TÉCNICO LISBOA

Blockchain-based automatic energy efficiency model for Performance Contract application

Pedro Manuel Ferreira Sismeiro

Thesis to obtain the Master of Science Degree in

Engineering Physics

Supervisor(s): Prof. Doutor Carlos Augusto Santos Silva
Prof. Doutor Ilidio Pereira Lopes

Examination Committee

Chairperson: Prof. Doutor Horácio João Matos Fernandes
Supervisor: Prof. Doutor Carlos Augusto Santos Silva
Members of the Committee: Prof. Doutor Luís Filipe Moreira Mendes
Doutora Diana Pereira Neves

December 2020

Aos meus pais, que sempre souberam quem sou.

À família, pela confiança incondicional.

Aos amigos de antes e aos de agora, que não deixaram que as luzes se apagassem.

À Angie, os dias todos.

"Straight up!"

Travis Scott

Acknowledgments

I'd like to express my deepest thanks to my supervisors and the IN+ staff, for the trust and autonomy they've laid on my hands with such natural easeness. Without their supportive help, everything below this section would be of significantly less value and quality.

No words can be enough to thank all my university and childhood friends, for the unconditional support throughout this degree. You have made it all way easier to go thorugh these five years by just being there and sharing experiences with me, academically and leisurely. Thank you for helping me build a better version of myself every single day, one that surely will take a part of all of you with me. Thank you for knowing me so well. A special thanks to Maria Aveiro, my tech consultant, once again showing that high school years are forever.

To the people I've met on exchange, both in India and Colombia, thank you. You've shaped me into a globally aware individual and caressed me with your tender ways when the light seemed to fade.

To my close family, you've been my main support stone from day one. Thank you for once again showing your love in a healthy and supportive way. I hope to pay you back someday.

To Angie, my latin american gust wind, thank you for knowing exactly who I am, every day. You've made all of this a journey, when everything felt hazy. You helped me go through life with your tender and soft delicacy. With you, I won every day.

Resumo

A transformação e uso de energia ainda é a maior fonte de emissões de gases com efeito estufa. Do lado da procura, há um esforço global para investir em melhorias de eficiência energética, que tem abrandado. Melhorias de eficiência trazem benefícios para os países como aumento de segurança energética, menor gasto em combustíveis fosseis e redução de emissões. As empresas de serviços energéticos (ESCOs) implementam medidas de eficiência energética (ECMs) através de contratos de performance energética (EPCs), que garantem um certo nível de poupança energética/económica. A medição e verificação (M&V) é central para auditar os termos contratuais e eficiência das ECMs, em EPCs. Estruturas de M&V deficitárias podem gerar desconfiança entre partes e cálculos de poupança ambíguos. Esta tese visa aumentar a transparência na implementação de EPCs calculando, guardando e protegendo informações de poupança devidamente. Desenvolveu-se um modelo de consumo base para dois edifícios do IST onde foram instaladas ECMs, usando o algoritmo XGBoost. Estimam-se as poupanças pela diferença com o consumo medido, no período homólogo. Os modelos usados apresentaram CV(RMSE) inferior a 7.8% e retornaram poupanças de 16.9$\pm$7.3% e 20.6$\pm$6.3%. Os cálculos são publicados numa *blockchain*, composta por nós-edifício. O critério de validação de publicações verifica se o modelo usado para determinar poupanças é suficientemente preciso. Uma plataforma mais clara e fiável para a execução de EPCs foi desenvolvida. Conclusões, limitações e futuras melhorias são discutidas.

Palavras-chave: Eficiência energética, Contratos de performance energética, XGBoost, Blockchain, Serviços de energia

Abstract

Energy transformation and usage is still the major source of greenhouse gas emissions. On demand side, there is a global push to invest on energy efficiency improvements, which has been slowing down. Efficiency improvements have benefits for countries such as increased energy security, less spending on fossil fuels and emissions reduction. Energy Service Companies (ESCOs) deploy energy conservation measures (ECMs) through Energy Performance Contracts (EPCs), which guarantee a level energy/cost savings. Measurement and Verification (M&V) procedures are essential to EPCs, as to audit contract terms and ECM efficiency. Poor M&V frameworks can generate adversarial distrust between parties involved and unclear savings calculations. This integration thesis aims to increase transparency in EPC-mediated ECM implementations by properly assessing, storing and securing savings calculations. We develop a baseline model using XGBoost for two IST campus buildings which underwent retrofits and estimate savings from the difference to actual consumption data, for the same period. The used models presented a CV(RMSE) of under 7.8% and yielded savings percentages of 16.9$\pm$7.3% and 20.6$\pm$6.3%. Savings information are then posted in a blockchain ledger composed of building nodes. The transaction validation mecanism verifies if an accurate baseline model was used as basis for the calculations. A more clear and thrustworthy M&V platform for EPC execution was developed. Conclusions, limitations and future improvements are discussed.

Keywords: Energy efficiency, Energy Performance Contracting, XGBoost, Blockchain, Energy services.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The world is currently out of track to meet the 17 Sustainable Development Goals (SDGs), proposed by the United Nations [1] and, in particular, the 3 goals that are directly related to the energy sector, which consist on: tackling climate change, assuring universal access to energy and reducing health impacts of air-pollution [2].

At the same time, energy transformation and usage still remains the major source of greenhouse gas emissions, which need to quickly reduce and establish a net-zero balance between anthropogenic emissions by sources and removals by sinks [2, 3].

But not all hope is lost: some scenarios still propose optimistic achievements in the energy sector on the mid-run, such as the Sustainable Development Scenario (SDS), proposed by the International Energy Agency (IEA). If properly followed, their drawn recomendations guarantee that the global temperature rise can stay below 1.8° C with a 66$\%$ probability, without relying on global net-negative CO_2 emissions [4]. It also states that if global emissions fall from 33 billion tonnes in 2018 to less than 10 billion tonnes by 2050, we will be in route to net-zero emissions by 2070. But still, acting now to harness the power of net-negative emission technologies could still help achieve: net-zero emissions globally, universal access to modern energy and less 3.1 million premature deaths by 2050, related to air pollution by around 2050 [4].

In terms of economic push, capital is already moving from fossil to renewable to a significant extent. These improvements shall represent an increase in overall investment, which will be counterbalanced by reduced fuel costs on the consumer end, afterwards. On the supply side, the largest increase in investment comes from renewable-based power installments, which are expected to double until 2050, as well as additional spending on electricity grids and storage. On the demand side, there's a global push for further investment, in particular, on energy efficiency improvements (on buildings, industrial processes, transport and infrastructures). Still, this latter push is currently not in pace with supply side developments: we are focusing more on supply than on the demand side [4].

In fact, these efficiency improvements have been slowing down in recent years. In 2018, worldwide primary energy intensity improved by just 1.2%, the slowest rate since 2010, while optimistic scenarios like the SDS suggested a yearly improvement of 3%. A 1.2% improvement represents around 1,6 trillion

USD more in worldwide GDP and had it improved at the suggested rate, this value would have circled the 4 trillion USD milestone - around the size of the German economy. Major factors responsible for this slow pace include increased energy usage in industry, changing climate patterns, which are responsible for increased energy usage and flat policy progress/investment. Although they are lacking attention, efficiency improvements have significant benefits for countries, such as increased energy security - in which oil importers reduce their exposure to oil market instability -, finantial benefits from importing less oil - in 2018 both Japan and China spent around less 20 billion USD on oil due to technical efficiency improvements - and cutting CO_2 emissions. It's known that from 2015 to 2018, energy-related emissions dropped by 3.5Gt - the equivalent of the energy-related CO_2 emissions of Japan [5].

In conclusion, this transition to a low carbon economy requires a more efficient and synergistic energy system, that relies less on fuel combustion and more on renewable based power, but we still have to focus about energy efficiency improvements on the demand side, while not trading off energy security nor affordability. As such, boosting energy efficiency improvements in a reliable way, on the demand side shall be the main focus of this dissertation.

1.1 Motivation

The proposed solution takes an holistic approach towards improving energy efficiency and reducing carbon emissions by starting in a reverse way: dealing with it from the client demand side upwards. It helps securing better performance when trying to audit energy efficiency improvements, in a certain environment. Instead of relying on immediate carbon capturing capacity, we need to start optimizing energy systems from the bottom up, since that will also allow us to have greater ambitions on cutting greenhouse gas emissions and further improving supply.

Thus, I shall provide a tool for assessing efficiency improvements while helping to boost business models based on energy trading, energy certificates and/or energy performance contracting, well known to Energy Service Companies (ESCOs) [6–9]. It will guarantee safe energy transactions and a trusted auditing method that follows agreed terms by all parts involved.

In an ever more stressed energy environment, with rising electricity costs, additional environmental regulations and less marginal profit, buildings owners benefit from reducing their energy consumption and costs. Energy Services ESCOs provide clients with Energy Performance Contracts (EPCs), which are binding agreements under which energy savings are provided, verified and monitored during a certain period of time [10]. The ESCO designs and implements Energy Conservation Measures (ECMs), like changes in the lighting system or insulation envelope, and guarantees a given level of energy savings, over the period of the contract.

The level of energy savings secures the financial revenue that is used to fund the cost of improvements and services incurred at the ESCO side and financial savings are in general shared between the two parties from the start. Once the costs have been repaid, the client keeps the full savings generated from the ECMs [10, 11]. In case of failure in provisioning the contractually-agreed energy savings, financial penalties are applied to the service provider which reduces the contract revenue. A measurement and verification (M&V) procedure is then essential to EPCs, as to audit the contract terms and the proposed measures' efficiency. The energy saving is generally computed as the difference between a predictive baseline model and the real measured post-EPC, energy consumption, over the considered period [6, 10–12].

A poor framework for M&V can generate problems such as an unbalanced performance risk and unclear or inappropriate savings calculations [12, 13]. The International Performance Measurement and Verification Protocol (IPMVP) has been developed to provide guidance and standards for M&V procedures [14]. It tells us how external variables should impact the baseline model and how adjustments should be undertaken. This protocol is frequently used as guideline for M&V procedures on EPCs, but the lack of understanding impacts the client when it comes to grasping the way that the model works, leading them to opt on simple predictive models rather than more complex ones, which use a wider set of parameters (like meteorological ones), increasing its quality and lowering the commitment risk, for both parts [6, 15].

When considering the amount of data, there's also the risk of data tampering from both parts or from external providers, leading to inaccurate savings calculations, especially since depending on the modality of the EPC, the ESCO might be entitled to any excess savings. An adversarial relationship can be generated between parts, for example when the ESCO tries to reduce the expected savings level in order to receive more from that excess or when the client changes the building's consumption load without notifying the ESCO [10].

In essence, there is a problem of *trust* and understandibility when dealing with EPCs, which is limiting the widespread of this useful tool. Authors also claim lack of standardization or policy concerning EPC execution [7, 8, 16–18].

For its incorruptible and inmmutable character together with the lack of need for verification of a trusted third party, Blockchain's ability to track down transactions is getting increased attention on energy sector applications [9]. As such, this present work tries to evaluate its application on EPCs to overcome some of the limitations faced and to increase client trust in M&V procedures. By having it designed in a way that all model predictions and calculations happen inside it, I will develop a more trustworthy and standardized framework for executing and auditing EPCs.

1.2 Concept

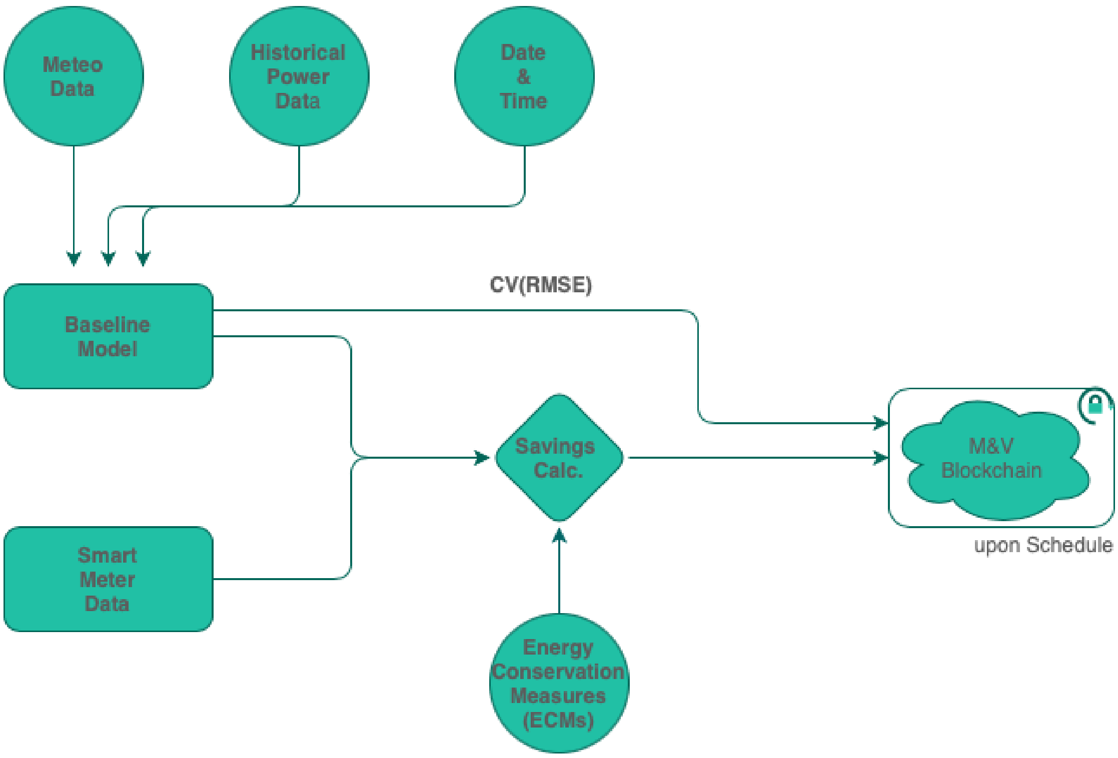


Figure 1.1: Proposed solution conceptual diagram.

The scope of this thesis is to bring transparency and security to EPC-mediated ECM implementations, by properly assessing, storing and securing savings calculations, following M&V 2.0 tendencies [43]. We shall develop an energy consumption forecasting model based on relevant data from selected features, called a baseline model. Then, actual data will be compared to the model predictions and the savings are calculated from the difference between them, under a given uncertainty. As to guarantee transparency, these results are stored in a prototype blockchain composed of building nodes posting under a closed network, which verifies that an accurate baseline model was used as basis for the calculations (Fig. 1.1).

It is expected to be a relatively new contribution towards the main cited challenges faced by ESCOs in EPCs and M&V when tackling energy efficiency, global primary energy intensity improvement and, ultimately, CO_2 emissions reduction. Khatoon et al. (2019) [19] and Gurcan et al. (2018) [20] is one of the few literature references that developed a solution like the one proposed by this dissertation, which usually rely on common blockchain platforms like Ethereum/Hyperledger or industrial-sector studies when facing this topic. The key values of this proposed prototype are simplicity, adaptability and effectiveness in the scope of the purpose of use, making it easier to adapt further details in this platform when faced with different real world applications and security requirements. As such, the main development objectives are laid in the section below.

1.3 Methodical Approach

Whether a client is feeding the grid by producing renewable energy, improving its own energy system efficiency or just trading green certificates based on saved emissions compensations, there is a need for predicting the baseline consumption, measuring the actual data and compare the difference, which can be translated in several relevant parameters and, ultimately, savings. Then, savings informations are stored in a secure, credible (and rather hyped) way: a Blockchain ledger. The milestones for this thesis are:

1. Properly understanding each one of the relevant topics to this dissertation.
2. Studying the pattern of energy consumption and ECMs of IST's *Alameda* Campus major buildings and draw relevant feature recommendations for building a regression model.
3. Building an accurate, effective and tested energy consumption forecasting model to serve as a baseline estimate.
4. Comparing the actual data with the prediction model and estimating savings for the reporting period.
5. Developing a secure, simple and adaptable blockchain algorithm to store the M&V information regarding savings.
6. Integrating all parts into a proper blockchain-based automatic performance contract auditor, testing and adjusting to the necessary extent.
7. Describing the prototype's limitations, possible future developments and results.

This thesis is expected to positively contribute to improve the state-of-the-art of M&V 2.0 and EPC applications, thereby improving primary energy intensity worldwide, cutting emissions and increasing the energy system's efficiency levels

1.4 Thesis Outline

Seven different chapters will be laid out in order to build the premise of this dissertation. The theoretical background is thoroughly exposed on chapter 3, concerning all relevant topics to the proposed solution. Chapters 3 and 4 focus on the adopted energy modeling and blockchain solutions, respectively. Chapter 5 describes the integration of all parts into a proper automatic EPC M&V blockchain. Finally, the last chapter lays out the retrieved results, limitations and conclusions of this endeavor.

1. **Introduction**: the beginning of this dissertation which includes the motivation, methodical approach and concept, as well as the thesis outline
2. **Background and Literature Review**: a section dedicated to giving background literature context about the main topics of concern. It provides information about the problems faced and the research hypothesis.

3. **Energy Modeling**: this chapter focuses on how to develop an energy consumption forecasting model from historical data, relevant features and energy consumption data analysis.

4. **Blockchain**, explaining the functioning of the proposed blockchain solution adapted to this application and relevant adjustments undertaken.

5. **Implementation and Usage**: a description of the proposed integrated prototype which encapsulates the energy forecasting model and blockchain algorithm frameworks. It shall technically explain the solution's behavior, feasibility and easiness to adapt to a changing environment.

6. **Conclusions**: summing up the achievements and limitations of this prototype and draw future work recommendations.

Chapter 2

Background and Literature Review

2.1 Energy Consumption & Efficiency

2.1.1 Global energy consumption

Global electricity production has grown each year continuously since 1974, except for between 2008 and 2009, when the global financial crisis caused a decline in production. In 2018, world gross electricity production was 3.9% higher than 2017 and consumption reached a new record of 23738.9 TWh [21] [22]. The production is dominated at 58% by non-OECD countries, which grew 4.8% a year on average from 2010-2018, compared to just 0.3% on the OECD. As of 2018, in terms of sources, OECD countries chose natural gas as the main gross electricity production fuel (Fig. 2.1) [23] but, globally, coal continues to be the main source (Fig. 2.2) [24].

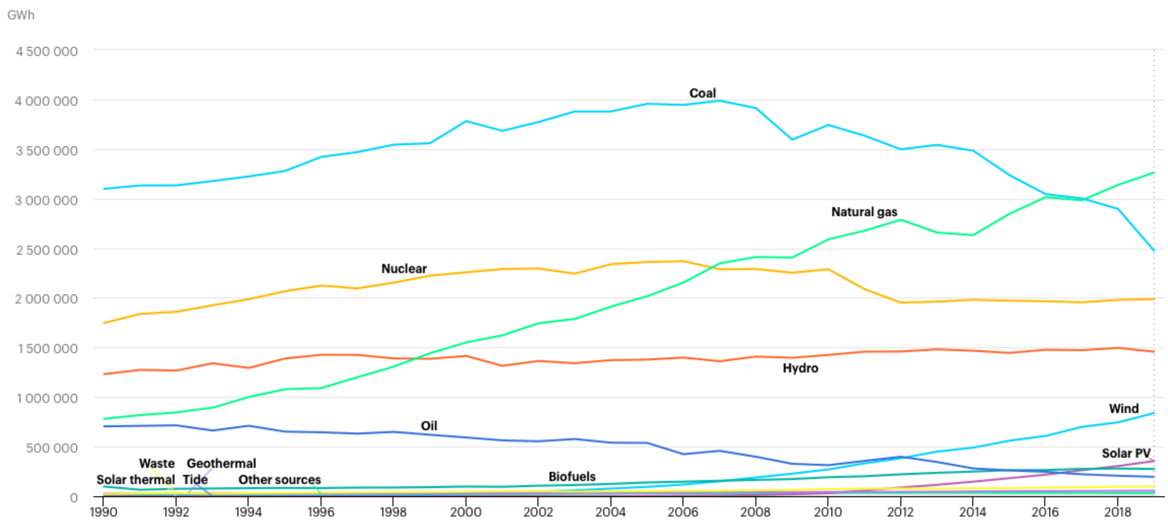


Figure 2.1: Electricity generation by source, OECD total 1990-2019 [23].

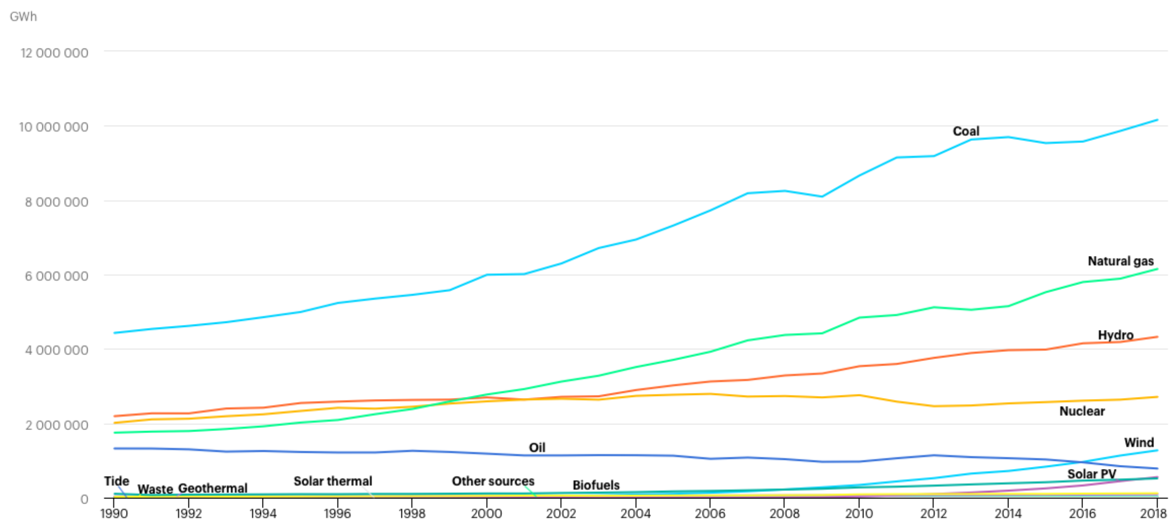


Figure 2.2: Electricity generation by source, world total 1990-2018 [24].

It is also clear from the graphs that currently, the dominant low-carbon emission sources are hydro and nuclear power, followed by wind and photovoltaic power.

On the demand side, the world total electricity final consumption in 2018 was 4% higher than in 2017, surpassing the increase in production, with a consistent higher increase in non-OECD countries of around 5.7% compared to just 1.8% increase in OECD countries [22]. In terms of electricity, most energy has been directed mainly to industry and residential, commercial or public buildings (Fig. 2.3), sectors where much of the growth in OECD electricity consumption has taken place since 1974 [25]. Industry has always been a key driver for electricity consumption (Fig. 2.4), but as of 2018, it is only marginally superior to that of residential/commercial/public buildings and even inferior in the US [26]. Road transport has recently experienced strong growth in electricity consumption as electric vehicles spread across OECD countries, in particular in Europe [22].

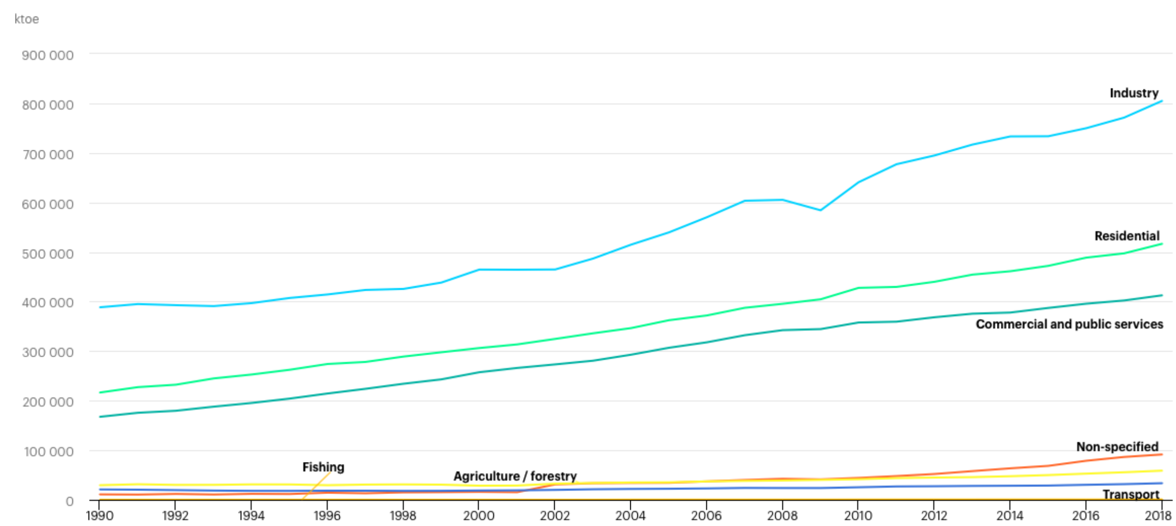


Figure 2.3: Electricity final consumption by sector, World 1990-2018 [25]. 1 ktoe (kilo-tonne of oil equivalent) is equal to 11.63 gigawatt-hours (GWh) [27].

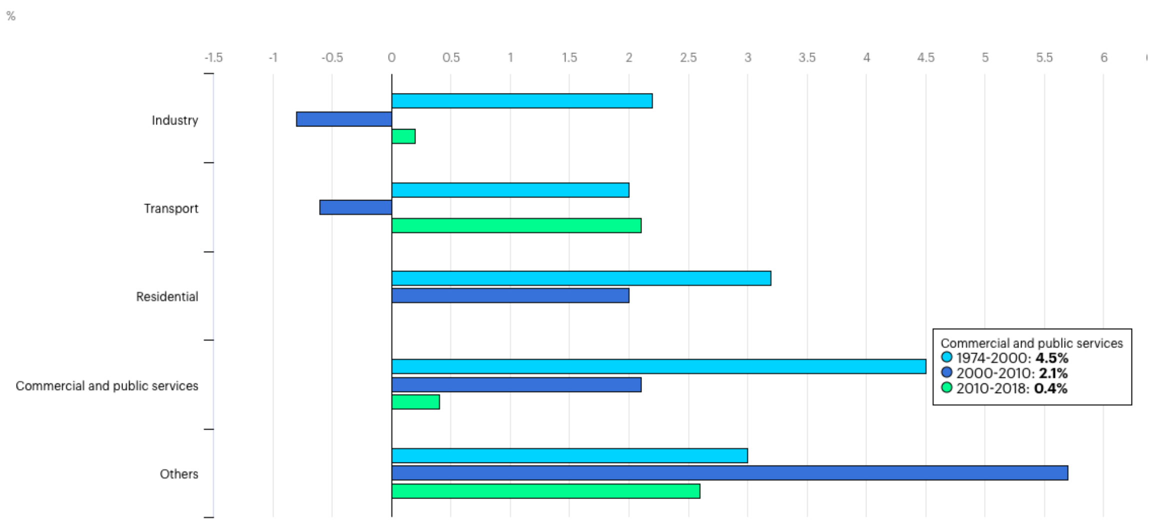


Figure 2.4: OECD average annual growth rate in electricity final consumption by sector, 1974-2018 [22].

The largest non-OECD consumers of electricity in 2018 were the People's Republic of China, India, the Russian Federation and Brazil, which represent 38.0% of global consumption (Fig. 2.5). Among these countries, China has the largest share, at 47.8% of total non-OECD consumption. Electricity use outside the OECD is dominated by industrial demand, which accounts for half of final electricity consumption.

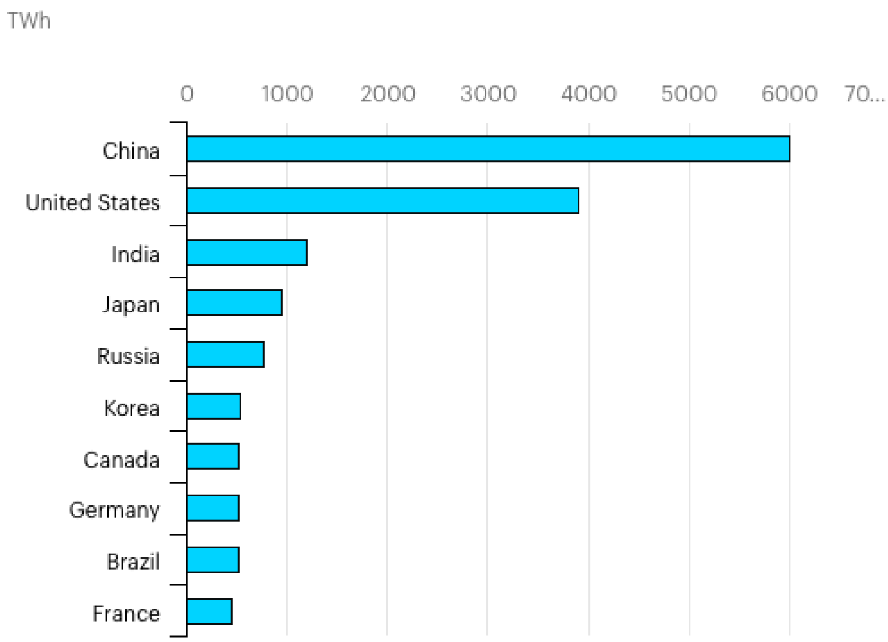


Figure 2.5: Top ten electricity consuming countries, 2018 [22].

In the EU, the main end use of energy by households is for heating (63.6 % of final energy consumption in the residential sector, Fig. 2.6). Electricity used for lighting and most electrical appliances represents 14.1% (this excludes the use of electricity for powering the main heating, cooling or cooking systems), while the proportion used for water heating is slightly higher, representing 14.8% [28]. Heating of space and water consequently represents 78.4 % of the final energy consumed by households (Fig. X). Throughout the 27 countries, the lowest proportions of energy used for space heating in 2018 were observed in Malta (20.4 %) and Portugal (28.2 %) and the highest in Luxembourg (78.7 %) and Belgium (73.5 %), which makes it a rather heterogenous matter [28].

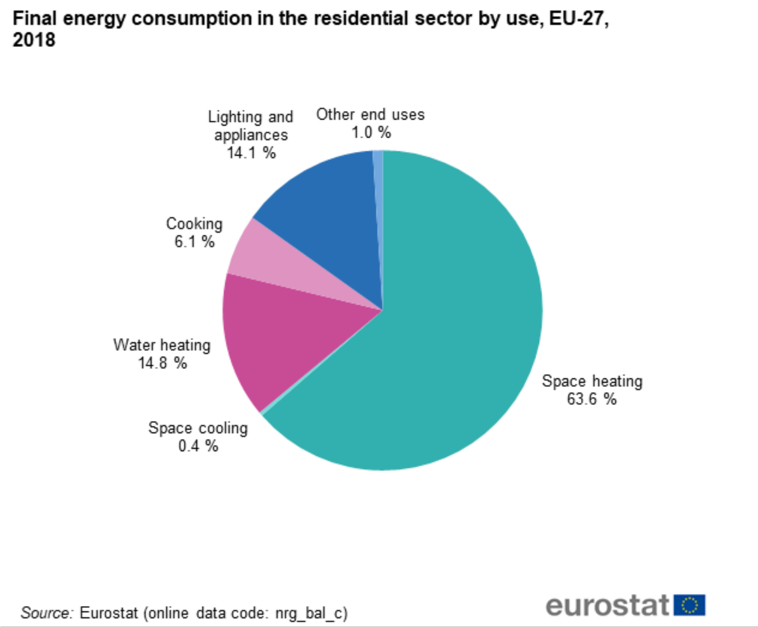


Figure 2.6: Final energy consumption in the residential sector by use, EU-27, 2018 [28].

Concerning energy source, electricity covers 100% of the energy needs for lighting and space cooling in the EU but also 83.4% of the other end-uses and 49.2% for cooking purposes. Renewables cover 27.0% of the energy needs for space heating, 12.6% for water heating and 5.6% for cooking purposes. Derived heat plays an important role in water heating (13.9%) and in space heating (10.6%), while oil products still cover 14.1 % of space heating energy use, 13.5% of cooking and 11.3% of water heating [28]. Fourteen out of 27 EU Member States use mainly renewable energies for space heating, with Portugal (80.9%), Croatia (65.0%), Bulgaria (59.3%) having the largest proportion. While the number of countries using mainly gas for this purpose is smaller (ten Member States), some of them are among the largest energy consumers of the EU – the Netherlands (86.2%), Slovakia (65.4%) and Italy (58.4%), where the proportion is the highest. Three Member states use mainly petroleum products for space heating: Cyprus (63.0%), Ireland (50.1%) and Greece (43.0%). Finally, one Member State (Sweden) mostly relies on derived heat (48.7%) and one Member State (Poland) uses mainly solid fuels for space heating (44.9%) [28]. Cooking is generally based on the use of electricity (in seventeen Member States) and gas (ten Member States) [28].

2.1.2 CO_2 Emissions

Global CO_2 emissions went up 1.7% in 2018 (560 Mt, equivalent to total international aviation emissions), the highest rate since 2013, reaching an astonishing 33.1 Gt (Fig. 2.7). That year, emissions from all fossile fuels increased but the power sector accounted for around two thirds of emissions' growth [29]. This increase was driven mainly by a global robust economy and weather conditions which boosted heating and cooling. From 2014-2016, we've seen a stagnation in emissions, mostly due to strong energy efficiency improvements and low-carbon technology deployment, but in 2017-2018 that just wasn't enough to compensate global demand increase. The result was a 0.5% growth in emissions per 1% gain in global economic output, in 2018, compared to 0.3% the average since 2010. Nonetheless, renewables and nuclear power made an impact, with emissions growing 25% slower than energy demand [29].

In fact, coal use in power generation surpassed 10 Gt worldwide, mostly coming from Asia. China, India and the US account for 85% of the net increase in 2018 (with relatively young coal plants), while emissions declined in Germany, Japan, Mexico, France and the UK [29]. Assessing the fossil fuel use, the IEA [29] found that coal combustion is responsible for 0.3°C of the 1°C increase in global temperature since pre-industrial levels. Finally, the global average anual concentration of carbon dioxide in the atmosphere scored 407.4 ppm, almost double the 180-280 ppm range in pre-industrial times [29].

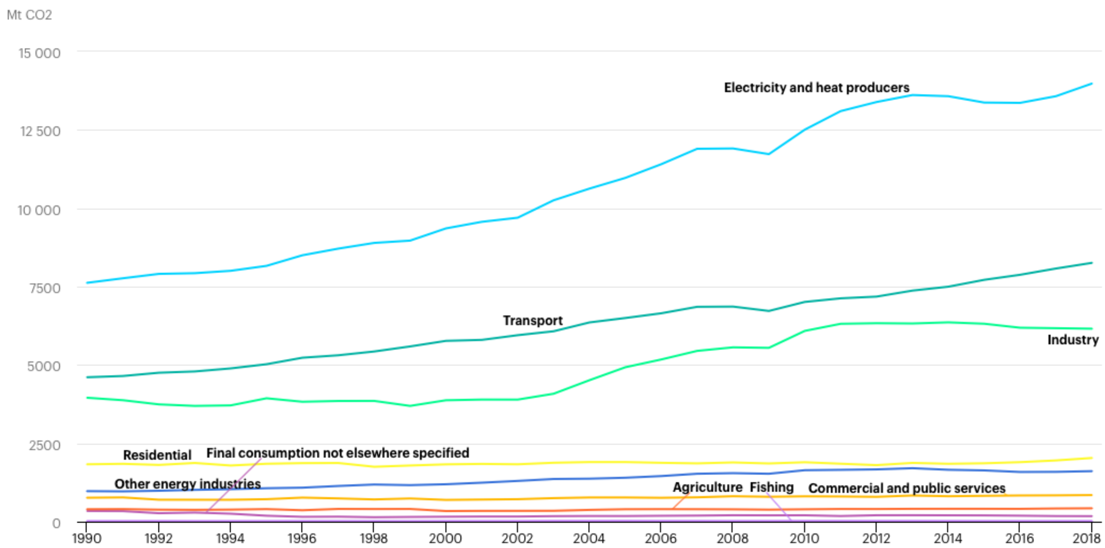


Figure 2.7: CO2 emissions by sector, World 1990-2018 [30].

As of the year of 2018, coal fired electricity generation was responsible for 30% of global emissions even though fuel switching between coal and natural gas accelerated in 2018. That change avoided 60Mt of coal demand and 95Mt of CO_2 emissions, mostly in China and the US [29]. In addition, as of 2016, energy usage in buildings accounted for 17.5% of the global greenhouse gas emissions, with commercial buildings representing 6.6% of global emissions [**?**].

But not all is bad news: carbon capture and usage facilities are expected to reduce the carbon dioxide present in the atmosphere up to 15%. Electricity generation from renewable sources also increased by over 7% in 2018, injecting an additional 450 TWh into global electricity networks. Output from nuclear contributed another 90 TWh of low-carbon generation [29].

Despite continued growth in emissions, the power sector has also seen transformation in recent years. The average carbon intensity of electricity generated is around 475 gCO_2/kWh, a 10% improvement on the intensity from 2010. Without this, global CO_2 emissions would have been 1.5 Gt (10%) higher [29]. When considering Portugal (our case study location), carbon intensity is significantly lower than the global average at around 265 gCO_2/kWh (2019-2020) [31] [32], benefiting from a dominantly renewable energy mix. This value is expected to lower after the shutdown of the two major coal plants in the country, scheduled to close until 2022 [33].

2.1.3 Energy Conservation Measures (ECMs)

Technical efficiency gains have been the key driver for improving the primary energy intensity. Global improvements between 2015 and 2018 avoided around 4% more energy consumption in 2018, nearly equivalent to the total primary energy demand of France and Italy combined. These gains more than doubled the global primary energy intensity improvement rate in 2018 (Fig. 2.8) [5].

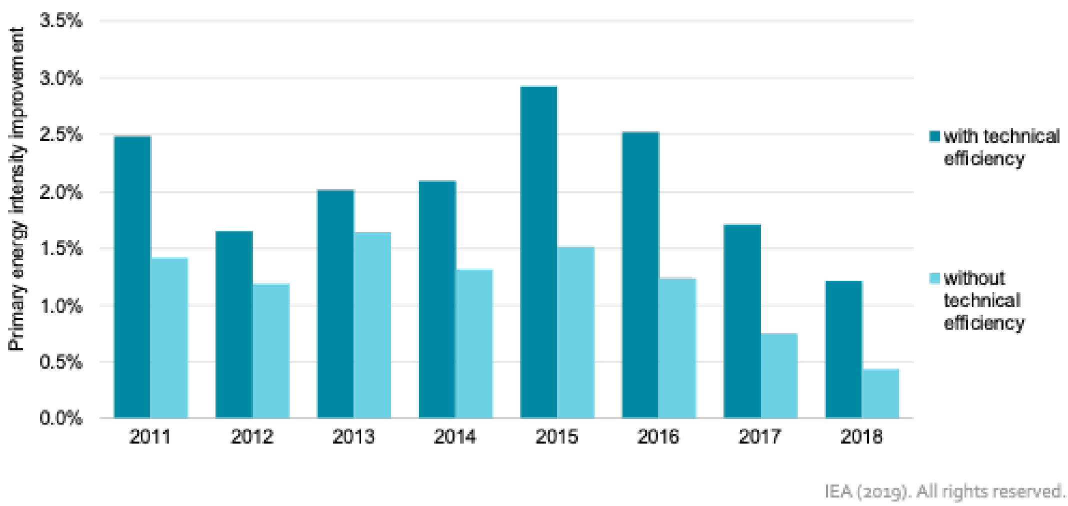


Figure 2.8: Impact of technical efficiency on primary energy intensity improvement (2011-18) [5].

Apart from reducing emissions and energy consumption, improvements in efficiency have a significant financial impact on GDP, while shielding countries from oil/fuel market instability and saving billions of USD (Figs. 2.9 and 2.10) [5]. Oil represents the largest proportion of import savings globally, saving of over 165 million tonnes of oil-equivalent from 2000-2018, similar to the energy used in the transport sectors of Japan, Canada and Italy combined.

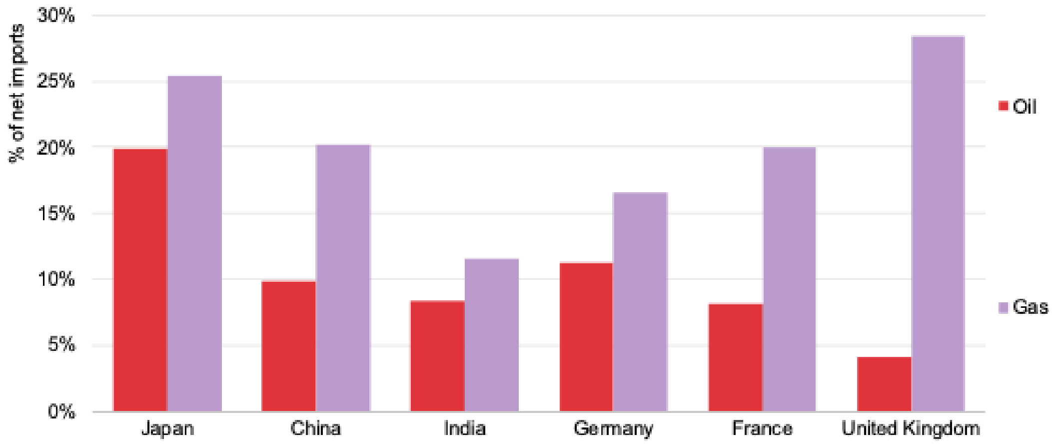


Figure 2.9: Avoided oil and gas imports in 2018 due to technical efficiency gains since 2000 [5].

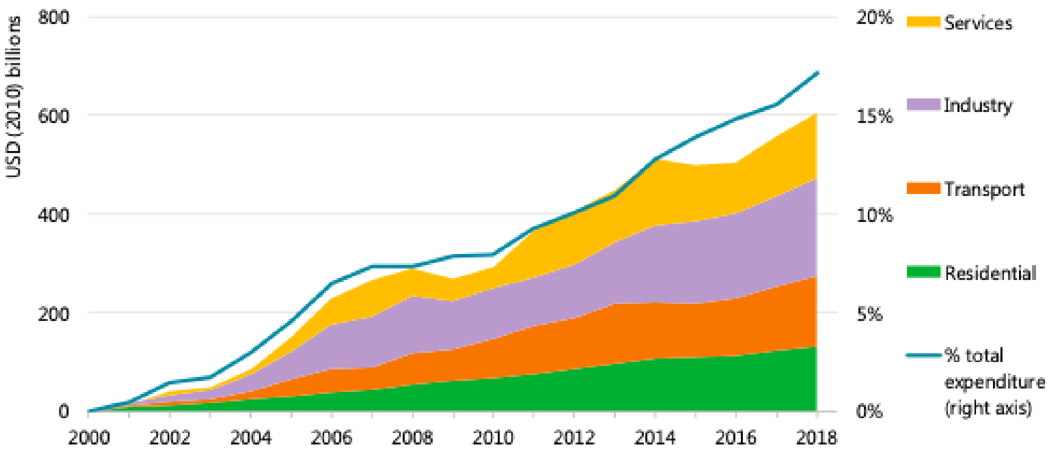
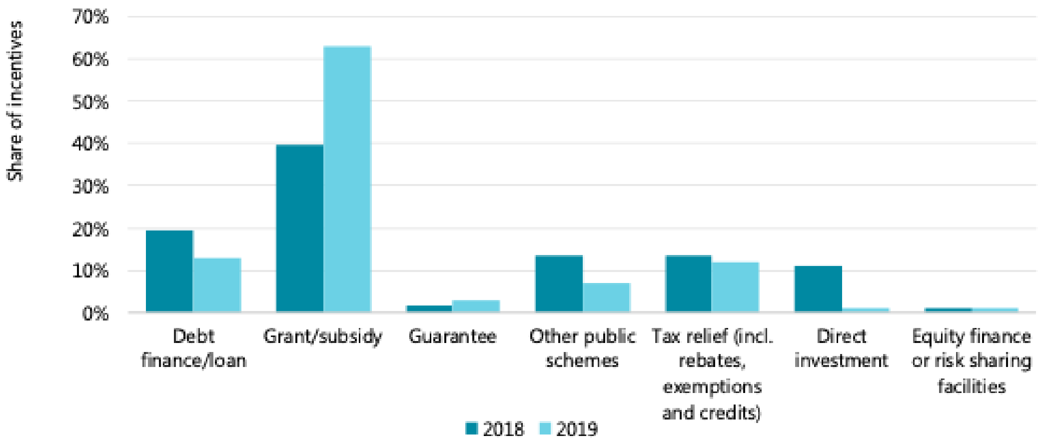


Figure 2.10: Avoided expenditure on energy due to efficiency improvements since 2000, by sector [5].

There are highly quoted studies which show us the impact this measures can have in reducing the energy consumption and carbon footprint of buildings worldwide. Kneifel et al. [34] showed that, in general, ECMs deliver consumption savings ranging from 20 to 30% without significant alterations to building design in a cost-effective way. They found that these improvements can also be responsible for up to 32% less carbon emissions over a 10-year period, when concerning energy consumption. They also exorted the comunity to adopt smarter and integrated HVAC (Heating, Ventilation & Air Conditioning) solutions in buildings. Popescu et al. [35], even shows that thermal retrofits can increase the property value of old residential buildings in between 2-3%, with up to 60% of the retrofit investment costs recovered on property transactions.

Government policies play a significant role in accelerating the development and adoption of energy-efficient strategies in all end-use sectors [5]. Several important developments in policy have been taking place in air conditioning and cooling, building certification/rating, carbon taxes and comunitary energy efficiency policy frameworks [5]. As of 2018, grants and subsidies continued to be the policy tool of choice (Fig. 2.11).



Notes: Countries included in this analysis for 2019 are Australia, Austria, Canada, Czech Republic, Greece, Ireland, Italy, Japan, New Zealand, Norway, Portugal, Slovak Republic, Spain, Switzerland, Turkey, United Kingdom and United States.

Figure 2.11: Government incentives for energy efficiency by type of incentive [5].

Confirming this data, Portugal launched a EUR 4.5 million program to improve building energy efficiency (Edificios + Sustentaveis, [36]), where consumers can get up to 7500€ per house (to a maximum of 2 houses) to undertake energy conservation measures, representing 70% of the investment cost, in 2020. The program focuses on window improvement, eco/recycled material thermal insulation , renewable heating/cooling systems, renewable deployment, water management and bioclimatic architecture.

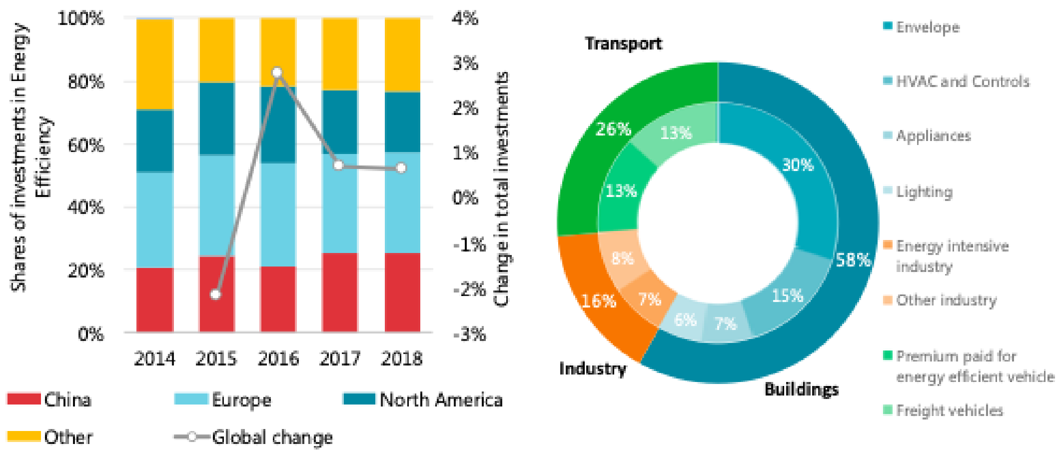


Figure 2.12: Energy efficiency investment by region, 2014-18 (left) and by sector in 2018 (right) [5].

At USD 240 billion, efficiency investments across the buildings, transport and industry sectors were 1.6% higher in 2018. Within this global picture, trends vary from sector to sector and between advanced and emerging economies (Fig. 2.12). Investments in the buildings sector, for example, dropped 2%, however, at USD 139 billion, buildings still attracted the highest share of global investments. Industrial efficiency investments increased in China by 12% and in India by 5%, but have continued to decline in the US since 2015. Transport efficiency investments have increased slightly and mainly in freight [5].

2.1.4 The impact of the digital transition in the energy transition

Nowadays, digital technologies combine at a network level to improve energy efficiency in all sectors of energy consumption. Gathering huge amounts of real time data through sensors and smart meters and then storing them on distributed ledgers, like blockchain (the main focus of this tesis), can help in simplifying the process of establishing EPCs, for example. Through interfaces, simulations, data analysis/AI algorithms and actuators, consumers can more easily meet and assess efficiency standards. That is, the more interconected a building energy system is, the easier it gets to improve efficiency, since this allows for consumers to value efficiency accurately and in a fast way. The IEA listed the possible benefits of digitalisation on the energy sector and emissions below (Figs. 2.13, 2.14, 2.15 and 2.16) [5]:

Sector	Description	Possible benefits
Buildings	Between 2017 and 2040, increased digitalisation of both commercial and residential buildings, including 1 billion connected buildings and 11 billion connected devices.	Up to 10% less energy used. Cumulative energy savings of 234 EJ.
Transport	In urban transport, between 2015 and 2050 digitally enabled innovative technologies, including teleworking, massive shared mobility and autonomous vehicles, significantly reduce passenger kilometres travelled.	More than 50% lower CO_2 emissions in 2050.
Industry	Estimated cumulative impact from combining a range of digital technologies and advanced software applications.	Up to 30% energy savings.
Flexible demand capacity	With increased policy action prioritising digital strategies and smart infrastructure, flexible demand capacity increases from 40 GW to 450 GW.	Ten times more flexible demand capacity by 2040.

Sources: IEA (2019), *Perspectives for the Clean Energy Transition: The Critical Role of Buildings*, www.iea.org/publications/reports/PerspectivesfortheCleanEnergyTransition; IEA (2018b), *World Energy Outlook 2018*; IEA (2017b), *Digitalisation and Energy*; Schneider Electric (personal communication); ITF (2019), *ITF Transport Outlook 2019*.

Figure 2.13: Possible global benefits of digital technology [5].

Technology	Description	Possible benefits
Smart thermostat	Heating and cooling controlled remotely (or automatically) and temperature adjusted according to preferences or sensor inputs.	5-20% less energy use for heating or cooling.
Smart zoning	Individual rooms or zones heated or cooled to specific temperatures at specific times.	10% less energy use for heating or cooling.
Smart window control	Controls amount of light let through and can block heat or cold.	10-20% less energy use for heating or cooling.
Smart lighting (including occupancy control)	Remote control of lighting, automation, adjustment to occupancy.	1-10% less home energy use; 30-40% less lighting energy use.
Smart plugs	Turns unconnected products into connected devices.	1-5% less home energy use.
Home energy management system	Enhances control and automation of energy-using appliances and equipment.	8-20% less home energy use.
Extensive system of management and automation optimised for energy efficiency	Combination of technologies providing measurement, monitoring, dynamic benchmarking, information displays, management, control, automation, zoning, occupancy systems, maintenance management, etc.	30% less home energy use.
Smart district heating	Artificial intelligence combined with sensors to optimise district heating use in apartment blocks.	~10% less apartment block energy use. ~20% more apartment block peak energy savings

Sources: Adapted from 4E TCP EDNA (2018), Intelligent Efficiency: A Case Study of Barriers and Solutions; Leanheat (2019), More Efficient Energy Systems with Artificial Intelligence Controlled Buildings.

Note: Since multiple technologies can address the same energy end-use, the achievable combined savings potential is less than the sum of the potential of different technologies, however additional savings and benefits can be achieved by intelligently co-ordinating technologies.

Figure 2.14: Residential buildings: Possible benefits of digital technology [5].

Technology	Description	Possible benefits
Artificial intelligence algorithms	AI to predict the future performance of industrial equipment and alert plant operators to potential faults before they disrupt production.	Energy savings of up to 10% in energy-intensive industrial applications.
Advanced/intelligent energy management	A combination of concepts from traditional industrial energy management systems (e.g. ISO 50001) with digital technologies and advanced software applications.	Possible energy savings of over 10-30% depending on industrial process and technology.

Source: Schneider Electric (personal communication); General Electric (personal communication); IBM (2019), *Cognitive Manufacturing: An Overview of Four Applications that are Transforming Manufacturing Today*; IEA (2018) *Energy Efficiency 2017*.

Figure 2.15: Industry: Possible benefits of digital technology [5].

Technology	Description	Possible benefits
Road: connected and automated vehicles	Connected and automated vehicles (CAV) could reduce energy intensity of road transport at the vehicle, fleet, and urban systems levels. However, CAVs may result in rebound effects as a result of lower travel costs and new users.	Vehicle-level energy savings include platooning (up to 25%), eco-driving (up to 20%), and vehicle right-sizing (20-45%).
Road: shared mobility services	Shared vehicles (e.g. bicycles, scooters, cars) and shared mobility services (e.g. ride-sourcing) could help reduce energy use by shifting mobility from private cars to less energy-intensive modes.	Carsharing could reduce individual transport energy use by half. Distances driven, energy use and CO_2 emissions in 2050 are around one-third lower if vehicles are shared, electric and automated.
Road freight	Digital solutions for road freight include: GPS coupled with real-time traffic information for route optimisation; on-board monitoring and feedback for eco-driving; CAV platooning for fuel-efficient driving; and data sharing between companies across the supply chain to ship more goods with fewer trips.	Applying digital solutions to truck operations and logistics could reduce road freight energy use by 20-25%.
Rail	Automated train operations (ATO), communications-based train controls, real-time driver advisory services (DAS) and energy-efficient timetabling can reduce energy consumption by optimising driving patterns, increasing utilisation and promoting energy-efficient driving.	ATO can cut energy consumption by up to 20%, while potential energy savings from DAS range from 5% to 20%. Energy-efficient train timetabling can unlock energy savings of up to 35%.

Source: Taiebat et al. (2018), A review on energy, environmental, and sustainability implications of connected and automated vehicles; Wadud et al. (2016), Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles; Chen & Kockelman (2016), *Carsharing's life-cycle impacts on energy use and greenhouse gas emissions*; Fulton, Mason, & Meroux (2017), *Three revolutions in urban transportation: How to achieve the full potential of vehicle electrification, automation, and shared mobility in urban transportation systems around the world by 2050*; IEA (2017b) *The Future of Trucks*; Scheepmaker, Goverde, & Kroon (2017) *Review of energy-efficient train control and timetabling*; Douglas et al. (2015), *An assessment of available measures to reduce traction energy use in railway networks*; González-Gil et al. (2014), *A systems approach to reduce urban rail energy consumption*; Trentesaux et al. (2018), *The Autonomous Train*; Dunbar et al. (2017), *A tool for the rapid selection of a railway signalling strategy to implement train control optimisation for energy saving*; Urien (2013), *Energy Optimization for Public Transportation Applications*.

Figure 2.16: Transport: Possible benefits of digital technology [5].

2.1.5 Measurement & Validation (M&V)

The Efficiency Valuation Organization (EVO) defines Measurement & Verification as "the process of planning, measuring, collecting and analyzing data for the purpose of verifying and reporting energy savings within an individual facility resulting from the implementation of energy conservation measures (ECMs)" [37]. As savings can't be actively and/or directly measured, they are determined by comparing measured consumption before and after implementing ECMs, while making necessary adjustments for changes in conditions. These activities range from installing/calibrating smart meters to data analysis, automation and finally to reporting savings and quality assessing [37]. Although little to no doubt may exist about what the outcome of a savings project might be, it is wise to verify (more than once) that the solution is, in fact, able to produce the expected savings. Operational verification, which involves inspection, commissioning of equipment, functional performance testing and/or data trending allows the validation of the expected savings [37].

As such, M&V is not, in general, a simple list of tasks needed to meet requirements but instead part of the improvement of the facility operation and savings, being possible to find synergies with other project steps. EVO defined 8 possible purposes of M&V, listed below [38]:

1. **Increase energy savings** by properly evalutaing savings in order to change, improve or maintain ECM design or operations
2. **Document financial transactions** that serve as basis for financial payments and/or guarantee in EPCs in a transparent manner and subjected to independent verification.
3. **Enhance financing for efficiency projects** while increasing transparency and credibility of outcome reports which can increase the confidence that stakeholders have in energy efficiency projects, raising the chances of being financed.
4. **Improve engineering design and facility operations and maintenance** by encouraging comprehensive project design which considers all operational/cosy details. It also helps managers discover and reduce maintenance and operating issues.
5. **Manage energy budgets**, helping managers to assess and contain energy usage to account for variances from budgets.
6. **Enhance the value of emission-reduction credits**, improving emissions-reduction reports.
7. **Support evaluation of regional efficiency programs**, by better auditing utilty or government efficiency programs.
8. **Increase public understanding of energy management as a public policy tool**: enhancing savings increases credibility and public acceptance. Public credibility is expected to boost investment in energy-efficiency projects or emission credits. Good M&V practices highlight that the public benefits from good energy management by improving community health, reducing environmental degradation and increasing employment.

The same organization highlights that best practice M&V plans should be accurate in measurements, conservative when faced with uncertain data predictions, complete with all the relevant information, consistent between different situations or details, relevant in terms of content and transparent [39].

The International Performance Measurement and Verification (EVO, US Dpt. of Energy) Protocol's mission ever since 1995 has been to standardize and develop a framework for good M&V practice, and it is now the most widely recognized protocol in the world [40] [41]. It comprises the fact that savings cannot be acurately and directly measured since they represent an absence of consumption on demand [40]. Instead, it compares measured consumption before and after the implementation of ECMs, with needed adjustments, on a consistent basis (Fig. 2.17). This way, the general M&V equation [40] is used:

$$Savings = (BaselinePeriodEnergy - ReportingPeriodEnergy) \pm Adjustments \tag{2.1}$$

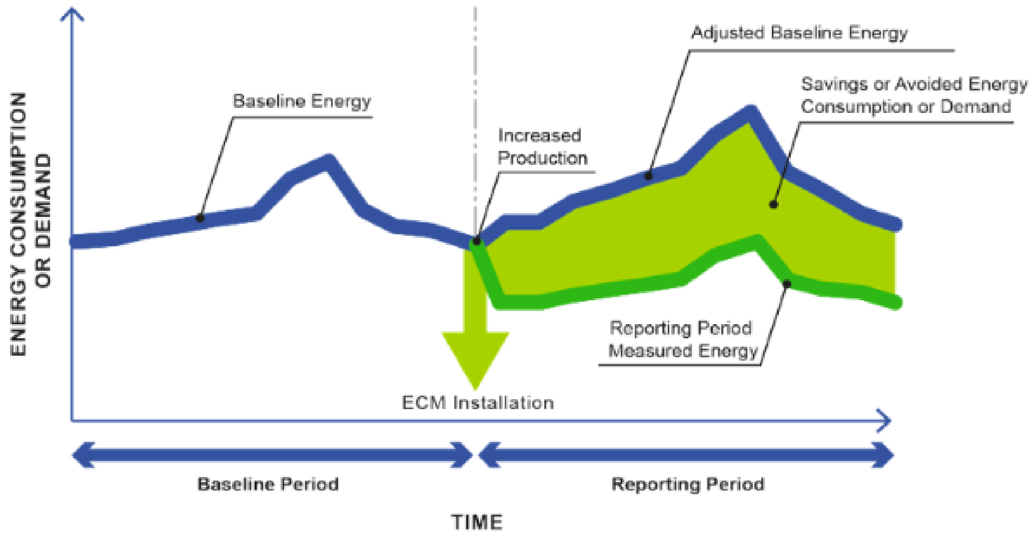


Figure 2.17: IPVMP savings calculation [40]

Therefore, as necessary terms of the previous equation, there's a deep focus on studying and developing a baseline consumption model from existing data and making relevant and regular adjustments which mirror current conditions.

When it comes to baseline, its important to properly define boundaries as to what energy perimeter we are studying and at what period [42]. The baseline period should be representative of a normal operating cycle, ranging from maximum to minimum operation, support ECM planning and coincide with the period immediately before the retrofit [42].

The adjustments should be driven by physical parameters and facts about the energy system and equipment and serve to adapt the model to more accurate real conditions. There are two types of adjustments [42]: Routine - in which energy-governing elements which usually change during the reporting period, like the weather conditions or production volume, and are modeled using either a constant value or a multiple parameter non-linear equation which correlates those elements with the energy consumption - and Non-Routine - like changes in operation cycles or facility size, which are not expected to change over a studied period and therefore need to be accounted for by those static factors. IPMVP offers four options to quantify savings (A, B, C and D) which allows clients to study ECMs from just key parameters to calibrated facility simulations, depending on the amount of data and automated computation available [42].

The recent massive deployment of energy automated systems and real time data handling is the foundation of what authors call M&V 2.0 or advanced M&V [43]. Benefits include a more modern audit system towards savings, which allows for greater data granularity and frequency, with stakeholders easily getting the pulse of the data and accurately taking decisions regarding ECMs, contributing for more credible M&V tecniques and EPCs. However, lack of standardization and regulatory policy is keeping this new age in M&V to reach its full potential and keeping a higher credibility level [43].

Granderson et al. [44] (2016) assessed these recent advanced technologies for predicting adjustments/baseline by testing ten existing models and comparing them to the actual measured data from 537 buildings. They confirmed that the available advanced interval data models hold great promise for the further scalability of M&V 2.0 practices applied with Advanced Metering Infrastructures. All models reached a coefficient of variation of the root mean squared error (CV(RMSE)) under 25% for 1-year and 6-month studies, which reveals that currently there is reason to be confident in the models' robustness [44].

Gallagher et al. [45] (2018) laid the fundamentals of a Machine Learning (ML) supported methodology for M&V 2.0, where data is collected, important features are selected through data analysis algorithms and, then, the baseline model is built from the existing data. In the considered biomedical manufacturing facility, they reached a CV(RMSE) of 11,23% with an optimal k-NN model.

Away from being a mere engineering requirement, M&V has a relevant mathematical background brought together by, for example, Xia et al. [46] (2013) where types of IPMVP options and baseline model/adjustment external parameters' dependance can be derived from the description theory behind M&V procedures and stochastic processes in a straightforward way.

Important metrics are relevant for model behaviour evaluation and model selection. IPMVP relies on the root mean squared error (RMSE) as the standard error (SE) but one other common used parameter in the literature are the coefficient of variation of root mean square error (CV(RMSE)) (Eqs. 2.2, in which y_i is the actual value, $\tilde{y_i}$ the predicted value, $\tilde{y}$ is the average of the actual values, and n is the total number of predictions in the studied period) [45].

$$CV(RMSE) = \frac{\sqrt{(1/n)\sum_{i=1}^{n}(y_i - \tilde{y_i})^2}}{\tilde{y}} \times 100 \tag{2.2}$$

The CV(RMSE) represents the level of variability between actual and predicted values and it shows the magintude of the error relative to the quantity (baseline energy consumption) [45].

2.2 Energy Service Companies (ESCOs)

2.2.1 Definition

Frequently, most facilities and operations face a significant waste of energy which can be avoided, whether it surges from the supply side, poor design or low energy prices. At the same time, energy providers focus almost solely on managing supply and demand, often in detriment of the best energy service value. This is the main goal of an Energy Service Company (ESCO): connecting energy with service in order to deliver the best overall value through waste avoidance by means of energy service upgrades, implementation of energy efficiency programs and savings deliverance to the end consumer [47]. The E3P defines an ESCO as a company that offers energy services which may include implementing energy efficiency projects (and/or renewable energy projects) mostly on a turnkey basis [48]. ESCOs therefore provide technical solutions without the consumer needing to own or pay direct operational costs for those systems, nor understanding the skills required to engage in such projects, making it acessible to the common consumer. This business model, backed by environmental, political and economic push, allows for savings of between 10-30% [47] [11] by deploying strategies such as: changes in lighting, behaviour change, systems control, process optimization and motor/drive upgrades, with most programs delivering payback periods of around 4 years or less [47]. These companies work on a solution based approach, providing the consumer with the best value decision which optimizes all of his objectives, considering both the capital cost of the equipments and their operational costs, while shifting risk away from the consumer side and incentivising energy reduction.

2.2.2 Digitalisation

Most of these energy efficiency projects became feasible due to general digitalisation and Internet of Things (IoT), for instance. These technologies brought greater control and optimization power as well as a highly significant amount of data and analytic power into picture, when concerning the energy sector. These features highly impacted the current connectivity between humans, machines and devices, through sensors and ICT (Information and Communication Technology) and even though more devices can translate into a larger energy consumption, it is possible to minimize the negative impacts associated with it and bring higher efficiency to sectors such as transportation, buildings and industry [15].

Energy efficiency projects are the result of a combination of technologies: sensors and acting devices gather and analyze data to implement changes in energy usage while maintaining dependence between them; then, data analysis algorithms flag or act upon changes needed to secure higher efficiency. A lot of smart appliances and transportation grids are already in use but energy system's efficiency improvements are slowing down and not up to speed with recent developments [15] [5].

2.2.3 Market, Policy and Challenges

Given the need to rapidly and significantly increase financing for energy efficiency, there's a growing interest in this business model: the global ESCO market grew 8% to USD 28.6 billion in 2017 [11]. However, this market remains somewhat underdeveloped in Europe compared to other major regions, accounting for just 10% of the global value, mostly in the non-residential buildings sector and industry [11].

Government policy and access to financial instruments are key drivers of ESCO activity, having an influence on whether projects are carried out in the private or public sector, with higher or lower credit rating. Subsidiary ESCOs usually consist of a small branch of a larger engineering firm or technology provider, such as Honeywell, Siemens or Schneider Electric, operating on a mature market which allows project aggregation and easier access to equity when needing to finance a project. Standalone ESCOs generally focus solely on delivering energy saving measures and have lower grade finacing capability [11].

Even though these companies propose an interesting business model, there are differences in their business development throughout the globe, mostly due to different government policy, contract definition and accessible technology [11] [49]. As such, they are seemingly stalling on a niche application *plateau*, comprising industries or other non-residential buildings.

In 2006, Bertoldi et al. [49] explored the inhomogeneity of the ESCO market throughout Europe despite the efforts of the European Union to foster this industry and laid out a strategy to further push this business models, insisting on increasing ESCO services/projects dissemination, standardizing contracts and M&V methods, developing financing sources for this type of company (such as Third Party Financing), creating an accreditation system for ESCOs and boosting governmental lead in public buildings. Seven years later, Hannon et al. [50] found that the dominance of EUCs (Energy Utility Companies) over the energy market and the energy system was still keeping ESCOs from reaching a broader market stance, but that ESCO's influence on the energy system was expected to grow in the coming years due to a change in the selection environment they operate in, with EUCs creating their own ESCO divisions, signaling a change on the energy sector's *status quo*. More recently, Kangas et al. [51] studied the barriers to energy efficiency as viewed by ESCOs and highlighted some of the same issues: lack of technological skills, disinterest in energy efficiency improvements and non-functional regulation. They also stated that public actors have a key role in overcoming these barriers by creating new possibilities for entrants to take part in decision-making.

2.2.4 ESCOs in Portugal

Bertoldi et al. [49], identified 7 major ESCOs operating in Portugal, as of 2006. However, Portugal is quoted as an ESCO emergent market in Soares et al. [52]. Growing steadily but at a slow rate, these companies have been thriving since 2008, mostly due to the PNAEE (National Action Plan for Energy Efficiency) - a plan to cut energy consumption in around 9.8%. In 2009 there were 10-12 companies which identified themselves as ESCOs, but that number grew to 89 in 2011 [52], thanks to plans like ECO.AP, which aimed at a consumption reduction of 20% in public administration facilities, partnering with ESCOs. Most companies reported to act mostly on the industrial sector and also in hotels, offices and swimming pools, preferably using non-performance contracts. After the financial crisis, some incentives were discontinued (raising VAT from 13% to 23 % on renewable energy gear and cutting renewable incentives of up to 30%) , which is expected to have had a stalling impact on the national ESCO market. Findley et al. (2018) [53] and Capelo et al. [54] showed that portuguese barriers to ESCOs thrival do not operate independently and must be addressed system wide, listing Portugal as one of the lowest maturity markets studied (22% maturity score, 65% being the maximum (Mexico)). It is also shown that this market is valued in USD 35 million in the country and that the number of ESCOs peaked in 2013 with 100 companies and dropped abruptly to 5 in 2017, according to the IEA [53]. In conclusion, the ESCO market is shown to be at a low maturity stage in Portugal and it is not expected to have a significant market share on the residential sector, since energy costs are highly taxed, which limits a significant cost reduction. Most (envelope) retrofits also aim to deliver a greater degree of thermal comfort, rather than energy savings.

2.3 Energy Performance Contracts (EPCs)

An Energy Performance Contract (EPC) is a commonly used ESCO mechanism for financing energy improvements, which allows for upgrades to be funded by the potential cost reduction they generate. Under the jurisdiction of this contract, the company implements an energy efficiency project and allocates cost savings/renewable energy produced to repay the capital costs of the project and its investment, which shifts the responsability of delieverance (of expected savings) to the ESCO side [10].

This type of contract is advantageous for consumers since it shifts the technical risks to the ESCO side, based on expected performance, that is, ESCOs' repayment is dependent on the attained level of energy savings. It is, then, a way of delivering infraestructural developments in enviroments where there is lack of manpower, engineering skills, management time or capital funding making it a creative business model for underdeveloped countries with some degree of creditworthy organizations wishing to improve energy efficiency [10].

There are two main types of EPCs [10]:

1. **Shared savings** in which cost savings are shared (for repayment purposes), for a determined length and in a pre-determined percentage, between the lender, consumer and ESCO, depending on the capital cost of the project, contract length and associated risk.

2. **Guaranteed savings** in which the ESCO guarantees a certain level of energy savings from the EPC in exchange for an upfront payment and/or regular service fee, shielding the client from performance risks.

2.3.1 Shared savings

Under this contract, the client assumes some degree of performance risk and, as such, credit risk is at least partially guaranteed by the ESCO, which can ultimately put these companies in default risk, since they become too indebted to keep lending money from financial institutions, hence this being a type of contract in which short payback times are preferred. Thus, the shared savings contract helps boosting the market, since consumers assume little to no financial risk, but it can create barriers for smaller companies which will not be able to compete with highly leveraged ESCOs when needing to contract further debt for new projects [10].

Some variations of these contracts include the "first out", in which ESCOs are paid 100% of savings until the project costs and profit are fully covered, the duration being variable according to the level of savings attained; and also the "single energy price", in which both parts agree on the value of the service upfront and neither side gains from energy price differences that come about during the duration of the contract. In this latter case, if the actual prices are lower than the floor value, the consumer has a windfall profit that compensates the lower return. If the prices rise higher than the ceiling level, the return of the project is higher than expected but the consumer doesn't pay more for it, which resembles the guaranteed savings contract to a certain extent [10].

Often, when savings exceed expectations, an adversarial relationship between consumer and ESCO can be created (and that is, to a certain extent, the focus of this thesis), in which ESCOs try to underestimate savings and receive more from excess savings, which usally are contractly due to them [10].

2.3.2 Guaranteed savings

This type of contract places the performance and design risks under ESCO's responsibility, that needs to guarantee a certain level of energy savings to the consumer. The credit risk is usually undertaken by the client, through access to bank credit or financial institutions. Because of that, it is usually prefered in countries with a strong banking structure and high degree of familiarity with project management and technical expertise financing in the energy sector, which are usually seen as barriers for new markets, where the costumer tends to take investment risk.

In this case, if the savings come short to cover the debt service, ESCOs have to step in to cover the difference. On the other hand, if the savings exceed the estipulated level, the costumer pays an agreed percentage of the savings to the ESCO [10].

2.3.3 Implementation

On average, EPC projects have a capital outlay of EUR 1-5 million and a contract length of 5-10 years [47]. A typical EPC process goes through five implementation stages [47] [55] [56]:

1. **Identifying Requirements, Resources and Funding Approach**: the first stage focuses on identifying the potential project by defining what is the end goal of the project, the requirements, the resources needed and how will it all be funded. It also comprises contractor procurement approaches.

2. **Tender Phase** In order to deliver a proper tender package, gathering information on the properties within the scope of the project will be important to convey the basis of the opportunity. Bidders will be keen to understand areas such as energy use, current building condition, recent and future planned improvements and operational use. It is then a crucial phase to decide on feasibility, procurement routes, financing sources and, ultimately, to engage suppliers and select preferred partners. Can take 6-12+ months, depending on requirements and procurement.

3. **Investment Grade Proposal (IGP)** at this stage, ESCOs should have a detailed proposal consisting on the technical solution, contract structure, costs and benefits. It should also establish a contractly binding M&V protocol as well as detail the funding (internal or third party) confirmation. This stage involves a thorough survey of buildings so it is important that clients engage with the

appointed ESCO so it can develop a better understanding of building use. It might be useful to discuss ECMs being considered so that the final proposal is more likely to fit with requirements. Takes around 8-10 weeks.

4. **Energy Conservation Measures Installation** when the IGP is completed and agreed, the ESCO procedes with the instalation of the ECMs, starting to work on site. Depending on the complexity, this can take from some weeks to over a year.

5. **Service Delivery** Energy savings are a key benefit that should be achieved and operational and maintenance services may be provided by the ESCO. It is important that these savings are monitored through an appropriate measurement and verification plan. This should enable changes in operational and external factors to be accounted for in demonstrating overall performance. If underperformance does occur, aditional measures may be applied in line with the agreed contract.

2.3.4 Risks and Best Practice

In 2015, Lee et al ([57]) surveyed ESCOs and clients to understand EPC risks as perceived by each part and further elaborate measures to enhance the adoption of EPCs in the future, as well as to mitigate risk, agreed on by the respondents. His findings together with information from [47], are listed below (Table 2.1):

To ESCOs	To both parties	To Clients
Payment default after installation	Energy price changes	Long payback periods
Uncertainty in baseline measurement	Weather impacts	Project complexity
Instalation costs increase	Major changes in building use	Repayment ability
	Poor M&V protocol	
	Poorly documented assumptions	
	Poor energy audit	
	Poor contract	
	Economical environment and public policies	

Table 2.1: Main risks associated with EPCs [47] [57]

Back in 2004, Yik et al. [58] have listed the most relevant obstacles to energy performance contracting which still remain relevant today, which were: disputes over M&V protocols, changes in energy prices, occurences of extreme weather, changes in the building operation pattern, client perception that targets won't be met and hesitation/high interest rates from financial institutions, due to uncertainty in energy cost savings or payback ability. They've also exposed guidelines to access the viability of an EPC and the consequences of mismatches between expected and actual energy cost saving, culminating on the idea of a partnership firm between parties in order to solve conflicts.

In order to tackle risk perception and mitigation, several guidelines or success factors in EPCs have been studied over the years, so as to gather both parties' consensus. Lee et al. [57] drew some key measures to incentivise EPC adoption such as: promotion of successful projects, changing the government procurement practices to facilitate the use of EPCs and government partial backup of loans. They also highlighted the need to promote the true value for money of EPCs amongst building owners. Xu et al. [55] also established 5 critical success factors (CSFs) for EPCs in China, consisting on: accurate M&V, trust between parties, a control mechanism of sustainable development strategy, available technology and effective coordination.

On a bibliometric analysis about EPC research between 2008-2018, Zhang et al. [59] found 5 mainstream research topics, amongst 127 papers mostly from China and the United States: implementation, mechanisms for effective EPCs, risk management and ESCO/stakeholder behaviour in EPCs and decisions. When considering research challenges or further work, the paper found 6 research challenges: lack of an effective M&V, limited studies in the residential/buildings/tertiary sector, lack of risk prevention, limited research on dual stakeholder relations, poor consumer attitude towards ESCOs (especially in the EU) and ineffective mechanisms to ensure post-EPC savings.

As expected, much of the recommendations of researchers lie within the field of regulatory and transparency issues. Nuñez et al. [60] say that one of the key issues regarding funding for EPC projects in South East Europe is the need for a simpler contract model that assures transparency and the urging public institutions and general public to further regulate and boost this practice through creation of associations that gravitate towards EPCs. When Frangou et al. [7] studied the impact of these contracts on the tertiary sector in Southern European countries, that usually have a hard time securing credit-rating, researchers found that these financial barriers have a negative impact on the implementation of these contracts and appoint reasons for it: scarcity of examples, lack of knowledge and skills, lack of financial expertise or products designed for ESCOs, high guarantees, poor legislation and interest rates. Applying their studies to the Greek scenario [8], authors exhorted the community to support legislation changes and standardization, since that is expected to have a significant impact on credit, which is key for EPCs.

2.3.5 Successful Examples

Ever since they first appeared to the public in the 1980s, EPCs have been increasingly thriving as mechanism to secure energy savings in most sectors. European Union's programmes like SUNsHINE (Latvian Baltic Energy Efficiency Facility) aim to boost EPC adoption by easing the financial process inherent to these contracts and presenting a formal investment guide [61]. The programe lasted for 1 year and a few months after the beggining, a total EUR 12.8 million investments were foreseen and 1300 MWh energy savings per year or 26 000 MWh in 20 years [62].

The Ener-In-Town wished to establish greater control over energy consumption in municipally-owned buildings, by providing a solution which eliminate two existing barriers: lack of detailed knowledge on consumption figures and lack of qualified local personnel with the capacity to propose improvement actions [63]. The execution results yielded savings of around 16% on sports buildings, 6% on swimming pools and public facilities, 9% on office blocks and 8% on residential/school buildings, across 32 european municipalities.

Particularly, in Portugal, there's a continuing investment to cut energy consumption on buildings, expressed by the PNAEE [64] (Plano Nacional de Ação para a Eficiência Energética - National Plan of Action For Energy Efficiency). The plan proposed cuts of 10 and 20% in building consumption over the years, and it's considered to be a change engine in the country, with programs like ECO.AP [65] deploying ESCO activities throughout several public buildings.

Austria and Germany also report a distinct acceptation of EPCs in Europe [66].

Overseas, in some of the world's most industrialized countries like the US and China, there are also success case studies of EPC implementation. The Empire State Building retrofit project started in 2008 and hoped to reduce the building's energy consumption pattern. It focused on improving the efficiency of the building's exterior side, changing the ventilation system to a smarter one and promoting efficient lighting and behaviour changes [52]. In 2012, it surpassed the energy savings expectations by saving 38% of the building's energy consumption, which translates to USD 4.4 million savings per year, all while creating 252 jobs [67]. The impact on lighting and ventilation matched to around 50% of the savings [67]. A report by the Clinton Foundation highlights four reasons why the Empire State Building Retrofit Model works, them being: ownership with credibility, replicability to other buildings, increased transparency and the attention drawed towards the famous building [68]. Federal US Department of Energy programs like the Federal Energy Management Program, also helped foster EPC adoption, by implementing them to enhance public buildings [69].

On the other side of the world, China also reports a similar degree of success in the application of EPCs in hotel/building retrofits [55], with [17] highlighting contractual bond as a main barrier to a better framework.

2.4 Energy Consumption Forecasting Models

Predicting energy consumption as a variable is just as any data forecasting routine (Fig. 2.18): first, the data is gathered from meters and treated, getting split and "cleaned" through data analysis algorithms. Then, other parameters (features) dependence is studied through feature selection/data clustering routines, which quantify and qualify the degree of relevance of the parameters to the evolution pattern of the consumption load. Finally, data predicting models are trained with historical data (of both x - features - and y - energy consumption), so as to create a mathematical description for the evolution of the features with the energy consumption. After the model is trained, it is able to predict energy consumption values from new feature data sets, with a measured uncertainty. The two main parts of this process are described in the sub sections below.

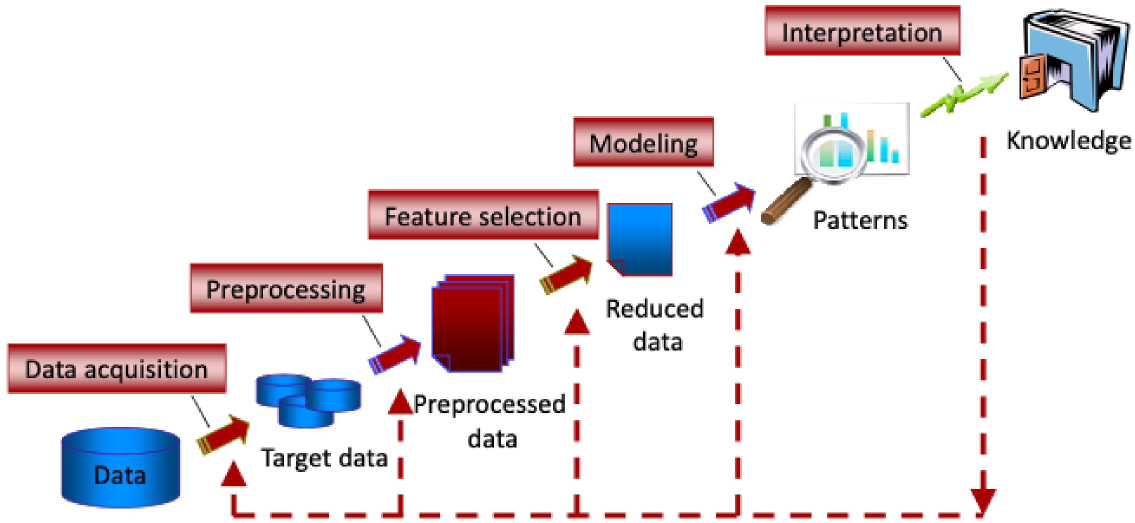


Figure 2.18: Data discovery process [70]

2.4.1 Feature Selection

This stage represents the process of identifying relevant features to the evaluation of a certain variable of study in order to reach a model with better accuracy, being indispensable for Machine Learning (ML) data processing [71]. This way, data used to train forecasting models is better fit by removing redundant, irrelevant or loosely relevant variables from the data set, reducing noise predictions, model complexity and computing time. Mainly, feature selection is about determining the best subset of candidate features, by spotting the irrelevant or weakly relevant ones, following one of the three approaches: optimizing an evaluation measure (Filter approach), satisfying restrictions on those measures (Wrapper approach) or reaching the best commitment between subset size and evaluation measure (Embedded approach) [71].

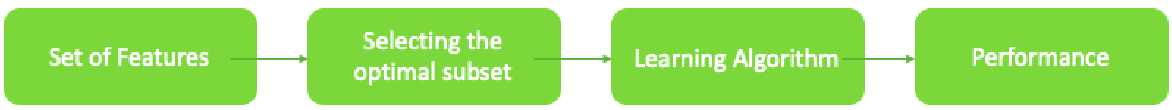


Figure 2.19: Filter approach [72]

Filter methods (Fig. 2.19) are widely used as preprocessing steps and they select candidates (ignoring the prediction model) by evaluating general parameters like correlation or statistical test scores (Pearson's Correlation, LDA, ANOVA or chi-squared) in regard to the modeled variable and suppressing the least relevant features after a certain stopping criterion is met [72] [71] [73]. The method doesn't consider relations between features and, as such, becomes highly efficient in reducing computation time and avoiding overfitting. Selecting redundant variables is frequent and, therefore, improved versions of this method try to eliminate features highly correlated to each other [73]. Some examples of these algorithms include FOCUS, ABB and relief [71].

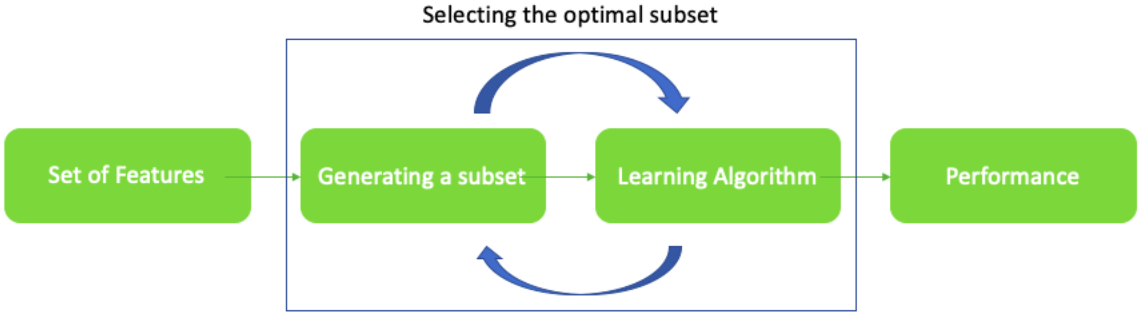


Figure 2.20: Wrapper approach [72]

The Wrapper approach (Fig. 2.20) is quite similar to Filter apart from the evaluation criteria and that it can identify interactions between variables. It works by using a subset of variables to train a model and, based on the results, deciding to add or remove features from the previous subset [72]. The main setbacks of this method is that it increases computational time and the risk of overfitting. It chooses the optimal subset of variables that are better suited for the learning algorithm and usually presents a better behavior too [71] [73]. Common examples of wrapper methods are Forward Selection, Backward Selection - the first starting with no features and adding them until an addition stops improving the model and the last doing the same thing backwards - and Recursive Feature Elimination- which works by repeatedly creating new models and labeling the best and worst performing features until they all are exhausted, presenting a ranking of features by elimination order [72].

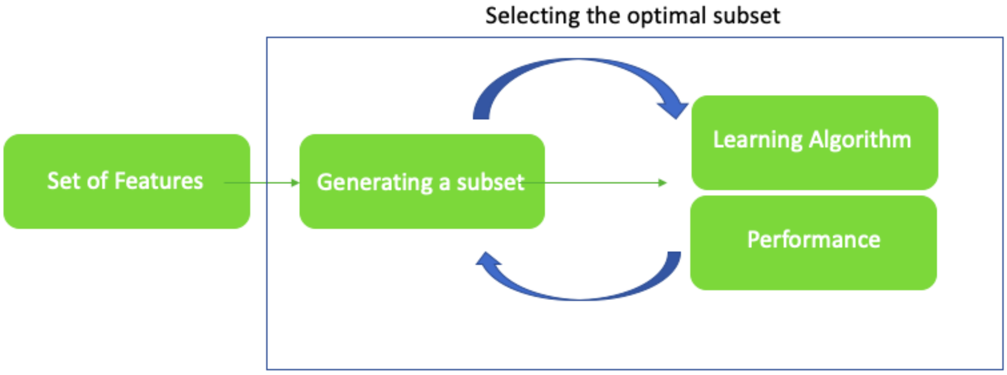


Figure 2.21: Embedded approach [72]

Finally, the Embedded approach (Fig. 2.21) combines the advantages of the latter two: lower computational time cost and feature dependency capture. In fact, not only it studies dependencies between input and output variables but it also searches locally for features that are important to improve local discrimination. A prediction-independent criteria decides the optimal subsets for a certain cardinality and then, the learning algorithm selects the final optimal subset among the optimal subsets across different cardinalities [71] [73]. Some popular examples of these methods are LASSO and RIDGE regression which consider penalization functions to reduce overfitting [72].

Some real world practical examples of feature selection include remote sensing and browsing of images, text classification, genomic data analysis and intruder (spam) detection [71].

2.4.2 Regression

Predictive modeling is the process of creating a model that follows the evolution of a certain variable, based on historical data, in order to predict the future behavior of that variable [74]. It is then a mathematical problem of approximating (training) a function from input variables (X) to output variables (y). In the scope of predictive modeling, there are classification and regression problems [75]. Classification focuses on labeling parts of the data, splitting it in groups by probability of belonging to it, depending on the application goal. It consists of a discrete variable study in which results are evaluated by the accuracy of the model [75]. In this dissertation, the focus are turned towards the problem of modeling a continuous quantity (energy consumption of a certain building) from multiple input parameters, which known as a multivariate regression predictive modeling. The skill of these models is usually assessed by knowing the root mean squared error (RMSE) or the mean absolute error (MAE), which constitute error parameters in the same units as the predicted value [75]. Some of the most common regression algorithms are Linear Regression (LR), Random Forests (RF) and Extreme Grandient Boosting (XGBoost).

Linear Regression attempts to describe the relationship between variables by fitting a linear equation to the observed explanatory variable (X) [76]. It is usually preceded by some form of data correlation study in order to avoid overfitting (that is, the model becoming to adjusted to the data and, therefore, unstable). Least squares is the most common method for fitting a first degree linear equation to observed data by minimizing the sum of the squares of the vertical (y) deviations from each point to the modeling function [76]. LR avoids cancellation between positive and negative interference by first squaring deviations and then summing them.

Random Decision Forests Regression [77] works by constructing several decision trees at training phase and outputting the average prediction of the individual trees. Studying a random selection of features, it uses a technique called *bagging* or Bootstrap Aggregating [78], in which n subsets of the data fit n models, in order to reach a prediction, which is then averaged from those n model results, reducing variance and avoiding overfitting. Decision trees (Fig. 2.22) [79] work by splitting data into trees that classify the behavior of the output variables (Y) based on the range of values of exploratory variables (X), until a variance minimization algorithm reaches its optimal point. This method is widely used for its little configuration requirements and reasonable predictions, especially in multivariate problems, which come just about short of gradient boosted trees.

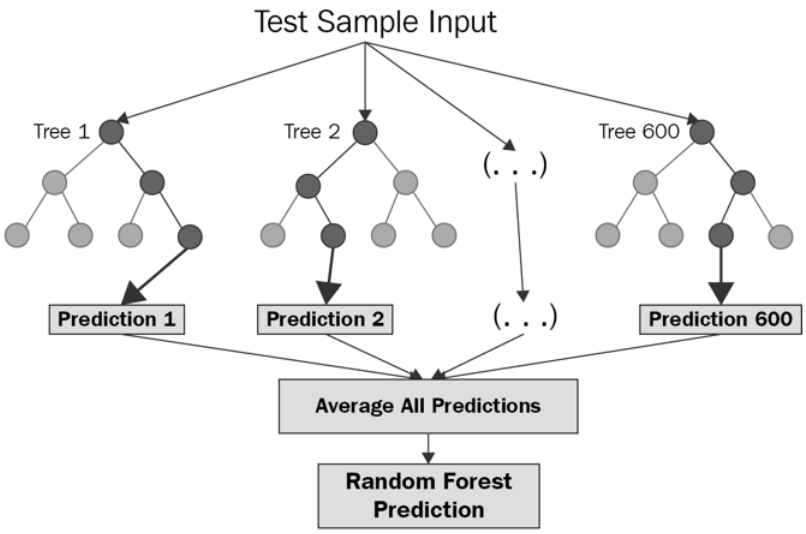


Figure 2.22: RF decision trees [80]

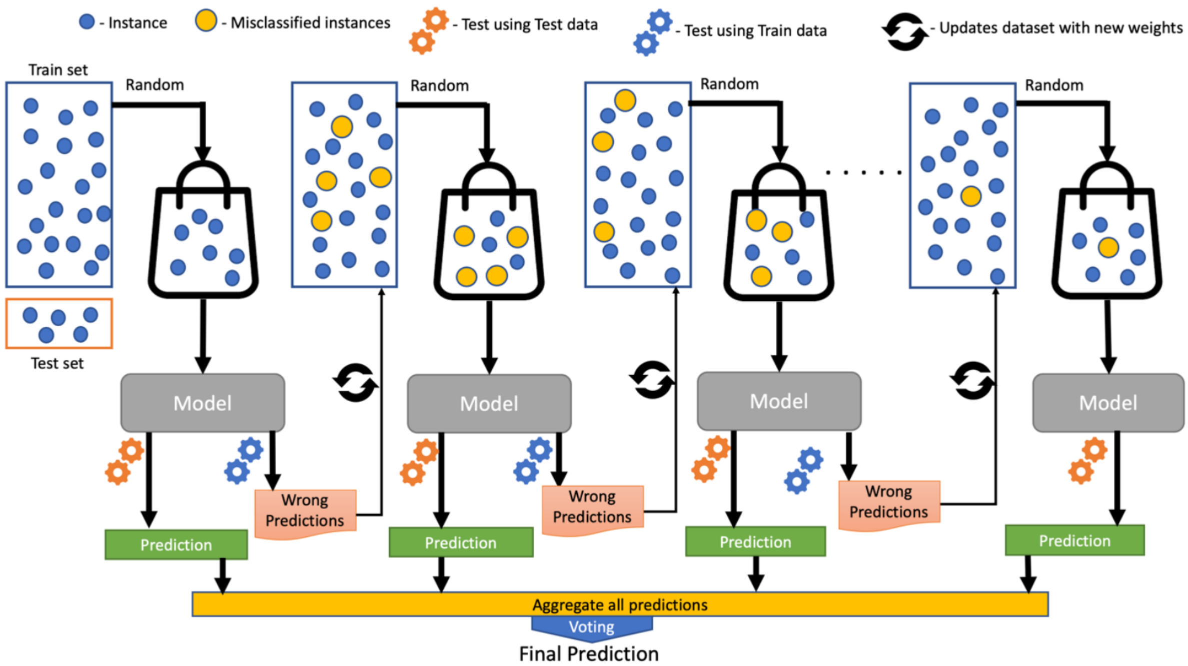


Figure 2.23: Extreme Gradient Boosting prediction routine [81].

Extreme Gradient Boosting (XGBoost, Fig. 2.23) [82] was first proposed by Chen and Guestrin and its central idea of *"boosting"* combines gathering all "weak" learners in order to feed a "strong" learner, simplifying objective functions that join predictive and regularization terms [83]. Put in a simple way, XGBoost fits one model to the whole input data subset and another to this one's residuals (deviation from data to model). It then averages variations and data predictions to induce corrections on fitting predictions [83]. XGBoost is known for efficiently preventing overfitting while optimizing computational effort and using automatic parallel calculation, making it one of the new most promising regression algorithms for data prediction [83].

2.5 Blockchain

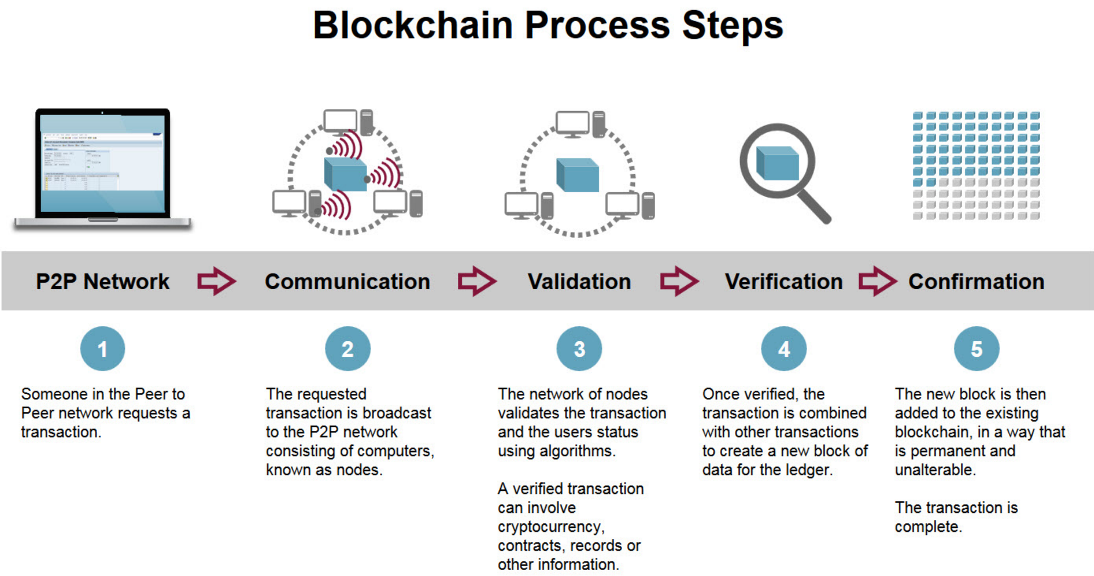


Figure 2.24: Blockchain flow process [84].

First proposed in 2008, Blockchain is a technology which can be briefly explained as a public book-of-records (ledger) in which all committed transactions are stored in blocks. The chain is then formed and expanded as new blocks are appended to it. The security advantages of this technology come from asymmetric cryptography and distributed consensus algorithms in each node (the ”poster”), which maintains a consistent chain. It's a topic of research for several sectors where efficiency can be improved and costs reduced since it is characteristically decentralized, anonymous, persistent and easily auditable [85].

A Blockchain is a shared list of recorded transactions between parties in a network, in which actors with permission can act upon without the control of a central authority. It's made up of nodes which store the list of transactions, called ledger. In this structure, a block is a group of transactions from the same time period. Each block has the data it contains, the timestamp and certain hashes that identifiy each block and the previous one, in order to keep an theoretically immutable chain. To add a block it may need first to be mined (created) by the node and then approved by a number of nodes, through a consensus mechanism [86] (Fig 2.24).

Allowing payments to be handled without a third party (bank or intermediary), Blockchain has attracted financial services' attention for managing digital assets, remittance and online payments, the most famous application being the *Bitcoin* cryptocurrency [85].

The Blockchain ledger is virtually immutable and tampering with transactions becomes quite difficult for intruders, that's why its applications reach Internet of Things (IoT), security systems, smart contracts and public services. Businesses which need to keep reliability and honesty levels are also keen to attract new customers using these technologies, since it is distributed and immune to single point of failure situations [85].

Although it presents great potential for network systems, it also faces technical issues of scalability (*Bitcoin* can only handle 7 transactions per second, which is unfit for high frequency trading, and raising block size to more than 1MB would induce slower network propagation, for example), selfish mining strategy (where miners hide their mined blocks and achieve larger revenues than their fair share) and even shown privacy leakage. In addition, the most famous consensus algorithms, Proof of Work (PoW) and Proof of Stake (PoS), are either highly energy consuming or generate the ”rich get richer” phenomenon [85].

2.5.1 Nomenclature

Blockchain systems are split into three types of applications: public, private and consortium blockchain. The main difference between them is the visibility of the records and consensus process role, as seen in Table 2.2.

Characteristic	Public	Consortium	Private
Consensus determination	All miners	Selected set of nodes	One organization
Read permission	Public	Public or Restricted	Public or Restricted
Immutability	Alm. impossible to tamper	Could be tampered	Could be tambered
Efficiency	Low	High	High
Centralized	No	Partially	Yes
Consensus process	Allowed to all nodes	Permissioned	Permissioned

Table 2.2: Comparison chart between public, consortium and private blockchain [85]

Let us now clearly define these characteristics and distinguish between each type of blockchain [85].

1. **Consensus determination**: tells us who has the power to take part in the consensus process, ranging from all nodes (in the public blockchain), some nodes (consortium) or just one organization (private) which the determines the final consensus.

2. **Read permission** answers the question: ”to whom are the transactions visible?”.

3. **Immutability**: when working with a large network, such as in a public blockchain, it becomes nearly impossible to tamper with transactions since records are stored on a large number of nodes. The inverse happens in more reduced configurations.

4. **Efficiency**: propagation latency becomes higher in a network with a large number of nodes, generally making private and consortium blockchain systems more efficient than public networks.

5. **Centralized**: more restrict networks with fewer users become more centralized by defintion, making public blockchains the most honest example of a decentralized ledger.

6. **Consensus process** answers ”who can join the consensus process of this network”.

Public and private blockchain are attracting many users and communities in all kinds of applications. Consortium systems have significant potential in many business applications, with Hyperledger and Ethereum providing tools and frameworks to establish these networks [85].

2.5.2 Consensus Algorithms

Establishing consensus between nodes on a blockchain is a variation of the Byzantine Generals problem. It states that a group of army generals circle a certain city: some prefer to attack and the others prefer to retreat. However, the attack would only be successful if all generals attack the city together. Thus, they have to reach an agreement in a distributed environment, just like a blockchain, in which there is no central authority that ensures that the distributed ledgers are all the same [85]. Therefore, some protocols are required to ensure ledgers stay consistent between nodes. Here are some common approaches to guarantee consensus in a blockchain [85]:

1. **Proof of Work (PoW)**: as the name suggests, when a node wants to post a block of transactions, it has to undergo a proof of work done to ensure that the node is not likely to attack the network. This is the consensus mechanism used by *Bitcoin*. Each node calculates a hash (coded) value that identifies the block, containing a nonce, which is frequently changed to get different hash values. The process requires that this calculated value be equal or smaller than a certain given value. Whenever the node reaches the right value, it transmits it to the other nodes, which then verify the righteousness of the hash value. When validated, miners would append this new block to their registered blockchains. Nodes that perform calculations are called miners. Due to its decentralized nature, valid blocks might be created at the same time when multiple nodes find the right value nearly at the same time. As a result, branches can be created, where the chain has two valid options to establish consensus, while it is unlikely that two competing branches will generate the next block simultaneously. When this happens under PoW, the longer chain is judged as the authentic one. This method is highly resource wasting.

2. **Proof of Stake (PoS)** is a less energy-consuming alternative to PoW. Mostly used in financial systems, the miners following PoS have to prove ownership or right to post transactions. It becomes quite unfair since the richest/most credible person on the network is bound to be dominant, even though it is believed that people with more digital currency are less likely to attack the network. As a result, several approaches to PoS use stakes proof combined with other metrics such as currency age, lowest hash. It is more effective than PoW, but as the mining cost becomes near zero, attacks might become a constant reality. Many blockchains start using PoW to then gradually change to PoS, like Ethereum and Peercoin did.

3. **Practical Byzantine Fault Tolerance (PBFT)** is a famous replication algorithm that tolerates byzantine faults, by tolerating up to 1/3 malicious byzantine replicas. This stragtegy determines new blocks by rounds, composed of three stages: pre-prepared, prepared and commit. In each phase, a node needs 2/3 of other nodes' approval to enter the next one, requiring every node to be known by the network. It is used by Hyperledger Fabric.

4. **Delegated Proof of Stake (DPoS)** improves transaction confirmation time by having stakeholders elect delegates that generate and validate blocks. Dishonest delegates are easily voted out. Used by Bitshares.

5. **Ripple (currency)** determines collectively-trusted subnetoworks within the larger one. The nodes are divided into servers - which participate in the consensus process - and clients - for transferring funds. Each server contains a Unique Node List, that it queries upon logging a transaction into the ledger. The transaction is valid if agreement reaches 80%, keeping the ledger safe for under 20% of faulty nodes in the UNL.

6. **Tendermint (currency)** which is quite similar to PBFT but nodes have to lock their balance (coins, interest) to become validators. Faulty validators are punished.

In short, different consensus strategies vary on node identity requirement, energy consumption and tolerated faulty nodes [85]. Recent advances focus mainly on splitting block creation and transaction confirmation roles, block generation speed stability and valid branch election [85].

2.5.3 Challenges and Future Directions

Currently, blockchain faces issues of scalability, privacy and selfish mining. All these topics and possible solutions are studied in [85].

Possible future directions in blockchain research suggest an increased focus on blockchain testing and standardiztion, stopping centralization tendency, big data management/analytics and new blockchain applications, cited in [85].

2.5.4 Energy Sector Applications

It is only natural that the recent hype surrounding blockchain technologies has reached the energy sector. In terms of concept, the energy blockchain network is briefly defined as every energy subsystem constituing a node, with energy transaction records collected in each block. Smart contracts are responsible for automatic transmission, validation and execution of relevant information [87]. The main characteristics of this type of network are that they are [87]:

1. **Descentralized**, based on peer-to-peer (P2P) energy networks, with data being stored in each node;

2. **Transparent** in terms of rules and data/transaction visibility;

3. **Automatic**, ensuring that the systems run with little to no human intervention;

4. **Traceable**, since data posts are signed and timestamped. Modifications require node consensus and confirmation.

Zhao et al. (2019) [87] has found the United States to be a good example of a favorable environment for the development of energy blockchain technologies. They presented examples of application that range from microgrid energy trading, downed poles (acting as nodes) control and shared electric vehicle (EV) charging pile systems. It was suggested that in order to foster deployment, these practices should have good policy support, favorable financing environment, advanced technical equipment access and an active *tech-geek* atmosphere [87].

Several other articles study the application of blockchain to P2P energy trading and its business model, such as [88], [89], [90] and [9].

Blockchain's usage in the energy efficiency industry was well studied by Khatoon et al. (2019) [19]. Just like this dissertation, they claim a vast territory for blockchain to explore when concerning EPC application by ESCOs. The highlighted benefits of an EPC application blockchain were: savings exchange and encryption, increased transparency, reliability and security, improved client trust and market success, transaction cost reduction (from no intermediaries being involved) and properly valued savings [19]. They study two examples of application: the italian White Certificates and the UK's Energy Company obligations scheme [19].

Chapter 3

Energy Modeling

3.1 Exploratory Data Analysis

First, we analyze collected energy consumption data from 2017-2018 by smart meters deployed throughout four Instituto Superior Técnico - Alameda campus' buildings: Civil, Central, North Tower and South Tower. Most considerations will take the Civil building as basis since it is the most representative one of the whole campus, in terms of data homogeneity and consumption pattern. We need to visualize the data and, for that purpose, a bar histogram of the recorded hourly values of consumed power for Civil was plotted (Fig. 3.1).

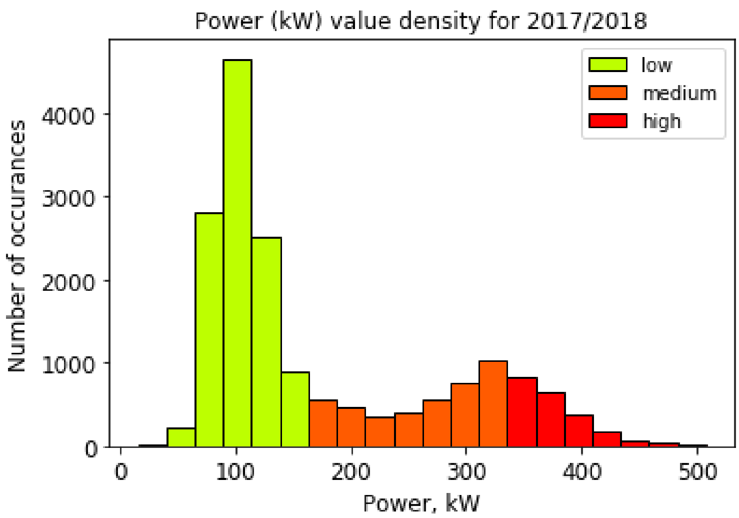


Figure 3.1: Hourly power consumption (KW) histogram for the Civil building, from 2017 to 2018. Two highly populated regimes of consumption with a transient one in the middle are distinguishable - a bimodal type distribution.

It is clear from this analysis that there are two immediate regimes of consumption - a peak higher one and a flat lower one. The two consumption regimes constitute a bimodal-type distribution with a lower transient regime in between them. Most counts lie on the 50-150 kW range, with half of the values laying under the 110 kW mark, which goes accordingly with the average hourly values of consumed power in kW, listed below (Table 3.1):

	2017	2018
Power (kW)	164,3	163,7

Table 3.1: Average hourly power drain (kW) for Civil building. There was a 0.4% reduction from 2017 to 2018.

Let us now take a look at a 10-day consumption pattern by hour, in Civil building, so as to explore the type of consumption cycle present at this university. For that, let us plot the hourly values of power with the respective timestamp, on Figure 3.2.

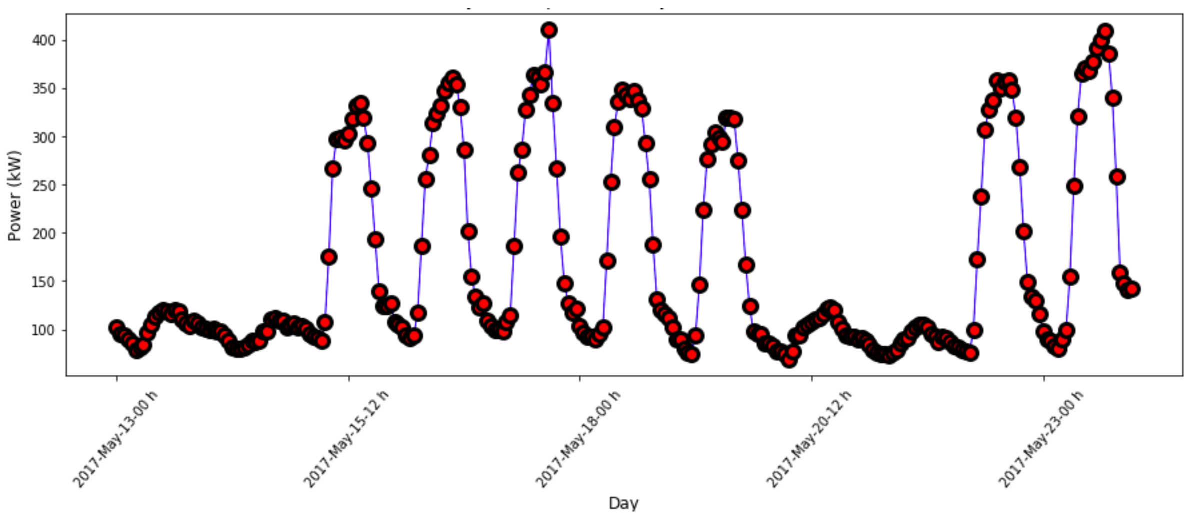
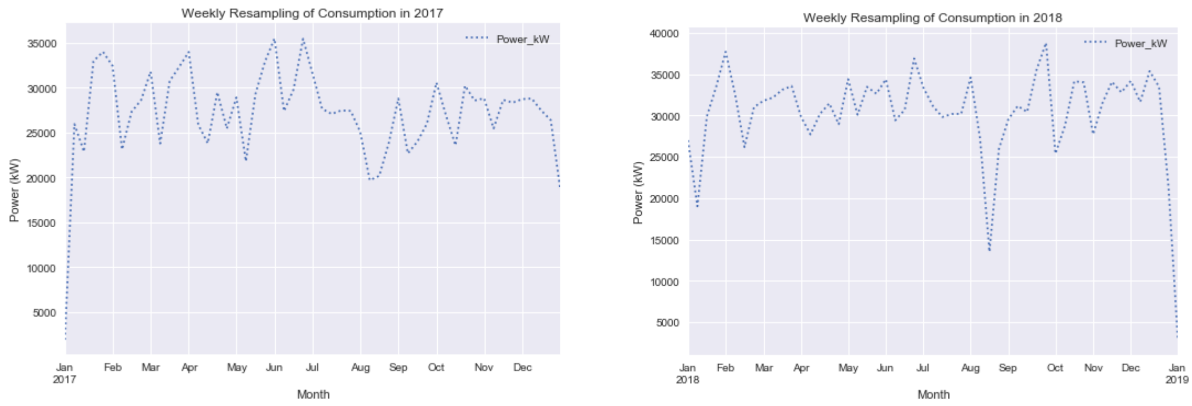


Figure 3.2: Scatter plot of the hourly power consumption (KW) pattern for the Civil building from 13 to 23 of May 2017. Two patterns of consumption arise: normal operations days where the power consumption peaks at around midday and weekends/holidays when there is significant power consumption reduction.

It is possible to see that there is low consumption values on weekends and one could retrieve that this behavior would also occur on school holidays (mainly in the month of August, when there is little to no activity in the campus). There are then two types of consumption patterns: business days (where the school goes about its activities as usual) and weekends/holidays (where the values drop to around 25% of the peak power consumed on business days). All this inferences were verified to be valid for the year of 2018 too.

In order to confirm this second lower pattern of consumption, let us resample the data to a weekly basis, for both years, and see if the values actually drop in key months like January and August (school holidays).



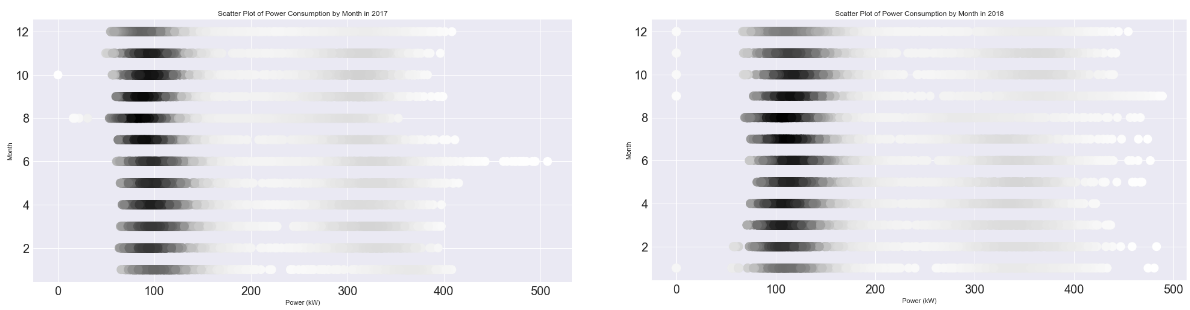
(a) 2017

(b) 2018

Figure 3.3: Weekly sampling plot for the power consumption at Civil in 2017 and 2018. Power consumption clearly drops during school exams (January and June) and holidays (August).

From observing the graphs, one can clearly identify the school holidays drop pattern (in January and August) and even a middle, less pronounced plateau regime on exam season (January/February and June/July), coming from idle operations in class buildings, like this one. It is then safe to say that August records the lowest values of consumption throughout the year.

This way, one already knows how the consumption varies throughout the year, its months and weeks. But from Figure 3.2, it is clear that the daily pattern of consumption starts climbing towards the maximum peak consumption value, which occurs around lunch time, and then goes down again, towards idle power levels ($\approx$ 100 kW). Let us now explore that idea, by analyzing scatter and plateau plots for hourly and daily consumption, respectively (Fig. 3.4).

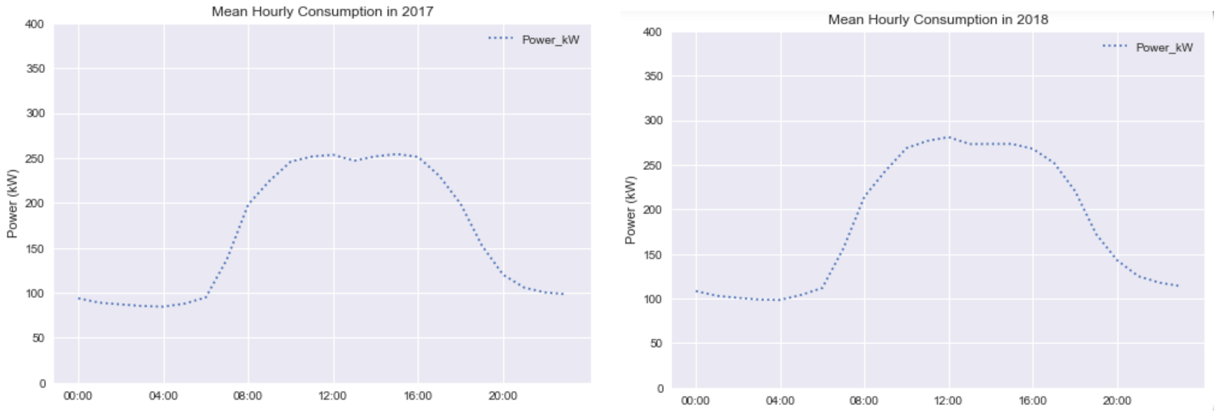


(a) 2017

(b) 2018

Figure 3.4: Density plots for hourly power consumption at Civil building in 2017-2018. Here it is possible see again two highlighted consumption regimes.

There are clearly two densely populated regions of data, just as witnessed on Fig. 3.1 - the heavily dense $\approx$ 100 kW idle regime and the $\approx$ 300 kW power peak, happening every day. Let us now see how that behavior is expressed in terms of mean hourly power consumption, in the graphs below:



(a) 2017

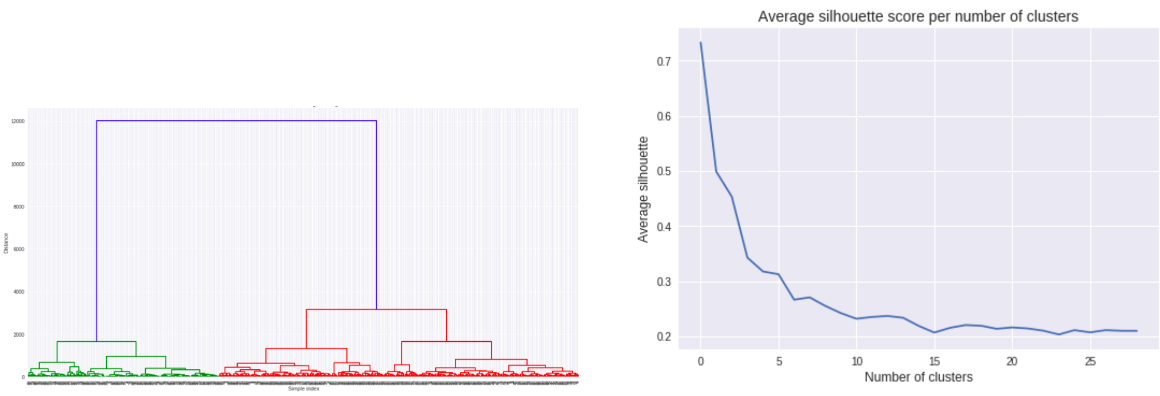
(b) 2018

Figure 3.5: Mean hourly power consumption at Civil building, in both years. A typical power consumption curve goes up in the morning, peaks at lunch time and starts going down again until the evening.

As expected, the power consumption begins to rise until around 11 AM, where it reaches a plateau of peak consumption. Around 5 PM, power begins to drop until it reaches idle levels. There is an idle regime drive throughout all days in which there's little to no activity in the campus. In particular, the mean hourly energy consumption has gone up by around 20% from 2017 to 2018 ($\approx$ +30kW average offset).

3.2 Clustering & Feature Selection

Considering one can distinguish two consumption regimes by performing exploratory data analysis, it becomes interesting to study data clusters - that is, separating data into groups and labeling them, using Machine Learning Classification algorithms. In fact, building an Hierarchical Clustering Dendrogram and retrieving kMeans [91] silhouette scores using Python's *scikit-learn* [92] library confirms this hypothesis, as seen on the graphs below (Figure 3.6).



(a) Hierarchical Clustering Dendogram

(b) Average Silhouette Scores by cluster number

Figure 3.6: Hierarchical Clustering Dendrogram and Average Silhouette Scores (kMeans) for hourly power consumption at Civil building, for both years. This confirms that there are, most certainly, two types of data points in this study.

From the kMeans study, one can now establish a proper plateau plot of both consumption regimes, differentiating them by cluster labels. The result plot is presented below, on Figure 3.7.

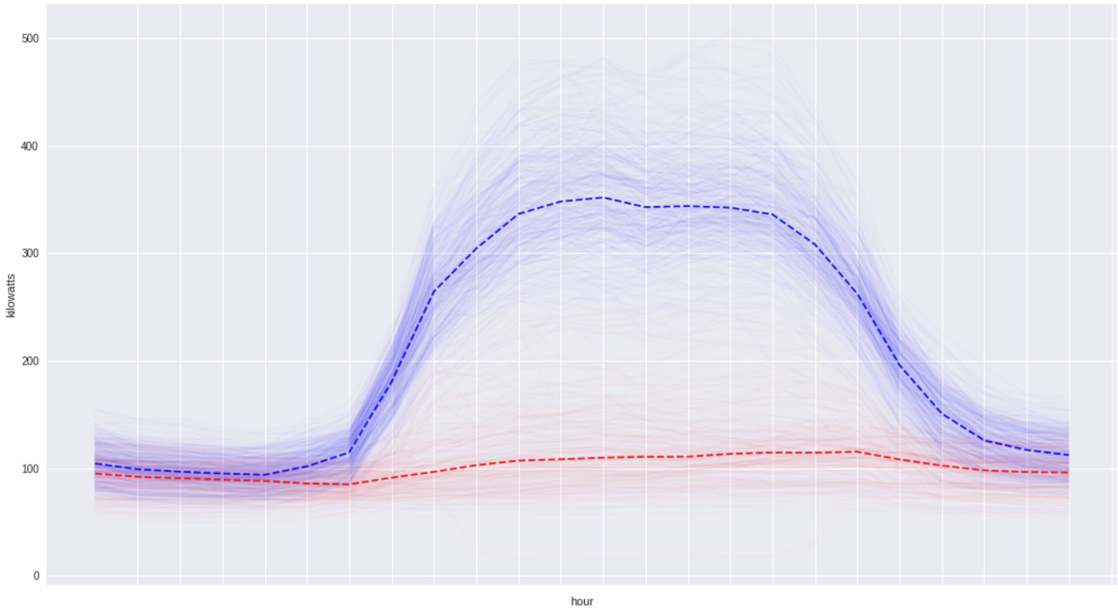


Figure 3.7: Plateau plot for hourly power consumption at Civil building in both years, during 24 hours, with the two highlighted consumption trends. Typical operation day (in blue) and lower operation day (in red).

When it comes to feature selection, one already knows that a parameter *WeekDay* (= 1 on week days and =0 on weekends) can be useful as it can distinguish between idle and business day consumption, so I shall add it to the feature dataset. In addition, two feature columns are loaded to this dataset to convert *timestamps* into discrete values of hour and month. Most of the input features, such as in other studies of the kind [93], are meteorological parameters, which are known to have an impact on energy consumption and positively influence models. The received data for the period 2016-2019 was retrieved from IST Meteorological Services [94]. Their data files include the following features:

1. Temperature (in degrees Celsius)
2. Relative Humidity (in percentage, %)
3. Wind Direction (in magnetic orientation, degrees)
4. Wind Gust (in m/s)
5. Wind Speed (in m/s)
6. Solar Radiation (in W/m^2)
7. Atmospheric Pressure (in mbar)
8. Precipitation (in mm/h)
9. RainDay (=1 in a day with reported precipitation, =0 otherwise)

To conclude, feature columns *Power - 1* and *Power - 2* are added, correspondent to the power consumption of the previous two entries, that is, of the previous two hours. I shall further explain the use of these features but basically, they describe the autoregressive behavior of an energy time series (i.e. the consumption at a certain period depends also on the previous level of consumption).

One can now head on to find the most relevant features to this study, so as to find the perfect balance between model complexity, computational time and accuracy. I shall try all three approaches to feature selection in order to boost the ultimate decision to extract certain features. Starting by the filter methods, I used the kBest routine, which uses an ANOVA classifier function, available in Python's *scikit-learn* library [92]. The highest scores are laid on this following table:

Feature	kBest Score
Power-1	44,2
Precipitation	6,6
Solar Radiation	1,68
Temperature	1,39

Table 3.2: kBest algorithm highest scores and features for k=4.

Then, moving towards the wrapper approach, RFE (recursive feature elimination) supported by a linear regression model [92] was used. Choosing to find the two main important features, the following affinity ranking was attained:

Feature	RFE Affinity
Wind Gust	1
Wind Speed	1
Hour	2
Power - 1	3

Table 3.3: RFE algorithm 2-fold highest ranking features.

Finally, the ensemble approach is attained using an Extra Trees Regressor [92] to scan for feature importance. Those results are listed on the next table.

Feature	Feature Importance
Power - 1	8,34E-01
Solar Radiation	1,02E-01
Hour	4,64E-02
WeekDay	6,74E-03

Table 3.4: Extra Trees Regressor feature importances.

In conclusion, all three approaches highlight the need for just some of the features and I chose to use the ones who gathered wide consensus. As such, let us include the following features, which scored as important to all three methods: **Solar Radiation**, **Power - 1**, **Hour** and **WeekDay**. These parameters are, therefore, expected to have an impact on IST's energy consumption. While nearly all of the correlations found were self-explanatory, the Solar Radiation feature was found to be a good proxy value of the power consumption, going up in the morning and, afterwards, down until the end of the day.

3.3 XGBoost Regression Model

In order to build a prediction regressor model, several tests were conducted as to assess which one performed best using these data sets and features. The elimination criteria was based on the common metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). The models verified under a test size of 15% for the concatenated data of both 2017 and 2018 were: Linear Regressor (LR), Random Forest Regressor (RF), uniformized Random Forest Regressor (uRF), Multi-layer Perceptron Regressor (MLP - Neural Network) [92] and XGBoost [82]. In all four buildings, XGBoost outperformed the other models, except for the North and South Tower, where the absence of some data points is expected to might have impacted the results.

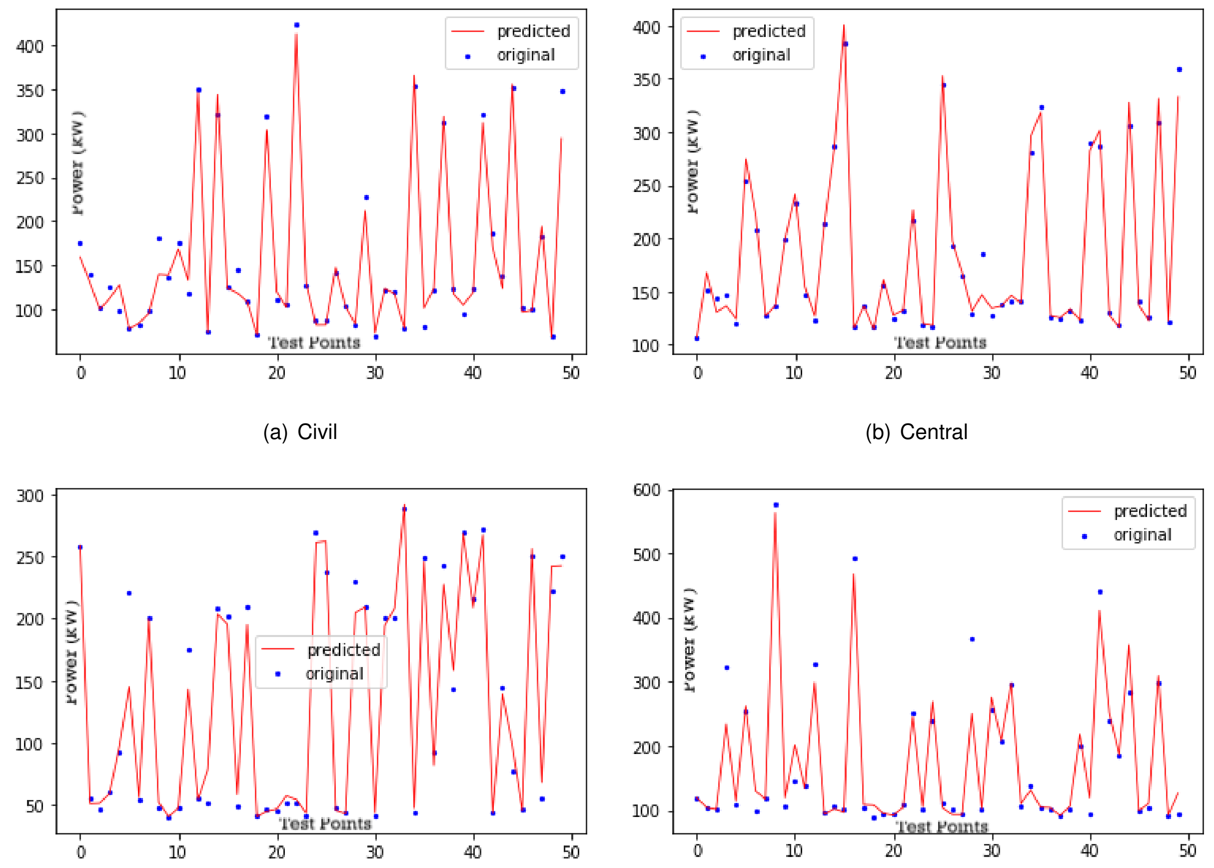
Upon using the previously trained model, the Power - 1 feature was replaced in its essence. Instead of representing the previous hour power consumption, and since the model should be used *a posteriori*, it was actively loaded as the previous year consumption at the same hour, day and month. One expects this difference to help us build the model in a quicker and non-recursive way. The error results for the XGB model which was used to determine the 2019 baseline consumption are listed on Table 3.5. The results were averaged from 10 runs of the model, for each building, at a test size of 12,5%, which was found to minimize errors.

Building	$\overline{17}$ (kW)	$\overline{18}$ (kW)	MAE (kW)	MSE (kW^2)	RMSE (kW)	MAE (rel.)	CV(RMSE)
Civil	164,3	183,3	8,23	194	13,5	0,047	0,078
Central	189,2	182,4	7,27	140	11,8	0,039	0,064
N. Tow.	102,5	114,4	8,88	294	17,1	0,082	0,158
S. Tow.	177,6	173,9	17,5	1393	37,2	0,100	0,212

Table 3.5: IST's buildings average hourly power consumption for both years, XGB basline model error parameters (MAE, MSE and RMSE), and error parameters relative to the 2-year average (MAE and CV(RMSE)). The behaviour at the Towers shows that additional features should be considered, to lower the CV(RMSE). The model was trained with a concatenated 2-year-long data set, which helps reducing meteorological induced variability.

Regarding the relative errors, it's possible to confirm that the industry adopted metric - CV(RMSE) - is well below the accepted standard of 25% for energy baseline models, which indicates a high degree of model precision. Both Tower metrics, due to their own glass envelope and high demand HVAC scheme, suggest that there is the need more features to study its energy baseline, such as non-linear temperature dependencies.

Here, some data plots of the first 50 test data points and the model's own prediction, for the four buildings, are exposed (Fig. 3.8):



(a) Civil (b) Central

(c) North Tower (d) South Tower

Figure 3.8: Data plots of the first 50 test data points vs. XGB regressor model prediction, for a given run, in all four buildings.

The model feature importance percentages are displayed on the table below, to show us how deeply did input variables influenced these predictions, as retrieved by the XGBoost library. Ten rounds of modeling were conducted as to better retrieve the final averaged feature importance for each building and a globally averaged feature importance (Table 3.6).

Building	Solar Rad.	Power - 1	Hour	WeekDay
Civil	0,018	0,884	0,092	0,006
Central	0,022	0,869	0,102	0,006
North Tower	0,017	0,88	0,094	0,009
South Tower	0,027	0,846	0,111	0,016
Average	2,1%	87%	10%	0,9%

Table 3.6: Feature importances for each building and global averaged feature importance, in percentage. Power - 1 and Hour are, hence, the most relevant features in this study.

Let us now take a look at the buildings feature correlation heat map (Figs. 3.9 and 3.10), built from all initial features, to explain and validate this feature group, as well as to confirm or undertake further adjustments to the model. Here, I used Python's *corr()* function together with *Seaborn* to build this map.

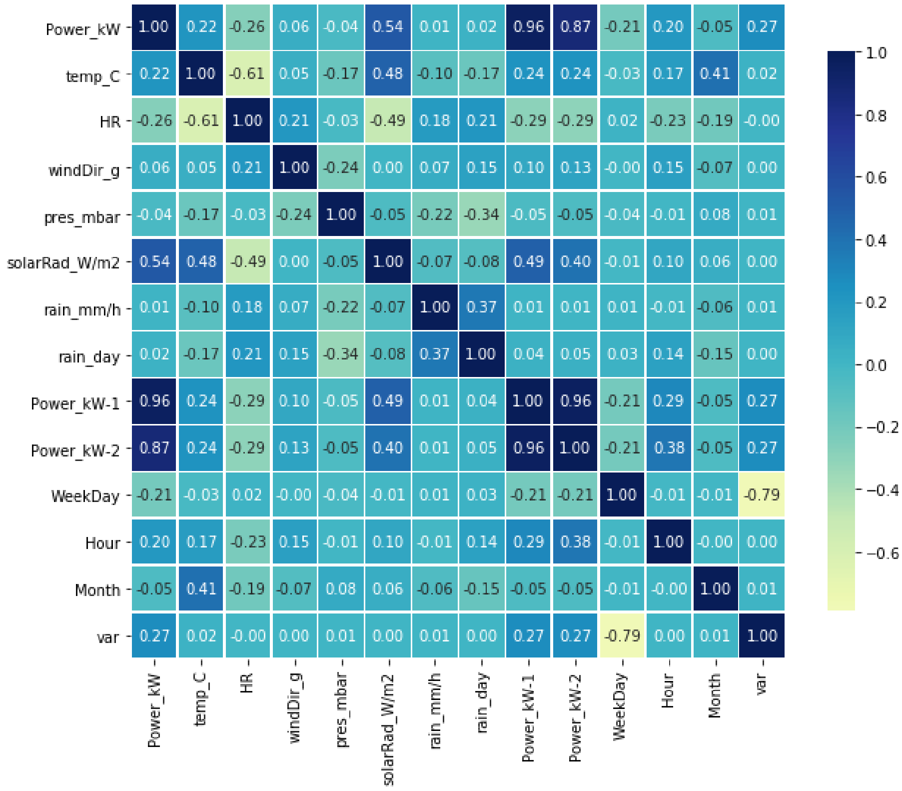


Figure 3.9: Civil building input feature correlation heat map. Here, ”WeekDay” stands for the number of the day relative to Sunday while ”var” stands for the used *WeekDay*

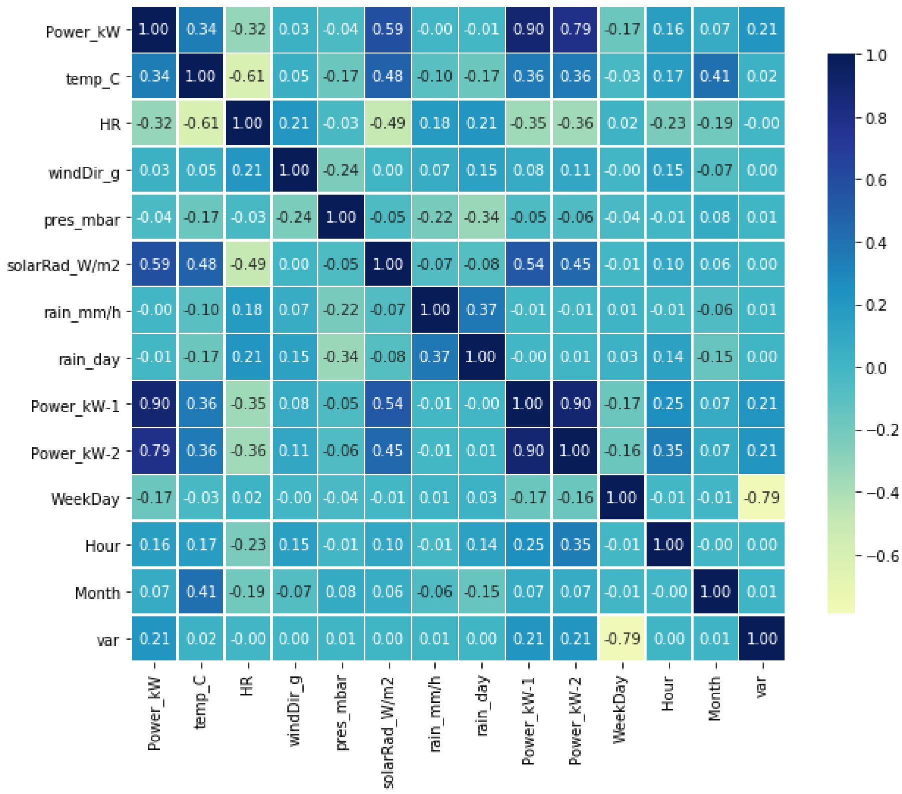


Figure 3.10: South Tower input feature correlation heat map. Here, ”WeekDay” stands for the number of the day relative to Sunday while ”var” stands for the used *WeekDay*

In fact, the model is using the top four correlated features to the "Power_kW" output parameter (apart from Power - 2, which would turn out redundant), which boosts model reliability. But when considering the Towers, these are getting higher error parameters. Actually, when looking at the South Tower heat map (Fig. 3.10), one can see that there's a higher correlation with temperature, which was confirmed by the North Tower's heat map. This can be explained by the building's architecture (glass window coated tower) and HVAC systems, which is a big contributor to the total load, on a logarithmic dependence with temperature. To improve the model, several training rounds were conducted admitting the temperature as a feature, to see if it would reduce the error. The result came out to be negative, with error factors increasing upon the consideration of temperature. It is then assumed that the lower accuracy of the tower models can be due to the lack of data points, because of faulty smart meters on both towers, a non-linear dependence with temperature and lack of completion of ECM implementation. Concerning Civil and Central, the ECM impacted buildings, it's safe to assume that the error parameters are good enough to effectively establish a baseline model, with CV(RMSE)s under 8%, below the standard 25% accepted in this sector. In addition, the 5-fold cross-validation r2 scores yielded a mean value of 98% for both buildings. As such, let us drop the savings study on both towers and focus on the two buildings which underwent ECMs on time for the considered reporting period - Civil and Central.

3.4 Savings M&V

To assess the measured savings, let us now integrate all parts of the data model. Loading the model with the IST's meteorological features from the year of implementation (2019) and plugging previous year's consumption data, at the same time and day, as *Power - 1*, the previously trained model retrieves its value predictions for the hourly consumption at each building. Then, let us compare it to the actual smart meter retrieved consumption data for the year of implementation, on a hourly basis. From converting the energy data from output current (A) to output power (kW), I was able to determine IST's buildings load factor (Table 3.7). Additionally, to calculate the CO_2 emissions reduction, let us refer to the value obtained from [31] [32] of 0,265 Kg/kWh, referent to Portugal.

Building	Load Factor
Civil	0,83
Central	0,91
North Tower	0,95
South Tower	0,95

Table 3.7: IST's buildings load factor.

It is known that ECMs began being implemented across Civil and Central buildings during the month of April, by *Campus Sustentável - IST* [95]. It focused on changing the lighting scheme to more efficient LEDs on the highest consuming buildings. The reporting period for the retrieved savings in this section was considered to be from 01/06/2019 to 31/12/2019. This way, the period leaves one month for ECM impact stabilization and then analyze a period consisting of three months of normal scholar activity, three months of holidays and one month of exam season, so one can better estimate an overall savings percentage, for an university campus. Below are the used formulas for percentage savings, actual savings and the error parameter, followed by a table in which these calculated values are displayed. The sums are performed on a hourly basis and the used average RMSE was the value present in Table 3.5, for each building. Representative plots of the model behaviour versus predicted data, hourly savings and CO_2 kilograms saved for a week in October 2019 at the Civil building are also displayed in Figure 3.11.

$$Savings = \sum_{1/6/19}^{31/12/2019} (Model - Actual) \tag{3.1}$$

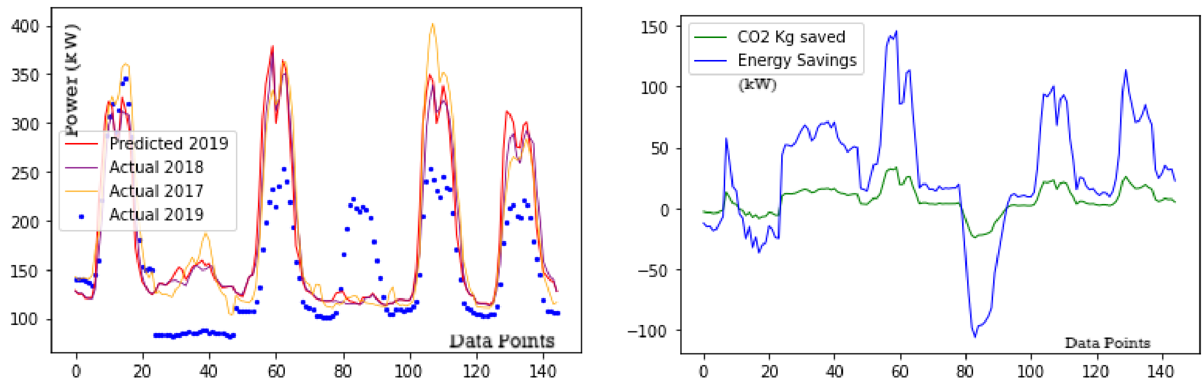
$$Savings(\%) = \frac{\sum_{1/6/19}^{31/12/2019} (Model - Actual)}{ModelAverage} \times 100 \tag{3.2}$$

$$Error = \frac{\overline{RMSE}}{ModelAverage} \times 100 \tag{3.3}$$

After ten averaged runs, the retrieved savings results were:

Building	Savings (%)	Savings (MWh)	CO_2 saved (ton)
Civil	16,9 $\pm$ 7,3	147,8	39,2
Central	20,6 $\pm$ 6,3	179,6	47,6

Table 3.8: Civil and Central buildings reported savings from 01/06/2019 to 31/12/2019, according to the XGB model.



(a) Baseline model

(b) Energy and CO_2 savings

Figure 3.11: Model hourly behaviour (a), hourly energy savings and hourly CO_2 savings (b) for a week in October 2019 at the Civil building. Here it is clear that calendar differences between years of study compensate each other sequentially.

Chapter 4

Blockchain

In order to deploy a ledger algorithm, an IBM Developer public blockchain prototype was used. The code and its significant updates are hosted on *git-hub* [96] and a thorough tutorial on how to program and use it is present in [97]. This application shall allow users to share information by posting on the network using a simple web interface (Fig. 4.1). Each building constitutes a node, which will run the savings script. Those calculations are validated by the baseline model's CV(RMSE) by the remaning nodes upon posting. A transaction comprises then a savings string and the model's RMSE and mean consumption value, these latter two used to calculate CV(RMSE). The used web framework was Flask [98]. In this chapter, I shall explain the code and adjustments that have been made, in order to deploy this M&V solution and properly assess EPC execution.

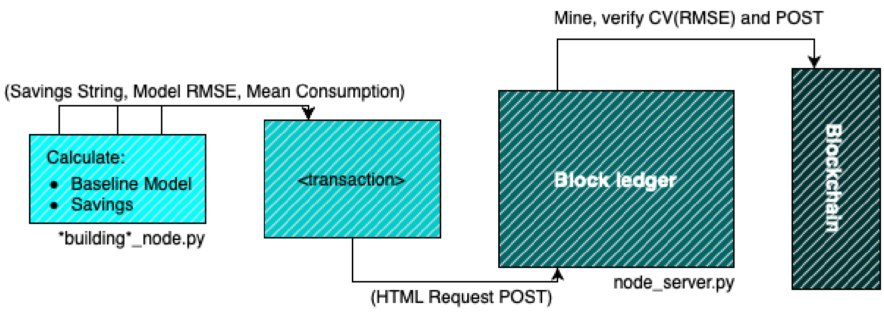


Figure 4.1: Application Posting Scheme

4.1 *Blockchain* and *Block* classes

The *node_server.py* application contains the blockchain implementation code for each node (Fig. 4.2). It begins by defining the *Blockchain* and *Block* classes, central to the development, where the routines needed to act upon the ledger are established.

The savings data shall be stored using JSON posts, saving the author, content and time stamp of the transaction. These transactions are then packed in blocks, which can contain one or many transactions at a time. These are the blocks that will further be added to the chain.

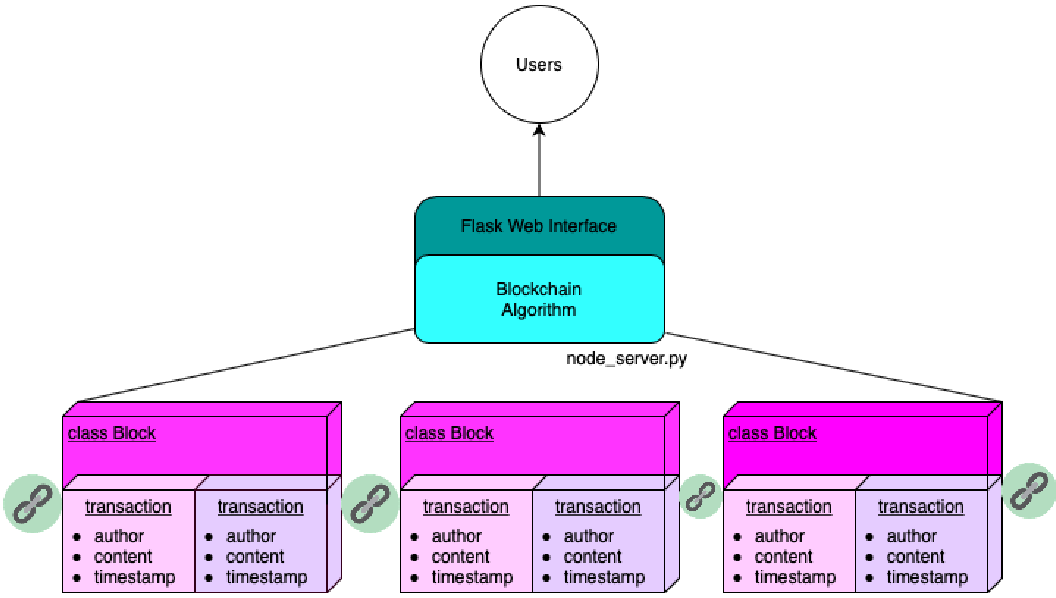


Figure 4.2: Blockchain network scheme.

The Block object is initialized by the function __init__ (Listing 4.1), where it takes an index (which serves as an unique identifier), the transactions array, a time stamp, the previous hash string and a default zero *nonce* as arguments.

In a network like this one, it's desirable to prevent transaction data tampering. To encode transaction information, a cryptographic hash function is used. A hash function is a function that takes input data of any size and retrieves fixed size data from it (the hash), which is used to identify the input [97]. These functions have to be easy to compute, deterministic (in the sense that the same data must retrieve the same hash) and uniformly random regarding changes in input. This way, it is virtually impossible to figure out the input data from the hash (the only way being to compute all possible input combinations) but, having the input and the hash, one can simply pass the input through the hash function to verify a provided hash. This is known as effort asymmetry (Fig. X).

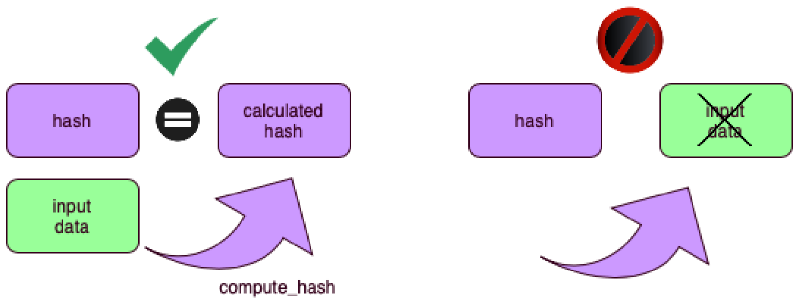


Figure 4.3: Effort Asymmetry.

In the context of the application, the function called compute_hash() is responsible for encrypting the data referent to the transaction string and, in this case, encoding the savings information text to be put on the block. The cryptographic hash function used to encode the strings was the Secure Hash Algorithm 2 (SHA-256, [99]) (256 bits), designed by the US National Security Agency (NSA), considered to be safe.

In order to avoid chain tampering, the blocks are chained together by having each of them store the previous block hash. This way, one makes sure that any changes in the previous blocks invalidates the whole ledger.

```
class Block:
  def __init__(self, index, transactions, timestamp, previous_hash, nonce=0):
    self.index = index
    self.transactions = transactions
    self.timestamp = timestamp
    self.previous_hash = previous_hash
    self.nonce = nonce

  def compute_hash(self):
    block_string = json.dumps(self.__dict__, sort_keys=True)
    return sha256(block_string.encode()).hexdigest()
```

Listing 4.1: *Block* class [96][97].

In the *Blockchain* class, let us start by the initializer function, the genesis (first) block creator and the last block retriever property (Listing 4.2). The initializer creates two empty arrays: *chain* and *unconfirmed_transactions*. The function that retrieves the last block returns the element -1 of the *chain* array. To initialize posting on the chain, the first block is created using function *create_genesis_block* that joins an empty *block* object to the current chain with index ”0”, so as to ensure coherence between blocks.

```
class Blockchain:
  # difficulty of the PoW algorithm
  difficulty = 4
  def __init__(self):
    self.unconfirmed_transactions = []
    self.chain = []

  def create_genesis_block(self):
    genesis_block = Block(0, [], 0, "0")
    genesis_block.hash = genesis_block.compute_hash()
    self.chain.append(genesis_block)

  @property
  def last_block(self):
    return self.chain[-1]
```

Listing 4.2: *Blockchain* class [96][97].

At this point, it is still possible to tamper with data by just changing the previous block and easily recomputing all the blocks that follow. One avoids this by exploiting effort asymmetry upon calculating the hash, making it difficult and random. In this case, let us add the constraint that the hash should start with n leading zeros. To prove that this computation was performed, a nonce (dummy) variable is stored on the blocks, that is incremented until the calculated hash satisfies this constraint (Listing 4.3).

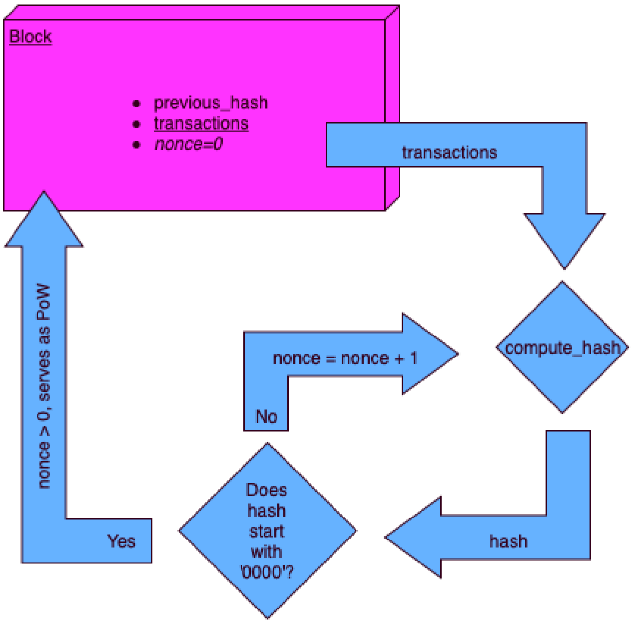


Figure 4.4: Proof-of-Work.

The *proof_of_work* function makes sure that the hash is retrieved following this constraint and its difficulty, saving proof of computation afterwards (Fig. 4.4). Correspondingly, there is a boolean function called *is_valid_proof* that checks if the computed hash for the block matches its input content. After that, *add_block* routine receives a block and its hash which then appends to the main chain, after it confirms previous hash coherence (perserved order of transactions) and computation proof (data tampering). In addition, the *add_new_transaction* function appends new transactions to the unconfirmed_transactions array.

```
def add_block(self, block, proof):
    previous_hash = self.last_block.hash
    if previous_hash != block.previous_hash:
        return False
    if not Blockchain.is_valid_proof(block, proof):
        return False
    block.hash = proof
    self.chain.append(block)
    return True

@staticmethod
def proof_of_work(block):
    block.nonce = 0
    computed_hash = block.compute_hash()
    while not computed_hash.startswith('0' * Blockchain.difficulty):
        block.nonce += 1
        computed_hash = block.compute_hash()
    return computed_hash

def add_new_transaction(self, transaction):
    self.unconfirmed_transactions.append(transaction)
```

```
    @classmethod
    def is_valid_proof(cls, block, block_hash):
        return (block_hash.startswith('0' * Blockchain.difficulty) and block_hash == block.
        compute_hash())
```

Listing 4.3: *Blockchain* class [96][97].

The process of appending the unconfirmed transactions to a block and computing proof-of-work is called mining. Once the hash constraints are met, a block is said to be mined and can be added to the ledger. In most cryptocurrencies, this mining computational activity is compensated by a share of cryptocurrencies [97].

The *mine* function appends all unconfirmed transactions to a block and adds it to the chain, after undergoing PoW and previous block hash coherence, resetting the unconfirmed transactions' array.

The *check_chain_validity* routine is used on the consensus mechanism for different chains (Listing 4.4). It checks if the computed hashes for the blocks in the chain match what they were supposed to, according to this particular cryptographic scheme.

On a longest-chain (more produced work) consensus approach like this one, one thereby validates each conflicting chain while checking what is the longest one (Fig. 4.5). This method is used on the further explained web-interface function *consensus* for that exact purpose.

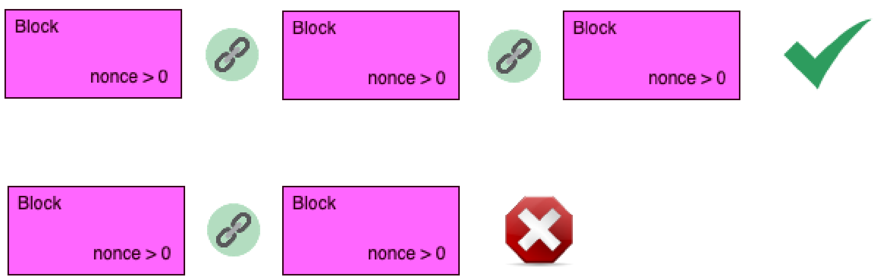


Figure 4.5: Conflicting chains consensus mechanism. The longest chain, which shows more evidence of work done via nonce, is the valid one.

[1]

```
    @classmethod
    def check_chain_validity(cls, chain):
        result = True
        previous_hash = "0"
        for block in chain:
            block_hash = block.hash
            delattr(block, "hash")
            if not cls.is_valid_proof(block, block.hash) or \
            previous_hash != block.previous_hash:
                result = False
                break
            block.hash, previous_hash = block_hash, block_hash
```

[1] It is important to understand the difference between block coherence, referent to previous hash matching, and conflicting chain consensus, in which a decision mechanism to chose between two different chains is designed, hosted by different nodes.

```
    return result

  def mine(self):
    if not self.unconfirmed_transactions:
      return False
    last_block = self.last_block
    new_block = Block(index=last_block.index + 1, transactions=self.unconfirmed_transactions,
    timestamp=time.time(), previous_hash=last_block.hash)
    proof = self.proof_of_work(new_block)
    self.add_block(new_block, proof)
    self.unconfirmed_transactions = []
    return True
```

Listing 4.4: *Blockchain* class [96][97].

4.2 *Flask* framework

Concerning the web interface, I shall explain how this solution handles the blockchain instructions on the client side, by using submitting HTML requests through *app routes*, using Flask [98] to create a REST API that invokes operations in the blockchain node [97].

To add new transactions to a block, the *new_transaction* routine saves the author and the content of the transaction (hence, the savings information), recording the current time stamp (Listing 4.5).

Aditionally, this adapted version of this blockchain requests the RMSE for a given used baseline model and the mean energy consumption value during the baseline period (in kW), in order to confirm that the model CV(RMSE) is satisfactory. This way, only transactions that guarantee a certain level of model accuracy can be added to a block. In this solution, I chose to demand a CV(RMSE) of under 15%. The submission of these parameters by the nodes is further discussed on Chapter 5.

```
@app.route('/new_transaction', methods=['POST'])
def new_transaction():
  tx_data = request.get_json()
  required_fields = ["author", "content","RMSE","Mean"]
  for field in required_fields:
    if not tx_data.get(field):
      return "Invalid transaction data", 404
  tx_data["timestamp"] = time.time()
  if (float(tx_data.get("RMSE")/tx_data.get("Mean"))) < 0.15:
    return blockchain.add_new_transaction(tx_data)
  return "Success", 201
```

Listing 4.5: *Flask* web framework [96][97].

If intentional manipulation or network latency occurs, the copy of the chain in some nodes can be compromised and differ from the other nodes. In that case, the network needs to agree upon some version of the chain to maintain integrity [97].

The *consensus* function clarifies just this latter point, by checking chain length when chains of different nodes appear to diverge. This way, it's agreed that the longest chain corresponds to largest amount of work (PoW) done and, hence, valid.

In order to submit the mining command, the *mine_unconfirmed_transactions* app route uses the *mine* function of the *Blockchain* class, making sure it features the longest chain before anouncing it to the network, which enforces the consensus criteria (Listing 4.6).

```
@app.route('/mine', methods=['GET'])
def mine_unconfirmed_transactions():
result = blockchain.mine()
if not result:
  return "No transactions to mine"
else:
  chain_length = len(blockchain.chain)
  consensus()
  if chain_length == len(blockchain.chain):
    announce_new_block(blockchain.last_block)
  return "Block #{} is mined.".format(blockchain.last_block.index)

def consensus():
  global blockchain
  longest_chain = None
  current_len = len(blockchain.chain)
  for node in peers:
    response = requests.get('{}chain'.format(node))
    length = response.json()['length']
    chain = response.json()['chain']
    if length > current_len and blockchain.check_chain_validity(chain):
      current_len = length
      longest_chain = chain
    if longest_chain:
      blockchain = longest_chain
      return True
    return False
```

Listing 4.6: *Flask* web framework [96][97].

After a certain block is mined by some node, it is added to other nodes' chains. That is done by defining the *verify_and_add_block* function, which loads the PoW to the *Blockchain add_block* function (Listing 4.7).

Finally, there is the need for any node to announce to the network that it has mined a block, so everyone can update their blockchain [97]. This way, the other nodes can simply verify proof-of-work and add the mined block to their respective chains. The *announce_new_block* method is then called after every block is mined by a certain node so that peers can update their chains.

```
@app.route('/add_block', methods=['POST'])
def verify_and_add_block():
block_data = request.get_json()
```

```
    block = Block(block_data["index"], block_data["transactions"], block_data["timestamp"],
    block_data["previous_hash"], block_data["nonce"])
    proof = block_data['hash']
    added = blockchain.add_block(block, proof)
    if not added:
      return "The block was discarded by the node", 400
    return "Block added to the chain", 201

    def announce_new_block(block):
    for peer in peers:
    url = "{}add_block".format(peer)
    headers = {'Content-Type': "application/json"}
    requests.post(url, data=json.dumps(block.__dict__, sort_keys=True), headers=headers)
```

Listing 4.7: *Flask* web framework [96][97].

Finally, to establish a network, one needs to be able to securely register new nodes and put them up to date regarding the valid chain. The *register_new_peers* and *register_with_existing_node* functions guarantee those exact methods, enabling a node to register new other trusted nodes (Listing 4.8). This last method will allow the remote node to add a new peer to its list of known peers and initializing the blockchain of the new node with that of the network node.

```
    @app.route('/register_node', methods=['POST'])
    def register_new_peers():
    node_address = request.get_json()["node_address"]
    if not node_address:
      return "Invalid data", 400
    peers.add(node_address)
    return get_chain()
    @app.route('/register_with', methods=['POST'])
    def register_with_existing_node():
    node_address = request.get_json()["node_address"]
    if not node_address:
      return "Invalid data", 400
    data = {"node_address": request.host_url}
    headers = {'Content-Type': "application/json"}
    response = requests.post(node_address + "/register_node", data=json.dumps(data), headers=
    headers)
    if response.status_code == 200:
      global blockchain
      global peers
      chain_dump = response.json()['chain']
      blockchain = create_chain_from_dump(chain_dump)
      peers.update(response.json()['peers'])
      return "Registration successful", 200
    else:
      # if something goes wrong, pass it on to the API response
      return response.content, response.status_code
```

Listing 4.8: *Flask* web framework [96][97].

Chapter 5

Implementation and Usage

In the last part of the development, let us integrate the model and blockchain components of this solution, producing automated savings posts that can be seen in an HTML page. Here, it is explained how to run the *localhost* ports that host these two blockchain nodes and how each buildings' savings are posted in each node, after calculation. The average script execution time, in seconds, is listed below on Table 5.1.

	central_node.py	*civil_node.py*
Execution time (s)	5,576	5,416

Table 5.1: Average execution time (s) for each script.

5.1 Blockchain Node Server and Application

To deploy this application, let us assign a *flask* application to the *node_server.py* script. After that, I run the server ports in which savings informations are posted, using the *run* command. Let us initialize two *localhost* ports, one for each building to post on (Listing 5.1).

```
$ export FLASK_APP=node_server.py
$ flask run --port 8000 & flask run --port 8001
```

Listing 5.1: Initializing the server application and ports [97].

After this, an instance of a blockchain node is running at *localhost* ports 8000 and 8001. On a different terminal session, let us now run the *curl* commands to register a new node (port 8001) with a proxy node (port 8000) . Symmetrically, because of application constraints, there is the need to do the same thing to register the first node (port 8000) with the latter one (port 8001) (Listing 5.2). This will make the node at port 8000 aware of the nodes at port 8001 and vice-versa [97]. New nodes will also sync their chain with the existing node so that they are able to participate in the mining process.

```
$ curl -X POST
http://127.0.0.1:8001/register_with
-H 'Content-Type: application/json'
-d '{"node_address": "http://127.0.0.1:8000"}'
```

```
$ curl -X POST
 http://127.0.0.1:8000/register_with
-H 'Content-Type: application/json'
-d '{"node_address": "http://127.0.0.1:8001"}'
```

Listing 5.2: Registering nodes with a proxy node (ports 8000 and 8001) [97].

Now, one just needs to run the *run_app.py* front-end application on a different terminal session. This will start the HTML interface at `http://localhost:5000`, in which one can visualize chain transactions. By default, this application syncs with *localhost* port 8000, but that parameter can be changed by updating the *CONNECTED_NODE_ADDRESS* field in the *views*.py file [97].

5.2 Savings Script

The *central_node.py* and *civil_node.py* scripts which retrieve the energy consumption savings information were built to: 1) treat the existing datasets and train a XGB model 2) load the reporting period features to the same model 3) retrieve the savings information post (string) and model CV(RMSE) and 4) automatically post weekly savings information on the blockchain, corresponding to the same week in 2019.

To make automatic posts, *python*'s *schedule* library is used, forcing code execution every Sunday. To do this, I define a function which calculates and posts savings and scheduling it to post in a given weekday (Listing 5.3). Due to its functioning, the user needs to set up the blockchain application on a week before begining to store and check savings information on the chain.

```
schedule.every().sunday.do(civilsavings)
while True:
  schedule.run_pending()
  time.sleep(1)
```

Listing 5.3: Schedule function.

To post the information on the nodes, HTML requests are used and the *httplib2* library, running the code present on Listing 5.4, which retrieves information about the current week's consumption in the reporting year (2019) and previous years (2017 and 2018). Notice the "RMSE" and "Mean" fields, which are submitted to the application to check if the CV(RMSE) is under 15%.

```
savingsstring = f"[Week {str(then5)} to {str(then6)}] Mean Hourly Savings: {meansavingsh} kWh, Model Predicted Consumption: {modelpred} kWh, Actual Consumption: {actual} kWh, Savings: {savedenergy} kWh ({perc}%), CO2 Saved: {co2} Kg, Actual Consumption (2018, same week): {actual18} kWh, Actual Consumption (2017, same week): {actual17} kWh"

#code in central_node.py
h = httplib2.Http(".cache")
res = requests.post('http://127.0.0.1:8000/new_transaction',json={"author":"Central","content":savingsstring,"RMSE":RMSE_XGB,"Mean":actualmean})
```

```
res2 = requests.get('http://127.0.0.1:8000/mine')

#code in civil_node.py
h2 = httplib2.Http(".cache")
res3 = requests.post('http://127.0.0.1:8001/new_transaction',json={"author":"Civil","content":
    savingsstring,"RMSE":RMSE_XGB,"Mean":actualmean})
res4 = requests.get('http://127.0.0.1:8001/mine')
```

Listing 5.4: HTML requests that post and mine new transactions to the nodes (ports 8000 (Central building node) and 8001 (Civil building node)) [97].

In addition, it was possible to verify that the value variation between trained model results in different script executions was under 1%.

At all times, one can retrieve a certain node chain by running the command "curl -X GET `http://localhost:*portnumber*/chain`"

5.3 HTML Page

The visual interface runs by default at `http://localhost:5000`. There, one can visualize the savings information content, the "author" building and the timestamp of the node post. There are buttons to request mining, returning to homepage and refreshing the current page. Additionally, there are two disabled features, which can be used to further improve this solution's UI/UX: a Reply button, on each post, that can be used to insert observations/comments to each week's savings information and input boxes, which may allow us to submit a post directly on the HTML page through the *CONNECTED_NODE_ADDRESS* (default port 8000). Notice that the input box posts are (currently) disabled since we lack the "RMSE" and "Mean" model accuracy fields which validate these transactions (Listing 4.5). Below, there's a screenshot of the blockchain HTML page during usage (Fig. 5.1).

Figure 5.1: Blockchain application HTML page with weekly savings information from both buildings.

Chapter 6

Conclusions

6.1 Achievements

During the execution of this thesis, I able to deploy a blockchain solution that accurately estimates and stores savings in a transparent manner. By having the validation mechanism associated to the accuracy of the forecasting model, I developed a new framework for EPCs to be audited in a clear, safe and trustworthy fashion.

M&V 2.0 technology integrations like this one allow stakeholders to better examine and estimate energy efficiency improvements locally and globally. These represent the most significant reduction in energy-sector CO_2 emissions globally and help us get on track with the three energy-related SDGs, saving energy and balancing demand with supply side improvements.

This solution was applied to a real world scenario of lighting ECM deployment at the two main buildings of the IST *Alameda* campus - Civil and Central - to estimate savings on a weekly basis, throughout the reporting period. Resorting to a trained XGB [82] baseline model, it is shown that the ECMs have reached a savings level of 16.9 $\pm$ 7.3% and 20.6 $\pm$ 6.3%, in the Civil and Central buildings, respectively, and a cumulative 86.8 metric tons reduction in energy related CO_2 emissions, over a reporting period of 7 months (Jun-Dec 2019). The models yielded an average CV(RMSE) of 7.8% and 6.4%, for the Civil and Central buildings, respectively.

To store savings information in safe blockchain nodes, an adapted version of the IBM Developer was used [97] blockchain algorithm which only validates savings posts after verifying that the baseline model's CV(RMSE) is under 15%. The information is displayed to the user via a Flask web interface, which interacts with each node server on the established network.

Future applications, adjustments and limitations are discussed in the sections below, so as to drive future developments. These results allow us to state that the endeavor objectives were met and that this was a promissing innovation in the field of EPC auditing technologies.

6.2 Limitations and Future Work

During the development of this first prototype, I've came across some limitations which prevent us from presenting an universally accepted solution. These barriers shall become the key issues to address when gradually improving this solution.

Regarding the baseline model, it is known that the deployment of XGBoost regressor models can be more computationally costlier than similar performing algorithms, like Random Forests, which can raise concerns of scalability. In this initial study, the focus was on delivering the most accurate model in detriment of the most efficient one, which should be taken in consideration when engaging in wider applications.

In what concerns the integration of savings calculations with the blockchain node servers, there's a need to establish an actual network and securing a public-private key cryptography scheme [97], instead of running the scripts on *localhosts* and having virtually every possible user posting/changing data on the building nodes' ledger. This way, the posts are yet added another level of security, after which some data protection work is recommended on the model side too, when handling key parameters, like the model CV(RMSE).

Focusing on energy data and information, stakeholders didn't find any documentation available regarding the deployment of this particular ECM that I could compare this results against. This model was proven to have behave satisfactorily but I exhort the community to further boost smart-meter deployment in public buildings and thoroughly documenting ECM implementations in the future, helping to maintain data consistency and effectiveness.

For the prosperity of technologies like the one here presented, I further recommend the standardization and de-bureaucratization of energy savings and EPC procedures, making savings knowledge easily tangible to the end-user. Further adjustments can be employed in smart cities applications of this solution which additionally comprise water savings and renewable energy production/storage. To foster this type of application, miners could also be rewarded an unit amount of energy savings, just like it is done in most cryptocurrencies.

Bibliography

[1] U. Nations. Un sustainble development goals, 2020. URL `https://sdgs.un.org/goals`.

[2] U. N. Economic and S. Council. *Special edition: progress towards the Sustainable Development Goals - Report of the Secretary-General*. 2019. URL `https://undocs.org/E/2019/68`.

[3] IEA. *Global Energy and CO2 Status Report 2019*. 2019. URL `https://www.iea.org/reports/global-energy-co2-status-report-2019`.

[4] IEA. *World Energy Model*. 2019. URL `https://www.iea.org/reports/global-energy-co2-status-report-2019`.

[5] IEA. *Energy Efficiency 2019*. 2019. URL `https://www.iea.org/reports/energy-efficiency-2019`.

[6] Ö. Gürcan, M. Agenis-Nevers, Y.-M. Batany, M. Elmtiri, F. Le Fevre, and S. Tucci-Piergiovanni. An industrial prototype of trusted energy performance contracts using blockchain technologies. In *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pages 1336–1343. IEEE, 2018.

[7] M. Frangou, M. Aryblia, S. Tournaki, and T. Tsoutsos. Potential of energy performance contracting for tertiary sector energy efficiency and sustainable energy projects in southern european countries. In *The Role of Exergy in Energy and the Environment*, pages 733–747. Springer, 2018.

[8] M. Frangou, M. Aryblia, S. Tournaki, and T. Tsoutsos. Renewable energy performance contracting in the tertiary sector standardization to overcome barriers in greece. *Renewable Energy*, 125: 829–839, 2018.

[9] J. Hwang, M.-i. Choi, T. Lee, S. Jeon, S. Kim, S. Park, and S. Park. Energy prosumer business model using blockchain system to ensure transparency and safety. *Energy Procedia*, 141:194–198, 2017.

[10] E3P. Energy performance contracting. . URL `https://e3p.jrc.ec.europa.eu/node/246`.

[11] IEA. *Energy Service Companies (ESCOs)*. 2018. URL `https://www.iea.org/reports/energy-service-companies-escos-2`.

[12] J. A. Shonder and E. Morofsky. Best practice guidelines for using energy performance contracts to improve government buildings. 2010.

[13] P. Fennell, P. Ruyssevelt, and A. Smith. The impact of measurement and verification option choice on financial returns for clients in energy performance contracts. International Symposium to Promote Innovation & Research in Energy . . . , 2017.

[14] NREL. *International Performance Measurement and Verification Protocol*. NREL, 2002. URL https://www.nrel.gov/docs/fy02osti/31505.pdf.

[15] IEA. *Energy efficiency and digitalisation*. 2019. URL https://www.iea.org/articles/energy-efficiency-and-digitalisation.

[16] J. N. Ferrer. Leveraging funding for energy efficiency in buildings in south east europe. ceps policy insights no 2019-05?28 march 2019, March 2019. URL http://aei.pitt.edu/97078/.

[17] M. Zhang, M. Wang, W. Jin, and C. Xia-Bauer. Managing energy efficiency of buildings in china: A survey of energy performance contracting (epc) in building sector. *Energy Policy*, 114:13–21, 2018.

[18] B. Richter, E. Mengelkamp, and C. Weinhardt. Maturity of blockchain technology in local electricity markets. In *2018 15th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2018.

[19] A. Khatoon, P. Verma, J. Southernwood, B. Massey, and P. Corcoran. Blockchain in energy efficiency: Potential applications and benefits. *Energies*, 12(17):3317, 2019.

[20] O. Gurcan, M. Agenis-Nevers, Y. Batany, M. Elmtiri, F. Le Fevre, and S. Tucci-Piergiovanni. An industrial prototype of trusted energy performance contracts using blockchain technologies. In *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pages 1336–1343, 2018.

[21] IEA, 2020. URL https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20consumption&indicator=Electricity%20consumption.

[22] IEA. *Electricity Information: Overview*. 2020. URL https://www.iea.org/reports/electricity-information-overview.

[23] IEA, 2020. URL https://www.iea.org/data-and-statistics?country=OECDTOT&fuel=Energy%20supply&indicator=Electricity%20generation%20by%20source.

[24] IEA, 2020. URL https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20supply&indicator=Electricity%20generation%20by%20source.

[25] IEA, 2020. URL https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20consumption&indicator=Electricity%20final%20consumption%20by%20sector.

[26] U. E. I. Administration. *Monthly Energy Review*. 2020. URL https://www.eia.gov/consumption/.

[27] IEA, 2020. URL https://www.iea.org/reports/unit-converter-and-glossary.

[28] E. Commission, 2020. URL https://ec.europa.eu/eurostat/statistics-explained/index.php/Energy_consumption_in_households#Energy_consumption_in_households_by_type_of_end-use.

[29] IEA. *Global Energy & CO2 Status Report 2019*. 2019. URL https://www.iea.org/reports/global-energy-co2-status-report-2019.

[30] IEA. *CO2 Emissions from Fuel Combustion*. 2020. URL https://www.iea.org/subscribe-to-data-services/co2-emissions-statistics.

[31] *Boletim Electricidade Renovável*. URL https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjOxeG3-oHsAhWmxIUKHRMwD3QQFjAAegQIAxAB&url=https%3A%2F%2Fwww.apren.pt%2Fcontents%2Fpublicationsreportcarditems%2F02-boletim-energias-renovaveis-vf.pdf&usg=AOvVaw0ESx01NjvreF6Osgf3dAIy.

[32] E. Comercial. Origem da energia, 2020. URL https://www.edp.pt/origem-energia/.

[33] S. Noticias. Edp antecipa o encerramento de centrais a carvão. 2020. URL https://sicnoticias.pt/economia/2020-07-14-EDP-antecipa-o-encerramento-de-centrais-a-carvao.

[34] J. Kneifel. Life-cycle carbon and cost analysis of energy efficiency measures in new commercial buildings. *Energy and Buildings*, 42(3):333–340, 2010.

[35] D. Popescu, S. Bienert, C. Schützenhofer, and R. Boazu. Impact of energy efficiency measures on the economic value of buildings. *Applied Energy*, 89(1):454–463, 2012.

[36] G. de Portugal. Edificios + sustentaveis, 2020. URL https://www.fundoambiental.pt/avisos-2020/mitigacao-das-alteracoes-climaticas/programa-de-apoio-a-edificios-mais-sustentaveis.aspx.

[37] EVO. What is m&v - efficiency valuation organization (evo), 2020. URL https://evo-world.org/en/m-v/what-is-m-v.

[38] EVO. Purpose of m&v - efficiency valuation organization (evo), 2020. URL https://evo-world.org/en/m-v/purpose-of-m-v.

[39] EVO. Principles of m&v - efficiency valuation organization (evo), 2020. URL https://evo-world.org/en/m-v/principles-of-m-v.

[40] D. Tanguay. Ipmvp - efficiency valuation organization (evo), 2020. URL https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.

[41] EnergyWatch, 2020. URL https://energywatch-inc.com/ipmvp-options/.

[42] EVO. *IPMVP GENERALLY ACCEPTED M&V PRINCIPLES*. 2020. URL https://evo-world.org/images/corporate_documents/IPMVP-Generally-Accepted-Principles_Final_26OCT2018.pdf.

[43] E. Franconi, M. Gee, M. Goldberg, J. Granderson, T. Guiterman, M. Li, and B. A. Smith. The status and promise of advanced m&v: An overview of “m&v 2.0” methods, tools, and applications. 4 2017. doi: 10.2172/1350974.

[44] J. Granderson, S. Touzani, C. Custodio, M. D. Sohn, D. Jump, and S. Fernandes. Accuracy of automated measurement and verification (m&v) techniques for energy savings in commercial buildings. *Applied Energy*, 173:296–308, 2016.

[45] C. V. Gallagher, K. Leahy, P. O’Donovan, K. Bruton, and D. T. O’Sullivan. Development and application of a machine learning supported methodology for measurement and verification (m&v) 2.0. *Energy and Buildings*, 167:8–22, 2018.

[46] X. Xia and J. Zhang. Mathematical description for the measurement and verification of energy efficiency improvement. *Applied Energy*, 111:247–256, 2013.

[47] J. Brittain. Think of energy as a service.

[48] E3P. Energy service companies, . URL https://e3p.jrc.ec.europa.eu/communities/energy-service-companies.

[49] P. Bertoldi, S. Rezessy, and E. Vine. Energy service companies in european countries: Current status and a strategy to foster their development. *Energy Policy*, 34(14):1818–1832, 2006.

[50] M. J. Hannon, T. J. Foxon, and W. F. Gale. The co-evolutionary relationship between energy service companies and the uk energy system: Implications for a low-carbon transition. *Energy Policy*, 61:1031–1045, 2013.

[51] H.-L. Kangas, D. Lazarevic, and P. Kivimaa. Technical skills, disinterest and non-functional regulation: Barriers to building energy efficiency in finland viewed by energy service companies. *Energy Policy*, 114:63–76, 2018.

[52] I. M. F. F. C. Soares and F. P. M. Barbosa. Energy service companies in portugal and throughout the world. In *2012 ELEKTRO*, pages 251–255. IEEE, 2012.

[53] J. Findley. Growing energy efficiency investment: Global analysis of energy service companies. Master’s thesis, 2019.

[54] C. Capelo, J. F. Dias, and R. Pereira. A system dynamics approach to analyse the impact of energy efficiency policy on esco ventures in european union countries: a case study of portugal. *Energy Efficiency*, 11(4):893–925, 2018.

[55] P. Xu, E. H.-W. Chan, and Q. K. Qian. Success factors of energy performance contracting (epc) for sustainable building energy efficiency retrofit (beer) of hotel buildings in china. *Energy policy*, 39 (11):7389–7398, 2011.

[56] U. D. of Energy and C. Change. *Guide to Energy Performance Contracting Best Practices*. 2015.

[57] P. Lee, P. Lam, and W. Lee. Risks in energy performance contracting (epc) projects. *Energy and Buildings*, 92:116–127, 2015.

[58] F. W.-H. Yik and W.-L. Lee. Partnership in building energy performance contracting. *Building Research & Information*, 32(3):235–243, 2004.

[59] W. Zhang and H. Yuan. A bibliometric analysis of energy performance contracting research from 2008 to 2018. *Sustainability*, 11(13):3548, 2019.

[60] J. Núñez Ferrer. Leveraging funding for energy efficiency in buildings in south east europe. ceps policy insights no 2019-05/28 march 2019. 2019.

[61] E. Union. Investment guideline, 2018. URL `https://ec.europa.eu/energy/sites/ener/files/documents/002_1.2_nicholas-stancioff_seif-athens-roundtable_25-10-18.pdf`.

[62] C. EU. *Periodic Reporting for period 1 - Accelerate SUNShINE (Save your bUildiNg by SavINg Energy. Begin to move more quickly)*. 2017. URL `https://cordis.europa.eu/project/id/754080/reporting`.

[63] EU. *Ener-In-Town Project Summary*. URL `https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/enerintown_project_summary_en.pdf`.

[64] PNAEE, 2020. URL `https://www.pnaee.pt`.

[65] EcoAP, 2020. URL `https://ecoap.pnaee.pt/casos-de-sucesso/`.

[66] F. Seefeldt. Energy performance contracting–success in austria and germany–dead end for europe. *Proceedings of the European Council for Energy Efficient Economy 2003 Summer Study*, 2003.

[67] K. Vaughn. Empire state building retrofit surpasses energy savings expectations, 2020. URL `https://rmi.org/blog_empire_state_retrofit_surpasses_energy_savings_expectations/`.

[68] C. Foundation. Four reasons why the empire state building retrofit model works, 2020. URL `https://stories.clintonfoundation.org/four-reasons-why-the-empire-state-building-retrofit-model-works-1e4d9ebdbe9`.

[69] U. D. of Energy. Espc success stories, 2020. URL `https://www.energy.gov/sites/prod/files/2013/10/f3/espc_ss09_ssa.pdf`.

[70] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37, Mar. 1996. doi: 10.1609/aimag.v17i3.1230. URL `https://www.aaai.org/ojs/index.php/aimagazine/article/view/1230`.

[71] V. Kumar and S. Minz. Feature selection: a literature review. *SmartCR*, 4(3):211–229, 2014.

[72] S. Kaushik. Feature selection methods - machine learning, 2020. URL www.shorturl.at/oxFU3.

[73] Wikipedia. Feature selection. URL https://en.wikipedia.org/wiki/Feature_selection.

[74] Predictive modelling, 2020. URL https://en.wikipedia.org/wiki/Predictive_modelling.

[75] J. Brownlee. Difference between classification and regression in machine learning, 2020. URL https://machinelearningmastery.com/classification-versus-regression-in-machine-learning/.

[76] Y. University. Linear regression, 2020. URL http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm.

[77] A. Liaw, M. Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

[78] Bootstrap aggregating, 2020. URL https://en.wikipedia.org/wiki/Bootstrap_aggregating.

[79] Decision tree learning, 2020. URL https://en.wikipedia.org/wiki/Decision_tree_learning.

[80] Random forest, 2020. URL https://corporatefinanceinstitute.com/resources/knowledge/other/random-forest/.

[81] Xgboost - a deep dive into boosting, 2020. URL https://dzone.com/articles/xgboost-a-deep-dive-into-boosting.

[82] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322. doi: 10.1145/2939672.2939785. URL https://doi.org/10.1145/2939672.2939785.

[83] J. Fan, X. Wang, L. Wu, H. Zhou, F. Zhang, X. Yu, X. Lu, and Y. Xiang. Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in china. *Energy Conversion and Management*, 164:102–111, 2018.

[84] M. Global. *Blockchain Moving Beyond Bitcoin*. 2020. URL https://www.msg-global.com/blog-item/blockchain-moving-beyond-bitcoin.

[85] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang. An overview of blockchain technology: Architecture, consensus, and future trends. In *2017 IEEE International Congress on Big Data (BigData Congress)*, pages 557–564, 2017.

[86] OECD. Oecd blockchain primer. 2018. URL https://www.oecd.org/finance/OECD-Blockchain-Primer.pdf.

[87] Y. Zhao, K. Peng, B. Xu, Y. Liu, W. Xiong, and Y. Han. Applied engineering programs of energy blockchain in us. *Energy Procedia*, 158:2787–2793, 2019.

[88] X. Tai, H. Sun, and Q. Guo. Electricity transactions and congestion management based on blockchain in energy internet. *Power System Technology*, 40(12):3630–3638, 2016.

[89] F. Imbault, M. Swiatek, R. De Beaufort, and R. Plana. The green blockchain: Managing decentralized energy production and consumption. In *2017 IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pages 1–5. IEEE, 2017.

[90] Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, and Y. Zhang. Consortium blockchain for secure energy trading in industrial internet of things. *IEEE transactions on industrial informatics*, 14(8):3690–3700, 2017.

[91] K. Krishna and M. N. Murty. Genetic k-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439, 1999.

[92] scikit-learn, 2020. URL `https://scikit-learn.org/stable/`.

[93] M. G. Fikru and L. Gautier. The impact of weather variation on energy consumption in residential houses. *Applied Energy*, 144:19–30, 2015.

[94] Meteo tecnico, 2020. URL `https://meteo.tecnico.ulisboa.pt`.

[95] Campus sustentavel - ist, 2020. URL `http://sustentavel.unidades.tecnico.ulisboa.pt`.

[96] Ibm developer blockchain algorithm - github, 2020. URL `https://github.com/satwikkansal/python_blockchain_app/tree/ibm_blockchain_post`.

[97] Ibm developer blockchain algorithm, 2020. URL `https://developer.ibm.com/technologies/blockchain/tutorials/develop-a-blockchain-application-from-scratch-in-python/#conclusion`.

[98] Flask web interface, 2020. URL `https://flask.palletsprojects.com/en/1.1.x/`.

[99] Sha-2, 2020. URL `https://en.wikipedia.org/wiki/SHA-2`.