following paper, we address that gap by analyzing six predictive analytics projects executed in a Mexican DMC using Process Mining techniques. We perform process discovery using a detailed activity event log, which has not been possible in previous studies. Additionally, we perform a compliance evaluation versus the de facto methodology to identify the current process alignment gaps, and finally, we analyze the social networks present in the process execution. The research reveals that (1) process mining models are helpful to address management issues of PA/DM projects (2) PA/DM projects require alignment to mature methodologies to improve process performance and avoid execution problems (3) PA/DM project execution should be revised at the activity level to identify issues and to propose specific strategies. This study's findings can help project managers to perform process analyses and to make informed decisions in PA/DM projects. The following paper is an extension of the article "Applying Process Mining to Support Management of Predictive Analytics/Data Mining Projects in a Decision-Making Center" presented in the 2019 International Conference on Systems and Informatics (ICSAI 2019).

1. Introduction

Decision-Making Centers (DMCs) are immersive virtual environments used to understand complex problems, simplify decision-making, and visualize the results of predictive and scenario-based models [1]. These environments depend on the creation process of tools like: Predictive Analytics/Data Mining (PA/DM) models to operate [2]. Nevertheless, the authors have demonstrated in previous studies that DMC processes focus on high-level tasks and exclude detailed and standard PA/DM activities [2]. The absence of commonality in PA/DM project execution, generates issues, since (1) models are built using empiric methodologies and (2) managers cannot follow up specific technical activities since they are different in every project.

*Corresponding author: Marlene Ofelia Sanchez-Escobar, Tel.: +521 5544995733, Email address: A00704709@itesm.mx

In this research, we propose three approaches to overcome the mentioned issues and help managers and modelers make informed decisions about PA/DM projects. In the first, we apply process mining techniques to a set of PA/DM processes to discover the timing, flow, frequency, and performance of activities from diverse perspectives (e.g., process, organizational, and case). Second, we compare a real PA/DM project execution with an accepted PA/DM methodology, to identify how aligned are the real processes to the formal methodology (i.e., CRISP-DM) and what gaps need to be closed to achieve compliance. Third, we perform complementary human resources analyses to visualize the relationship between resources and communication channels during process execution.

We expect that managers in DMCs use the models presented in this study to evaluate their processes and to consider the implementation of specific management strategies.

The 2019 6th International Conference on Systems and Informatics (ICSAI 2019)

Applying Process Mining to Support Management of Predictive Analytics/Data Mining Projects in a Decision Making Center

Marlene Ofelia Sanchez Escobar
School of Engineering and Science
Tecnologico de Monterrey
A00704709@itesm.mx

Rafael Lozano Espinosa
School of Engineering and Science
Tecnologico de Monterrey
ralozano@tec.mx

Jose Martin Molina Espinosa
School of Engineering and Science
Tecnologico de Monterrey
jose.molina@tec.mx

Juana Julieta Noguez Monroy
School of Engineering and Science
Tecnologico de Monterrey
jnoguez@tec.mx

Genoveva Vargas Solar
Centre National de la Recherche Scientifique
Laboratory of Informatics of Grenoble
Genoveva.Vargas@imag.fr

***Abstract*—A Decision Making Centers (DMC) Environment facilitates the understanding of complex problems and simplifies the decision making process of policy makers and stakeholders by playing out several what if scenarios. One of the key elements in this environment is the management of Predictive Analytics and Data Mining (PADM) techniques required for real-time and accurate representation of data in scenarios. However, there has been a gap in understanding the best practices that facilitate the execution of PADM in environments such as a DMC. In this paper we aim to address this gap. Specifically, we apply process mining techniques to reveal: 1) process flow and executed activities in PADM projects and 2) The relationship and time distribution of activities within a PADM project execution . To this end two successful projects were analyzed in order to find relationships among activities and time metrics during execution. The present study aims to support managers in the execution of large scale PADM projects.**

***Index Terms*—Process design, Process modeling, Business process management, Project management, Data mining, Best practices, Decision making**

I. Introduction

Project Management (PM) has been studied for the areas of Data Mining and Predictive Analytics (DM/PA) since it is considered a key factor to successful project execution[1]. In fact, problems like: excessive time devoted to certain phases in DM process, poor data quality [2], lack of methodological support, work repetition, and complication in resource management for DM applications [1], has been documented by scholars as key points to be addressed in DM/PA projects. In this regard, to overcome these problems researchers have proposed the implementation of different strategies, for instance: Becker and Ghedini[1] propose the creation of a documentation infrastructure that support different methodologies during DM project execution. Likewise, in [2] Diop, Camara, Bah and Fall center their research in diminishing the time devoted to data preparation phase by means of a prior quality approach applied to a DM process. Nevertheless, studies that focus on the analysis and diagnosis of the entire execution[1] of a DM process within a project have not yet been conducted. So, this is where the following paper has its first contribution. We aim to generate a model of a complete DM process by means of a Process mining technique in order to have a real perspective of a DM process execution within a project.

We use Process Mining because this technique not only allow the discovery of process models from event logs[3], but also because it allows to study processes from different perspectives (e.g. process, organization or case)[4]. In our experiment we focus in two of them: (1)the process perspective since we are looking to discover the flow of activities and (2) the case perspective since we want to understand the properties of processes execution.

It is important to mention that the following experiment was developed in a Decision Making Center (DMC), which is defined as:

> a room with big format screens distributed in semi-circle that creates an environment for visualization of complex process models and that offers support for decision making.[5]

An example of the implementation of this concept is Arizona State University's Decision Theater®, where this environment is regarded as a tool that involves stakeholders in a decision-making process and that supports participatory planning[6].

[1]The entire execution of a Data Mining process contemplates the accomplishment of the phases proposed in a process framework. For instance, CRISP-DM framework recommends the execution of 6 phases with their respective activities.

978-1-7281-5255-1 ©2019 IEEE

Bibliography

[1] (IEA) International Energy Agency. World Energy Outlook 2019, 2019.

[2] Mark D. Levine, Jonathan G. Koomey, Lynn Price, Howard Geller, and Steven Nadel. Electricity end-use efficiency: Experience with technologies, markets, and policies throughout the world. *Energy*, 20(1):37–61, 1995.

[3] Rafael Lozano-Espinosa Marlene Ofelia Sanchez-Escobar, Julieta Noguez, Jose Martin Molina-Espinosa. Supporting the Management of Predictive Analytics Projects in a Decision-Making Center using Process Mining. *Advances in Science, Technology and Engineering Systems Journal*, 6(2):1084–1090, 2021.

[4] H. Doukas, K. Patlitzianas, A. Kagiannas, and John Psarras. Energy policy making: An old concept or a modern challenge? *Energy Sources, Part B: Economics, Planning, and Policy*, 3, 09 2008.

[5] Wilfrid L. Kohl. National security and energy. In Cutler J. Cleveland, editor, *Encyclopedia of Energy*, pages 193–206. Elsevier, New York, 2004.

[6] Andrew Chapman, Benjamin Mclellan, and Tetsuo Tezuka. Strengthening the energy policy making process and sustainability outcomes in the oecd through policy design. *Administrative Sciences*, 6:9, 07 2016.

[7] Léonore Halg, Sebastian Sewerin, and Tobias Schmidt. The role of actors in the policy design process: introducing design coalitions to explain policy output. *Policy Sciences*, 53:1–39, 06 2020.

[8] André Schaffrin, Sebastian Sewerin, and Sibylle Seubert. Toward a comparative measure of climate policy output. *Policy Studies Journal*, 43(2):257–282, 2015.

[9] Hiroshan Hettiarachchi and Chandrashekar Kshourad. Chapter 9 - promoting waste-to-energy: Nexus thinking, policy instruments, and implications for the environment. In Sunil Kumar, Rakesh Kumar, and Ashok Pandey, editors, *Current Developments in Biotechnology and Bioengineering*, pages 163–184. Elsevier, 2019.

[10] Frédéric Varone and Bernard Aebischer. Energy efficiency: the challenges of policy design. *Energy Policy*, 29(8):615–629, 2001.

[11] (IEA) International Energy Agency. Energy Efficiency Indicators 2020, 2020.

[12] Elizabeth Doris, Jaquelin Cochran, and Martin Vorum. Energy efficiency policy in the united states: Overview of trends at different levels of government. 01 2009.

[13] Kenneth Gillingham, Richard G. Newell, and Karen L. Palmer. Retrospective examination of demand-side energy efficiency policies. (1318-2016-103264):100, 2004.

[14] Catharina Wiese, Anders Larsen, and Lise-Lotte Pade. Energy efficiency policy: A review of instruments and potential interaction effects. In *Meeting the Energy Demands of Emerging Economies, 40th IAEE International Conference, June 18-21, 2017*. International Association for Energy Economics, 2017.

[15] (IEA) International Energy Agency. World energy balances: Overview, 2020.

[16] Lisa C. Schwartz, Max Wei, William R. Morrow III, Jeff Deason, Steven R. Schiller, Greg Leventis, Sarah Josephine Smith, Woei Ling Leow, Todd Levin, Steven Plotkin, Yan Zhou, and Joseph Teng. Electricity end uses, energy efficiency, and distributed energy resources baseline. Technical report, 01/2017 2017.

[17] Peter G. Taylor, Olivier Lavagne d Ortigue, Michel Francoeur, and Nathalie Trudeau. Final energy use in iea countries: The role of energy efficiency. *Energy Policy*, 38(11):6463 – 6474, 2010. Energy Efficiency Policies and Strategies with regular papers.

[18] Sara Moarif and Namrata Patodia Rastogi. *Market-based climate mitigation policies in emerging economies*. Center for Climate and Energy Solutions, 2012.

[19] Pawel Bartoszczuk. Tradable emission permits as efficient strategy for achieving environmental goals. 01 2004.

[20] Ugo Farinelli, Thomas B. Johansson, Kes McCormick, Luis Mundaca, Vlasis Oikonomou, Mattias Örtenvik, Martin Patel, and Federico Santi. "white and green": Comparison of market-based instruments to promote energy efficiency. *Journal of Cleaner Production*, 13(10):1015 – 1026, 2005. 10th Anniversary of the International Institute for Industrial Environmental Economics (IIIEE): Management and policy for sustainable consumption and production.

[21] Joseph Iwaro and Abraham Mwasha. A review of building energy regulation and policy for energy conservation in developing countries. *Energy Policy*, 38(12):7744 – 7755, 2010. Special Section: Carbon Reduction at Community Scale.

[22] Wokje Abrahamse, Linda Steg, Charles Vlek, and Talib Rothengatter. A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3):273 – 291, 2005.

[23] Silvia Rezessy, Paolo Bertoldi, and Agneta Persson. Are voluntary agreements an effective energy policy instrument? Insights and experiences from Europe. *American Council for Energy Efficient Economy (ACEEE) Summer Study on energy efficiency in industry*, 2005.

[24] Andrea Herbst, Felipe Toro, Felix Reitze, and Eberhard Jochem. Introduction to energy systems modelling. *Swiss Journal of Economics and Statistics*, 148(2):111–135, 2012.

[25] N. M. J. P. van Beeck. *Classification of Energy Models*, volume 777 of *FEW Research Memorandum*. Operations research, Tilburg, 1999.

[26] Dominic White, Niven Winchester, Martin Atkins, John Ballingall, Simon Coates, Ferran Mercader, Suzie Greenhalgh, Andrew Kerr, Suzi Kerr, Jonathan Leaver, Catherine Leining, Juan Monge, James Neale, Andrew Philpott, Vincent Smart, Adolf Stroombergen, and Kiti Suomalainen. Energy- and multi-sector modelling of climate change mitigation in new zealand: Current practice and future needs. *SSRN Electronic Journal*, 01 2018.

[27] Hans-Kristian RingkjÃ¸b, Peter M. Haugan, and Ida Marie Solbrekke. A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renewable and Sustainable Energy Reviews*, 96:440 – 459, 2018.

[28] N. Abbasabadi and J.K. Mehdi Ashayeri. Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment*, 161, 2019.

[29] Lukas G. Swan and V. Ismet Ugursal. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8):1819 – 1835, 2009.

[30] Matteo Giacomo Prina, Giampaolo Manzolini, David Moser, Benedetto Nastasi, and Wolfram Sparber. Classification and challenges of bottom-up energy system models - a review. *Renewable and Sustainable Energy Reviews*, 129:109917, 2020.

[31] Zahra Beheshti and Siti Mariyam Hj Shamsuddin. A review of population-based metaheuristic algorithms. *Int. J. Adv. Soft Comput. Appl*, 5(1):1–35, 2013.

[32] Ramachandran Kannan, Hal Turton, and Evangelos Panos. *Methodological Significance of Temporal Granularity in Energy-Economic Models—Insights from the MARKAL/TIMES Framework*, pages 185–200. Springer International Publishing, Cham, 2015.

[33] Luis Mundaca and Lena Neij. Energy-economy models and energy efficiency policy evaluation for the household sector. an analysis of modelling tools and analytical approaches. Technical report, Oct 2009.

[34] Toshihiko Nakata. Energy-economic models and the environment. *Progress in Energy and Combustion Science*, 30(4):417–475, 2004.

[35] J. S. Gifford and R. C. Grace. Crest cost of renewable energy spreadsheet tool: A model for developing cost-based incentives in the united states; user manual version 4, august 2009 - march 2011 (updated july 2013). 7 2013.

[36] John Bridger Robinson. Energy backcasting a proposed method of policy analysis. *Energy policy*, 10(4):337–344, 1982.

[37] Ellinoora Leino-Richert. Backcasting energy efficiency futures of the european union - case studies of finland and germany. Master's thesis, 2017.

[38] Xinyi Li, Runming Yao, Wei Yu, Xiangzhong Meng, Meng Liu, Alan Short, and Baizhan Li. Low carbon heating and cooling of residential buildings in cities in the hot summer and cold winter zone - a bottom-up engineering stock modeling approach. *Journal of Cleaner Production*, 220:271 – 288, 2019.

[39] Steven Firth, Kevin Lomas, and Andrew Wright. Targeting household energy efficiency measures using sensitivity analysis. *Building Research and Information*, 38, 02 2010.

[40] Michael A. McNeil, Stephane de la Rue du Can, Dan Hamza-Goodacre, and Prodipto Roy. Energy efficiency indicators and impact metrics. Technical report, 02/2016 2016.

[41] Sergey Paltsev. Energy scenarios: the value and limits of scenario analysis: The value and limits of energy scenario analysis. *Wiley Interdisciplinary Reviews: Energy and Environment*, 6:e242, 07 2017.

[42] S Zahran. The software process-what it is, and how to improve it. *WIT Transactions on Information and Communication Technologies*, 8, 1970.

[43] Rob Edsall and Kelli Larson. Decision making in a virtual environment: effectiveness of a semi-immersive "decision theater" in understanding and assessing human-environment interactions. 01 2006.

[44] Tasseda Boukherroub, Sophie D'amours, and Mikael Rönnqvist. Sustainable forest management using decision theaters: Rethinking participatory planning. *Journal of Cleaner Production*, 179:567 – 580, 2018.

[45] José Molina Espinosa, Sergio Ruiz-Loza, Julieta Noguez, Rafael Lozano, and Manuel Álvarez Madrigal. *Collaborative Decision-Making Centres for Sustainable Energy*, pages 271–297. 01 2022.

[46] Marlene Ofelia Sanchez Escobar, Rafael Lozano Espinosa, Jose Martin Molina Espinosa, Juana Julieta Noguez Monroy, and Genoveva Vargas Solar. Applying process mining to

support management of predictive analytics/data mining projects in a decision making center. In *2019 6th International Conference on Systems and Informatics (ICSAI)*, pages 1527–1533, 2019.

[47] Sergio Ruiz-Loza, José Molina Espinosa, and David Balderas. Decision-making laboratory for socio-technological systems. *International Journal for Interactive Design and Manufacturing (IJIDeM)*, 14, 12 2020.

[48] David H. Jonassen. Designing for decision making. *Educational Technology Research and Development*, 60(2):341–359, 2012.

[49] Alejandro Flores Benitez, José Martín Molina Espinosa, and Manuel Alvarez Madrigal. Una propuesta metodológica para el proceso de toma de decisiones aplicada en centros de soporte a la toma de decisiones. In María Lucia Barron Estrada, José Antonio Camarena Ibarrola, Karina Mariela Figueroa Mora, and Valeria Soto Mendoza, editors, *Aplicaciones de la Computación*, chapter 11, pages 93–101. Universidad Autónoma de Coahuila, Boulevard Venustiano Carranza sin numero Colonia República Oriente, C.P.25280, Saltillo, Coahuila, Mexico, 2020.

[50] Pete Chapman, Julian Clinton, Randy Kerber, Tom Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth. *CRISP-DM 1.0 step-by-step data mining guide.* 01 1999.

[51] Robert Nisbet, John Elder, and Gary Miner. *History of Phases of Data Analysis, Basic Theory, and the Data Mining Process*, page 1. 12 2009.

[52] Marlene Ofelia Sanchez-Escobar, Julieta Noguez, Jose Martin Molina-Espinosa, Rafael Lozano-Espinosa, and Genoveva Vargas-Solar. The contribution of bottom-up energy models to support policy design of electricity end-use efficiency for residential buildings and the residential sector: A systematic review. *Energies*, 14(20), 2021.

[53] O. Jridi, S. Aguir Bargaoui, and F. Z. Nouri. Household preferences for energy saving measures: Approach of discrete choice models. *Energy and Buildings*, 103:38–47, 2015.

[54] M. Aghamohamadi and N. Amjady. Modelling and optimisation for costly efficiency improvements on residential appliances considering consumers income level. *IET Generation, Transmission and Distribution*, 11(16):3992–4001, 2017.

[55] K. Hara, M. Uwasu, Y. Kishita, and H. Takeda. Determinant factors of residential consumption and perception of energy conservation: Time-series analysis by large-scale questionnaire in suita, japan. *Energy Policy*, 87:240–249, 2015.

[56] Berndt Lundgren and Marten Schultzberg. Application of the economic theory of self-control to model energy conservation behavioral change in households. *Energy*, 183:536 – 546, 2019.

[57] Onicha Meangbua, Shobhakar Dhakal, and John K.M. Kuwornu. Factors influencing energy requirements and co2 emissions of households in thailand: A panel data analysis. *Energy Policy*, 129:521 – 531, 2019.

[58] Lei Wen and Yang Cao. A hybrid intelligent predicting model for exploring household co2 emissions mitigation strategies derived from butterfly optimization algorithm. *Science of The Total Environment*, 727:138572, 2020.

[59] Saeidreza Radpour, Md Alam Hossain Mondal, and Amit Kumar. Market penetration modeling of high energy efficiency appliances in the residential sector. *Energy*, 134:951 – 961, 2017.

[60] Mahbod Heidari, DaÅ¡a Majcen, Nicoline van der Lans, Idelette Floret, and Martin K. Patel. Analysis of the energy efficiency potential of household lighting in switzerland using a stock model. *Energy and Buildings*, 158:536 – 548, 2018.

[61] Lei Wen and Yang Cao. Influencing factors analysis and forecasting of residential energy-related co2 emissions utilizing optimized support vector machine. *Journal of Cleaner Production*, 250:119492, 2020.

[62] A. Mohseni, S. S. Mortazavi, A. Ghasemi, A. Nahavandi, and M. Talaei abdi. The application of household appliances' flexibility by set of sequential uninterruptible energy phases model in the day-ahead planning of a residential microgrid. *Energy*, 139:315–328, 2017.

[63] Marzieh Jafary and Les Shephard. Modeling the determinants of residential appliance electricity use single-family homes, homes with electric vehicles and apartments. 04 2018.

[64] H. Liang, J. Ma, R. Sun, and Y. Du. A data-driven approach for targeting residential customers for energy efficiency programs. *IEEE Transactions on Smart Grid*, 11(2):1229–1238, 2020.

[65] W. Kleebrang, P. Bunditsakulchai, and W. Wangjiraniran. Household electricity demand forecast and energy savings potential for vientiane, lao pdr. *International Journal of Sustainable Energy*, 36(4):344–367, 2017.

[66] Bijay B. Pradhan, Bundit Limmeechokchai, and Ram M. Shrestha. Implications of biogas and electric cooking technologies in residential sector in nepal – a long term perspective using aim/enduse model. *Renewable Energy*, 143:377 – 389, 2019.

[67] Erica Marshall, Julia K. Steinberger, Valerie Dupont, and Timothy J. Foxon. Combining energy efficiency measure approaches and occupancy patterns in building modelling in the uk residential context. *Energy and Buildings*, 111:98–108, 2016.

[68] Marta Braulio-Gonzalo, Pablo Juan, María D. Bovea, and María José Ruá. Modelling energy efficiency performance of residential building stocks based on bayesian statistical inference. *Environmental Modelling and Software*, 83:198–211, 2016.

[69] Xia Wang, Wei Feng, Weiguang Cai, Hong Ren, Chao Ding, and Nan Zhou. Do residential building energy efficiency standards reduce energy consumption in china? a data-driven method to validate the actual performance of building energy efficiency standards. *Energy Policy*, 131:82 – 98, 2019.

[70] C. Davila, Nathaniel Jones, Adil Al-Mumin, A. Hajiah, and C. Reinhart. Implementation of a calibrated urban building energy model (ubem) for the evaluation of energy efficiency scenarios in a kuwaiti residential neighborhood. 2017.

[71] Moncef Krarti, Mohammad Aldubyan, and Eric Williams. Residential building stock model for evaluating energy retrofit programs in saudi arabia. *Energy*, 195:116980, 2020.

[72] T. Schütz, L. Schiffer, H. Harb, M. Fuchs, and D. Müller. Optimal design of energy conversion units and envelopes for residential building retrofits using a comprehensive milp model. *Applied Energy*, 185:1–15, 2017.

[73] Narjes Abbasabadi, Mehdi Ashayeri, Rahman Azari, Brent Stephens, and Mohammad Heidarinejad. An integrated data-driven framework for urban energy use modeling (ueum). *Applied energy*, 253:113550, 2019.

[74] Tobias Fleiter, Matthias Rehfeldt, Andrea Herbst, Rainer Elsland, Anna-Lena Klingler, Pia Manz, and Stefan Eidelloth. A methodology for bottom-up modelling of energy transitions in the industry sector: The forecast model. *Energy Strategy Reviews*, 22:237–254, 2018.

[75] Karl-Kiên Cao, Felix Cebulla, Jonatan Gomez Vilchez, Babak Mousavi, and Sigrid Prehofer. Raising awareness in model-based energy scenario studies-a transparency checklist. *Energy, Sustainability and Society*, 6:1–20, 09 2016.

[76] Mohammed Mahmoud, Yuqiong Liu, Holly Hartmann, Steven Stewart, Thorsten Wagener, Darius Semmens, Robert Stewart, Hoshin Gupta, Damian Dominguez, Francina Dominguez, David Hulse, Rebecca Letcher, Brenda Rashleigh, Court Smith, Roger Street, Jenifer Ticehurst, Mark Twery, Hedwig van Delden, Ruth Waldick, Denis White, and Larry Winter. A formal framework for scenario development in support of environmental decision-making. *Environmental Modelling Software*, 24(7):798–808, 2009.

[77] Tecnológico de Monterrey. Reporte del modelo de consumo eléctrico (proyecto 266632 “laboratorio binacional para la gestión inteligente de la sustentabilidad energética y la formación tecnológica”). unpublished confidential document. Technical report, 03/2018 2018.

[78] NU. CEPAL and German Agency for International Cooperation. Informe nacional de monitoreo de la eficiencia energética de méxico, 2018. Technical report, 05/2018 2018.

[79] Cliff T. Ragsdale. *Spreadsheet Modeling and Decision Analysis.* Cengage Learning, 6th edition, 2011.

[80] Panagiotis Drakoulogkonas and Dimitrios Apostolou. *A Comparative Analysis Methodology for Process Mining Software Tools*, pages 751–762. 08 2019.