Turnitin submission ID: 110650809

University of Exeter coursework header sheet

ECMM164 MSc Dissertation (A, TRM2+3 2019/0)

114141

1039874

680054221

Coursework: Dissertation and Interview

Submission Deadline: Fri 13th Sep 2019 12:00

Personal tutor: Ms Aileen Macgregor

Marker name:	N/A
Word count:	13468

First marker's comments	Indicative mark
Second marker's comments	Second mark
Moderator's comments	Agreed mark

MSc Dissertation

ECMM164

Dissertation Title:

Smart Electric Consumption Solution

Student Name: Umut Makascioglu
Programme: Postgraduate Taught
Student number: 680054221
Candidate Number: 031198
Student E-mail: um214@exeter.ac.uk
MSc Programme Name:
MSc Engineering Business Management
College of Engineering, Mathematics and
Physical Sciences
Supervisor: Dr. Barış Yüce

Declaration

I declare that the following work, regarding research and the results are products of my own, any and all outside information provided and presented in this dissertation comply to the ethical criteria of academic honesty, necessary and appropriate references were made where they were due.

Acknowledgements

I would like to thank my family, friends and all others who have supported me during the progress of this project, and especially my supervisor, Dr. Barış Yüce, for his endless support and guidance throughout the research and completion of this dissertation.

Table of Contents

Abstract

As the world progresses towards another Industrial Revolution and adopts it, it is crucial to recognize new ways to implement solutions into both existing and new patterns that govern the ways of life we have. One of the key factors in this regard are the houses that we spend so much time in and contain most of the devices that create our individual electric usage. In order to realize possible solutions that can optimize and lower energy costs in our homes, three experiments were conducted. First, an existing data set was analysed to depict existing user patterns regarding usage of electricity across a given time period and analysed data was extracted with regards to the specific time frame of the benchmark model. Secondly, a solution data set consisting of IDs from the previous data set was created as a solution. The IDs were chosen based on their relative error behaviour with regards to the experimental room that was set up by implementing Industry 4.0 technologies. Finally, through the utilization of an artificial neural network and its training, estimation results were obtained, which proved to be a better method for creating estimation based on multiple variables and for the creation of a framework for electric consumption.

Keywords: Industry 4.0, Internet of Things, Smart Homes, Energy Optimization, Artificial Neural Networks

1) Introduction

As the world's population grow so does the needs and costs to cover those needs. (BBC Bitesize, 2019) These needs cover both our domestic and industrial needs. This means that the desire for resources will increase as well as the energy costs that is needed to cover these needs. With Industry 4.0 manufacturing goods can be generated at a faster rate and with optimized costs thanks to maximized rate of obtaining data. (Business News Daily, 2019)

Moreover, it is also important to tell that our way of life, consumerism and use of natural resources for energy is damaging the environment and the world. In addition, due to our world having limited resources, the current consumer culture has a limited life span as well. Because of this, more sustainable ways to must be adopted for the continuation of our current lifestyles. (Balatsky, Balatsky and Borysov, 2015).

It is important to understand the necessities of our time to make progress for the future. If necessary actions are taken such as implementing Industry 4.0 technologies, reducing the effect on environment and use of resources and optimizing the energy costs can become a possibility. One possible way to implement such technologies is through the use of Smart Homes in our households, since our homes are the source of almost a third of our overall energy consumption. (Ec.europa.eu, 2019) These possibilities can be explored by using various simulation programs to depict scenarios and to provide case study results.

Before examining the problem thoroughly and providing a solution, first a Literature Review is provided to give necessary and vital information behind the main ideas and concepts behind the research's topics. Through this research the following questions will be answered;

- What is Industry 4.0 and its purpose and benefits?
- What are the relevant concepts in Industry 4.0?
- Industry 4.0 and how it is related to Internet of Things?
- What are Smart Homes and how can it be applied?
- How can Energy Cost Estimation be achieved for Smart Homes?
- Where else can these concepts be applied?

This is followed by the Methodology section, where the philosophy and the approaches that were adopted for this research are outlined as well as the data that were used and how it was collected. This section is ended by describing the objectives of the research.

Furthermore, the various experiments that were conducted for the research objectives are explained with regards to the inputs that were used and how these inputs were analysed for each experiment. Since each experiment is different, adequate information regarding their respective platforms and the utilization is explained as well.

The results obtained from the experiments are discussed and compared in the final section, and the best option for electricity estimation is shown as the final solution that should be used for the possible cost estimation framework.

Following the discussion of the results obtained from the experiments, the paper is concluded by making final remarks and providing ideas and possibilities for future research and project ideas by drawing from the material that was explained in the Literature Review and the results observed from conducted experiments.

2) Literature Review

In this chapter the concepts that form the back bone of the research and its objectives will be expounded upon and analysed by providing vital information and examples from existing academic papers and projects. For this purpose, the review structure follows a general to specific approach regarding the topics. For every subtopic, a literature review sub section is included that gives an overview about the subtopic and its importance. In a similar sense, the research for articles and other academic sources will follow the same structure, starting from the most general topic and focusing on converging topics as the research finalizes.

2.1) Industry 4.0

Industry 4.0 is considered to be the innovative way of combining production lines of old with the new age technologies such as Internet of Things and machine learning by making them “smart”. Thanks to the interconnectivity between machines and learned algorithms in them maximum gain can be achieved for production and supply chains and open the way for cost reduction in different ways. This also suggests that a faster rate of production as well as lowered costs can be achieved. Furthermore, since the accumulation of information within the system is vastly improved quality and efficiency will also be improved as a result. (Vaidya, Ambad and Bhosle, 2018) The term was first used by the initiative put forward during the Hannover Fair in 2011, in Germany, supported by the German government of the time, as the next tactical development and deployment of technologies in many fields, but the common denominator of all the concepts is that improving and progressing of the manufacturing ways that were and are still in use. (Hofmann and Rüsch, 2017)

2.1.1) Predecessors of the Revolution

In order to come to an understanding of Industry 4.0, first it should be understood how it came to be. In order to achieve this, a brief historical background regarding the previous revolutions will be provided as well as how each of them advanced technology and manufacturing forward.

The first revolution introduced mechanization and usage of burning coal to use water and steam power to a world where manufacturing relied on craftsmanship and muscle power. This increased the volume of produced goods tremendously with devices such as the Spinning Jenny and Watt’s steam engine. For industries such as textile and iron these new innovations were a breakthrough since the output was maximized and the labour and time cost were minimized. (Newworldencyclopedia.org, 2019) For instance, the innovation and the usage of the cotton gin saw the output of the raw material double with every passing 10 years and in half a century enabled the yield from Americas to cover the ¾ supply of the whole world. (Eliwhitney.org, 2019)

The second revolution followed almost a century later when other resources such as gas, oil and electricity were started to be used for manufacturing purposes, which throughout the century led to the rise of massive industrial powerhouses such as the German Empire and the USA. (Newworldencyclopedia.org, 2019) In addition to the diversification of resources, this period of time saw many technological inventions that defined 19th century and early 20th century, such as the telegraphy, railroads and combustion engines. (Niiler, 2019)

It can also be seen as a direct continuation of the first revolution, since the ideas and methods that shaped the first one spread across the world and the developing nations at the time used these ideas to transform their cities and societies from agricultural and basic ones to industrial, mass producing powerhouses. The rapid progress and implementation of these technologies in a short time period can be perhaps best depicted by the advancement of railroads in USA shown in Figure 2.1.

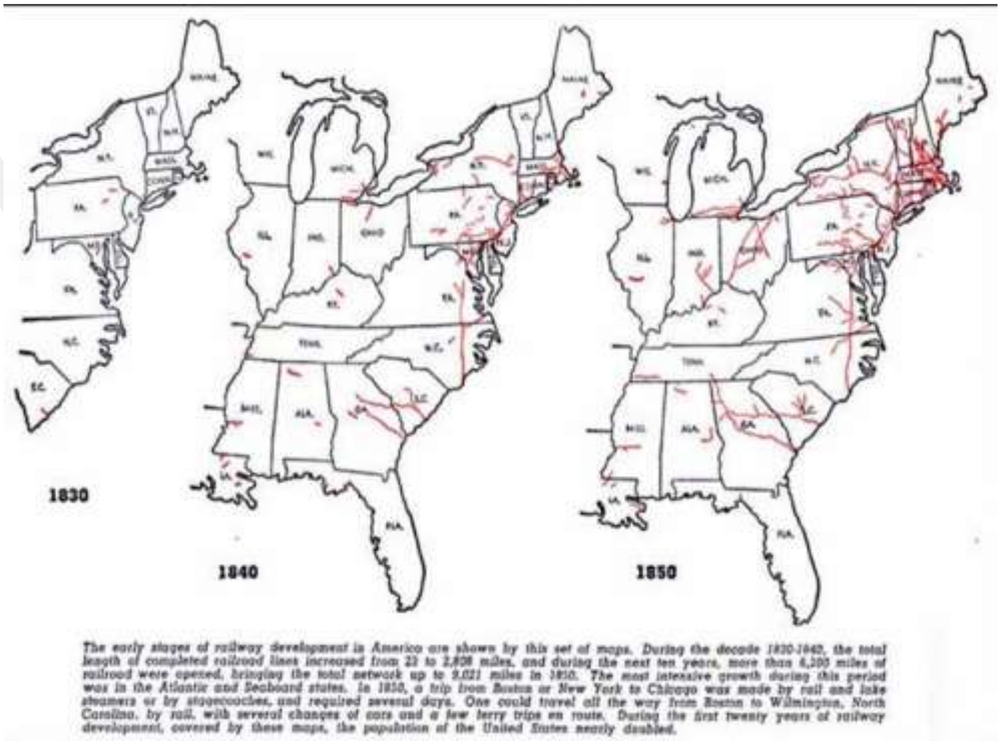

Figure 2.1: The Development of Early Railroads on the East Coast of USA. (Olson-Raymer, n.d.)

This era of societal change had reached its peak with the introduction of the moving assembly line by Henry Ford and the start of World War I where the effects of the products of the revolution were displayed in every aspect. (Niiler, 2019)

Figure 2.2: The Ford Moving Assembly Line (Eisenstein, 2013)

The third revolution does not have an exact start date. However, since the focus of the “digital revolution” is the digitalization of systems, it can be said that the invention date of the first practical transistor in 1948 might be considered as a start date, since in the following decades integrated circuits and microprocessors were made possible thanks to transistors. (Ross, 1998)

However, it could be considered more accurate if the start date is somewhere in 1950’s as the world was introduced to one of the first electronic calculators, thus opening the way for inclusion of computers in both our daily lives and the manufacturing world in addition to the communication technologies that are now an inseparable part of almost everyone. (Sanayidegelecek.com, n.d.) As it happened with the previous revolutions, the livelihoods of people, especially in the developed world where the fruits of the former revolutions were bore. The rise of consumer electronics throughout the Cold War era, with vinyl LP’s turning into CD’s in 1985and the rise and expansion of personal computers could be considered for some of the biggest signifiers for the effects of this revolution. (Museum of Obsolete Media, n.d.)

Similar to the second industrial revolution, the fourth industrial revolution may be viewed as a direct continuation of the third industrial revolution, since the origins of technological ideas that are being put forward by the current revolution, and are going to be implemented across the world are only available thanks to the implementation of the previous revolution.

Although the fourth revolution is on its way, right now, the world is still using and clinging to the 3rd era and the technological rewards and benefits. However, this does not mean that change will not happen in some distant future, on the contrary, the integration of the next industrial revolution may happen very soon. While it is true that the industrial revolutions have centuries of history behind them, the progression and the gap between each have been decreasing and the concepts that are covered and the new ones that are generated increase exponentially, as it can be observed in Figure 2.3.

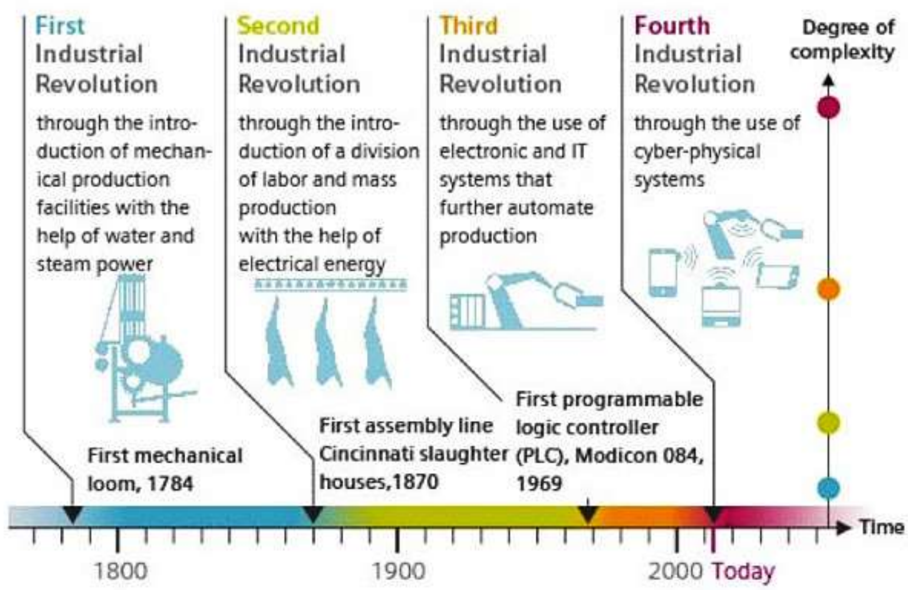

Figure 2.3: Progression of the Industrial Revolutions (Manufacturing-operations-management.com, n.d.)

2.1.2) Concepts of IR 4.0

Now that the humble beginnings of the current revolution have been established, the focus should be turned to the state of the art of the technologies and the impact and possibilities it can have in different areas.

As briefly mentioned at the start of this section, Industrial Revolution 4.0, also referred to as IR 4.0, covers a vast range of complex concepts. Although, the main drive behind the progress is, as it was in the previous revolutions, is of course improving manufacturing methods and maximizing yield, in addition to reducing costs. To be able to understand how the implementation of the next revolution in production and other areas can prove to be beneficial, first the concepts that form IR 4.0 used must be clarified.

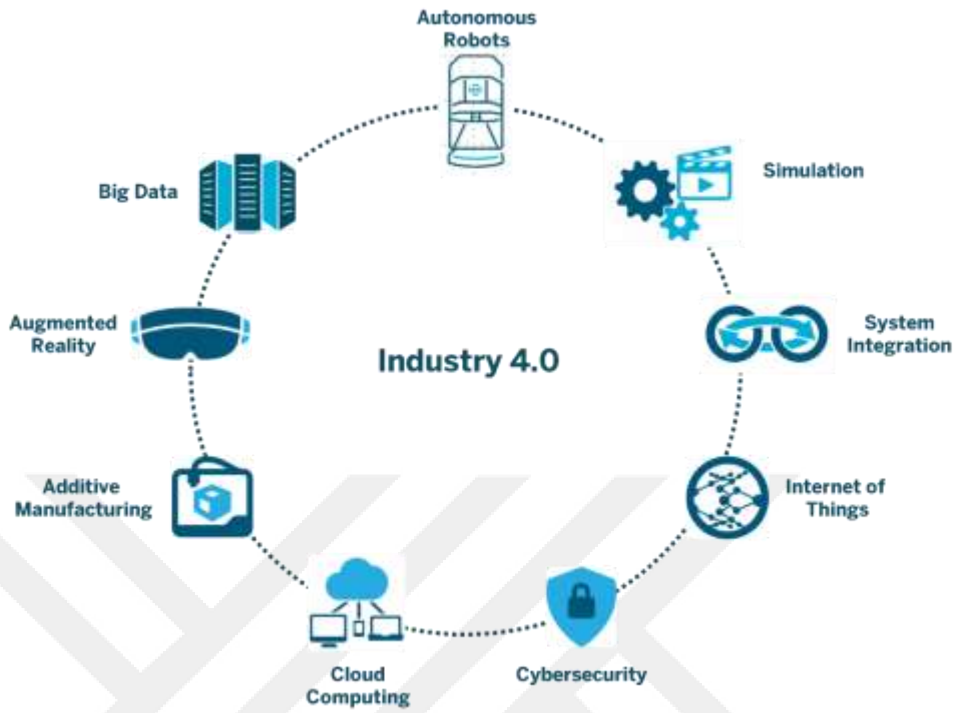

Figure 2.4: Concepts of Industry 4.0 (12CAD.com, 2018)

As it can be seen from Figure 2.4 above, there are wide range of areas of interest that constitute Industry 4.0. In addition to detailed explanations of the concepts, the connective nature of these concepts will be also described. Furthermore, implementation examples from the industry and possible scenarios will be offered to visualise how the concepts may be put to use in real life situations and possible benefits.

Additive Manufacturing

Additive Manufacturing is the usage of 3-D printers and compatible 3-D material to create a vast range of products. As Horst, Duvoisin, Vieira suggest (2018), the utilization of the technology enables mass customization and waste reduction for both individuals and manufacturers. Besides these benefits, by utilizing Additive Manufacturing, various complex designs can be achieved with ease and optimization of these designs can be created by improving upon the weaknesses of old models. Reiher and Koch state (n.d.) that, by improving the designs with Additive Manufacturing can remove unnecessary parts from the product, therefore obtaining waste reduction. Reiher and Koch also add (n.d.) that since waste material is reduced, overall weight of the design is reduced as well. This chain of optimization can be deployed upon other parameters of the designed products and even offer advantages such as reduction of inventory costs. A possible scenario could be, a factory that has successfully implanted, Internet of Things and Additive Manufacturing and where time management is at utmost importance. Through the connectivity of the machines on the assembly lines, output such as optimization parameters can be transferred within the shortest amount of time to the next machine and enable a consecutive execution of 3-D printers on the assembly line.

Augmented Reality

Augmented Reality allows the user an easy access to necessary digital data through the use of visual interfaces. To be more specific, the user is granted an "improved" version of reality by adding the data on to video footage taken by a camera on the used device all done in real time. (Masood and Egger, 2019)

The general application approach to Augmented Reality is to provide guides for workers in order to save time by faster rate of decision making and ease of access to useful data. (Rüssmann et al., 2015) Another possible scenario could be, an Augmented Reality application used by an employee in Additive Manufacturing, for the purpose of the visually analysing the 3-D model of a specific product for the purpose of providing ease to configure possible optimizations.

Autonomous Robots

The robots used in manufacturing offer more precise and efficient work compared to human workers while being able to operate under conditions that human limitations will not allow, also as the current robots keep on advancing, their operative capabilities will become more flexible and autonomous. (Rüssmann et al., 2015)

There are already existing Autonomous Robotics that are being used in various means of production. For instance, YumI, designed by ABB, an assembler robot, that has the necessary algorithms that enable learning and is able to operate with humans with added necessary safety measures, another example would be the Kuka LBR Iiwa, specifically designed for tasks that require extreme sensitivity with the capability of working and learning from humans that operate collaboratively. (Kamarul Bahrin et al., 2016)

Big Data

The collection of data can be made through various channels including electronic components like sensors or from digital platforms such as social media platforms. Through the collection and analysis of vast amounts of data from different categories, the best possible solution can be designed to be achieved whether it is for optimization or cost reduction in assembly lines. (Zhong et al., 2017)

Cyber Security

As the world becomes more and more digitalized and complex, so does the systems that operate resource management and facilities that partake in the storage and utilization of the information, such as energy grids, bank accounts and data storage units of companies like Google. Cyber-attacks to these facilities and units may results in colossal impacts such as loss of privacy for thousands, blackouts for cities. Such systems have not been invulnerable in the past as well as seen in Figure 2.5. (Cybersecurity for Industry 4.0 Cybersecurity implications for government, industry and homeland security, 2018)

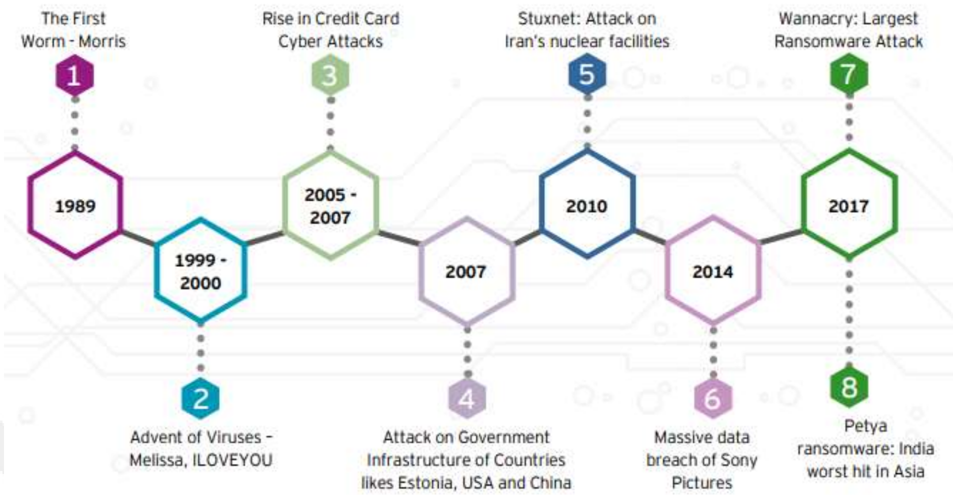

Figure 2.5: A Brief History of Cyber Terrorism (Cybersecurity for Industry 4.0 Cybersecurity implications for government, industry and homeland security, 2018)

Thus, Cyber Security is the notion that covers defence mechanisms and applications against threats and attacks that may come from the digital world. As well as digitalization, the connectivity is also increasing in the world by the employment of innovative IoT applications. This means that, with the increase in the rate of communication between devices and the number of channels, so does the risks which demonstrates the need for Cyber Security. (Cybersecurity for Industry 4.0 Cybersecurity implications for government, industry and homeland security, 2018) For instance, if a facility or a manufactory has not implemented any cyber defence system, the channels of communication between the control systems and the machinery will be unguarded against possible attacks. As Ervural and Ervural state (2018), in the example of the steel mill attack in Germany, the control systems were attacked and resulted in significant damages to the refinery.

Simulation

Simulations are used for a variety of reasons and in various areas of expertise. In most cases to provide abstract case studies to offer and validate possible solutions and optimize the implementation of such solutions. By applying desired adjustments through multiple runtimes in a simulation software, the user is able to obtain the best optimized case for specific requirements for a model. Furthermore, simulations can also be used to gather necessary data quantitatively through the creation of models in order to provide patterns, so that the inter relations within a system may be understood, but it should also be taken into consideration that the capabilities of a simulation and the expected outcome that is going to be observed rely on the capabilities and the specifications of the platforms it is being used and the variables that are being simulated. (Smith, 1998)

With regards to IR 4.0, the simulation systems will be able to provide real time data regarding the design and optimization of products possibly without them even leaving the plants where they are being manufactured. (Rüssmann et al., 2015)

Internet of Things

As it can be seen from the previous concepts regarding Industry 4.0 that were briefly explained, and the previous information regarding Industry 4.0, it can be said that the without the use of such a vital technology as Internet of Things, also referred to as IoT, the utilization of IR 4.0 would not be possible. It can also be said that, IoT is the element that started the transformation from the technologies and analog products of the third revolution into the current industrial revolution phase by relying more on the usage of digitalized data. (Vaidya, Ambad and Bhosle, 2018)

In the simplest terms, Internet of Things enables the wireless connection between components and the transfer of important data generated from these components. This is what is meant by the term interconnectivity, and since the machines are able to interact as if they were using a language, they are regarded as intelligent systems. (Rojko, 2017) Hence the reason why the term "smart" is used in almost every product related to Internet of Things, and has become a blanket term that can be applied to most consumer products. Observing from the Figure 2.6, it shouldn't be a surprise that the rate of increase in the number of connected devices correlates with that of the usage of the word smart on products.

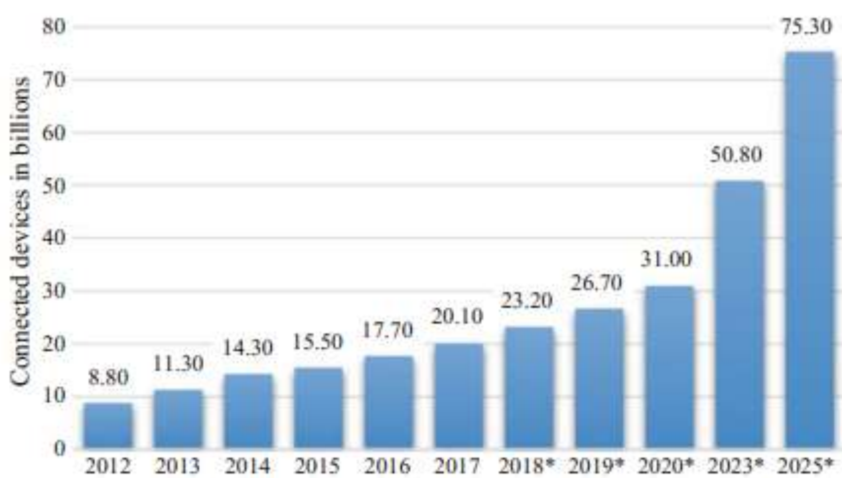

Figure 2.6: Number of Connected Devices through the Years (Ervural and Ervural, 2018)

As it was observed with the concepts that revolve around Industry 4.0, the areas that Internet of Things can be used may vary tremendously. One of the key application area was in manufacturing where its benefits will be described in the following section. Another "smart" implementation can be made through our households, where similar to manufactories, designated devices are able to transfer information freely within the household. This concept will be explained in more detail in its regarding sub section.

2.1.3) Current Stage and Usage of Industry Revolution 4.0

As of now, it has been 8 years, since the term Industrial Revolution 4.0 became the official term for the ongoing process of digitalization around the world, yet there are already a wide range of integrated solutions implemented by using IR 4.0. For an example, companies such as Festo, has already started the production and the sales of IR 4.0 related products. (Festo.com, n.d.) Not only the products but also the manufacturing processes of companies are going under a complete overhaul and the benefits are clearly observable, as stated by the CEO of Siemens, Joe Kaeser, production time was reduced by 24 fold and the output was boosted by 20 in 2016. (Greenfield, 2016)

Therefore, it is important to understand how much of a mass effect the new revolution is going to have in the world, especially with regards to manufacturing. Rojko (2017) states that, if the estimations are correct, logistics and production costs could be reduced between 10% and 30%, as well as a decrease in quality management costs up to 20%, if factories implement IR 4.0 solutions.

Besides the benefits that can be obtained by undergoing overhauls in manufacturing systems to be able to adopt IR 4.0, IR 4.0 solutions can also prove beneficial in management concepts such as supply chain managing and logistics. (Hofmann and Rüsch, 2017)

2.2) Smart Homes, Automation and Z-Wave

As mentioned in the previous sections, Internet of Things can be applied in various areas one of which is our households. By enabling the devices inside the households to transfer data and communicate, the concept of smart homes can be achieved. Furthermore, through this data transfer, applications and features inside the household can be regulated automatically.

Due to the interconnectivity that is available thanks to IoT, the devices do not need any outside interference to operate. The application areas may include heating and gas regulatory systems as well as the electrical usage of consumer products such as refrigerators and washing machines in homes, security components such as locks at doors and telephone applications. (Darby, 2017)

In addition to the ease provided by the utilisation of the concept, Smart Homes also reduce the energy wasted and increase efficiency on where they are integrated due to the fast paced feedback and self-monitoring that is provided. (Darby, 2017)

Now that the concepts, possible capabilities and benefits are briefly described, how such a technology can be integrated inside household should be described as well.

Z-Wave, is a technology specifically designed for home automation systems created by the Danish company Zensys in 1999. (Lamkin, 2019)

The concept fully uses Internet of Things, and enables the digital transfer of information within its configured automation system. Z-wave compatible electronics may consist of any sensors that are vital to a smart home system, these may include; temperature sensors, motion sensors and many more.

Z-wave devices use mesh networks to transfer vital information. Mesh networks operate differently compared to the usual and common Wi-Fi network configurations in regular households. Each and every device within the system is able to connect to each other, thus being able to propagate information from the source. (Fritz, 2019) To reinforce the understanding of a mesh network, the topology provided in Figure 2.7 is provided.

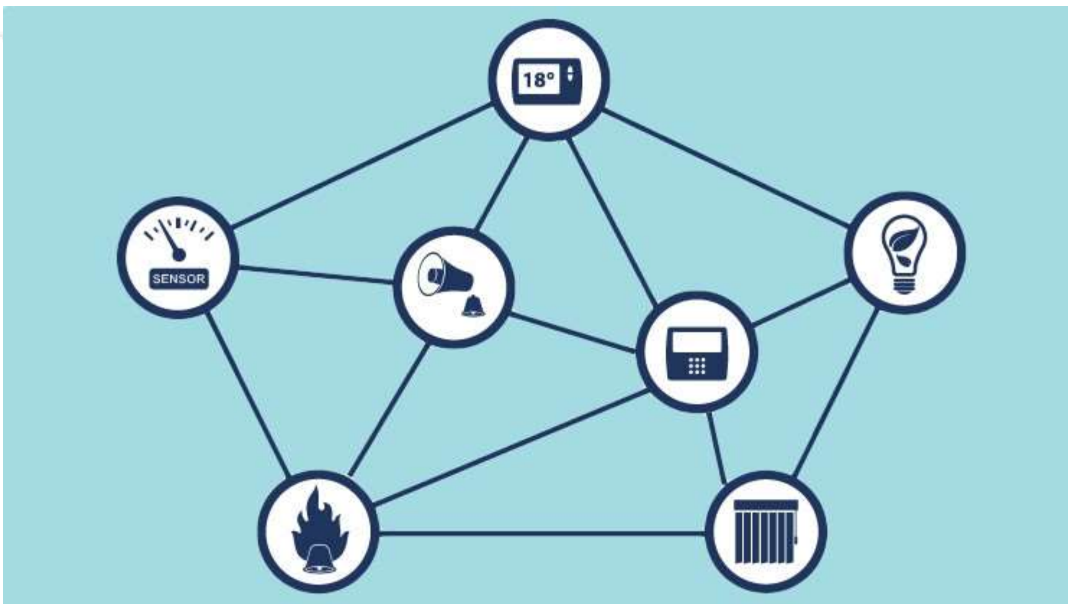

Figure 2.7: Mesh Network Topology (Lamkin, 2019)

By using the mesh network technology, the range difficulties that may occur in regular Wi-Fi and Bluetooth configurations are prevented. For an example, if a connected device is out of range from the source of information, another device within the range of the former device will be able to convey the necessary information, thus significantly preventing range issues.

In addition, it also operates on 908.42 low frequency band compared to 2.4GHz used in Wi-Fi and Bluetooth systems and implements ITU-T G.9959 global radio standard. This also proves beneficial, since most devices in a home that are connected to Wi-Fi networks use this band. As a result, any risk of possible interference that may arise from the incoming Wi-Fi signals becomes obsolete. (Lamkin, 2019)

2.2.1) Energy Estimation in Smart Homes through Artificial Neural Networks

Through the successful implementation of smart home concept and by using available low cost technological gadgets that are available with products such as Z-Wave, it is possible to monitor and record the certain types of data.

The recorded data can be used to optimize the energy usage of the household. One possible way of achieving this is through the utilization of ANN, artificial neural networks. An ANN is able to provide estimations regarding the behavioural pattern of the system that it is applied to. This is done by training the developed network by introducing it the necessary data sets that hold the vital information regarding the system the network is utilised for. It has also been noted that the higher the training, better performance can be observed from the results of the network, meaning more accurate representation of the desired behavioural pattern of the system. (Kowell et al., 2019)

By using the capabilities of the concepts mentioned before such as; Big Data and Internet of Things, huge amounts of data that was acquired can be used to create accurate estimations.

2.3) Agent Based Modelling and Simulations

Although there are many types of simulations that can be used for different scenarios and outcomes, there is one type that might prove more useful within the context of this research. An agent based simulation, refers to the practice of designing and implementing models through "agents"; individual components with definable and adaptable characteristics that can be used to create multi-dimensional, and diverse systems that can mimic real life applications. Such applications may include, power plants where not only the energy production is key issue, but the transportation of resources, environmental challenges and other logistical challenges are also prioritized. (Macal and North, 2005)

One counter argument to such a simulation technique could be, why bother so extensively to consider possible irrelevant side topics. But as it has been stated and explained in this review, the world is getting more digitalized and everything is getting connected by the implementations of concepts such as Internet of Things. By operating agent based systems on the complex and interconnected systems being put forward by IR 4.0, minimizing losses can become less risky and can be achieved with relevant ease. A possibility could the utilization of the concept through smart homes where the devices that are set up inside certain parts of the household, are able to operate freely while being able to transmit vital data for processing.

3) Methodology

In this section, the procedure of the research and the ways the data were gathered and processed will be explained thoroughly. In addition, the logic behind how and why the research data was chosen and analysed in order to be able to reach a conclusion for the research purposes will also be discussed.

3.1) Research Philosophy

There are four schools of thought that may be adopted before conducting a research. Briefly, these are as follows;

- Positivism: The research is about obtaining objective and scientific data through their statistical processing and done for the purpose of logic and dispatched form value systems.
- Realism: The research is about observing the world as it is. To be more specific, utilization of realism implies using direct data form observations from human senses which is referred direct realism as or using multi-layered techniques to reach data which is referred as critical realism.
- Interpretivism: Unlike realism and also positivism, interpretivism focuses on human expedience and values, thus relies on the collection of subjective data in order to achieve research objectives.
- Pragmatism: The research question is more prominent compared to the data that may be involved, therefore the methods and analysis is more flexible compared to the other philosophies. (Research-Methodolgy, n.d., a)

Thus, the procedure of adopting one of these schools of thought depend on the kind of data and the objectives of the research that is required for the research and the methods.

For the purposes of this research positivism has been chosen as the philosophy of the research. Since the objectives require mathematical observations and numerical conclusion deduction from the experiments that will be designed and simulated, this school of thought is the best suitable one.

3.2) Research Approach

There are several types of reasoning approaches that can be adopted by researchers regarding their research. The three types of approaches are; deductive, inductive and abductive approach. Briefly;

- Deductive: Using tests and calculations to prove or disprove a theory,
- Inductive: By following a bottom to top approach, observing research data and creating a theory as a result,
- Abductive: The goal is to achieve the best explanation for the outputs gained from the research (Research-Methodolgy, n.d., b)

In this research abductive reasoning has been utilized. The reasoning behind this decision was that the collected data will be used to get the best option available and not to confirm or reject a theory.

Instead, the best solutions will be used to configure a possible framework model for future use. After collection of the data, an algorithm will be followed as depicted in the flowchart below.

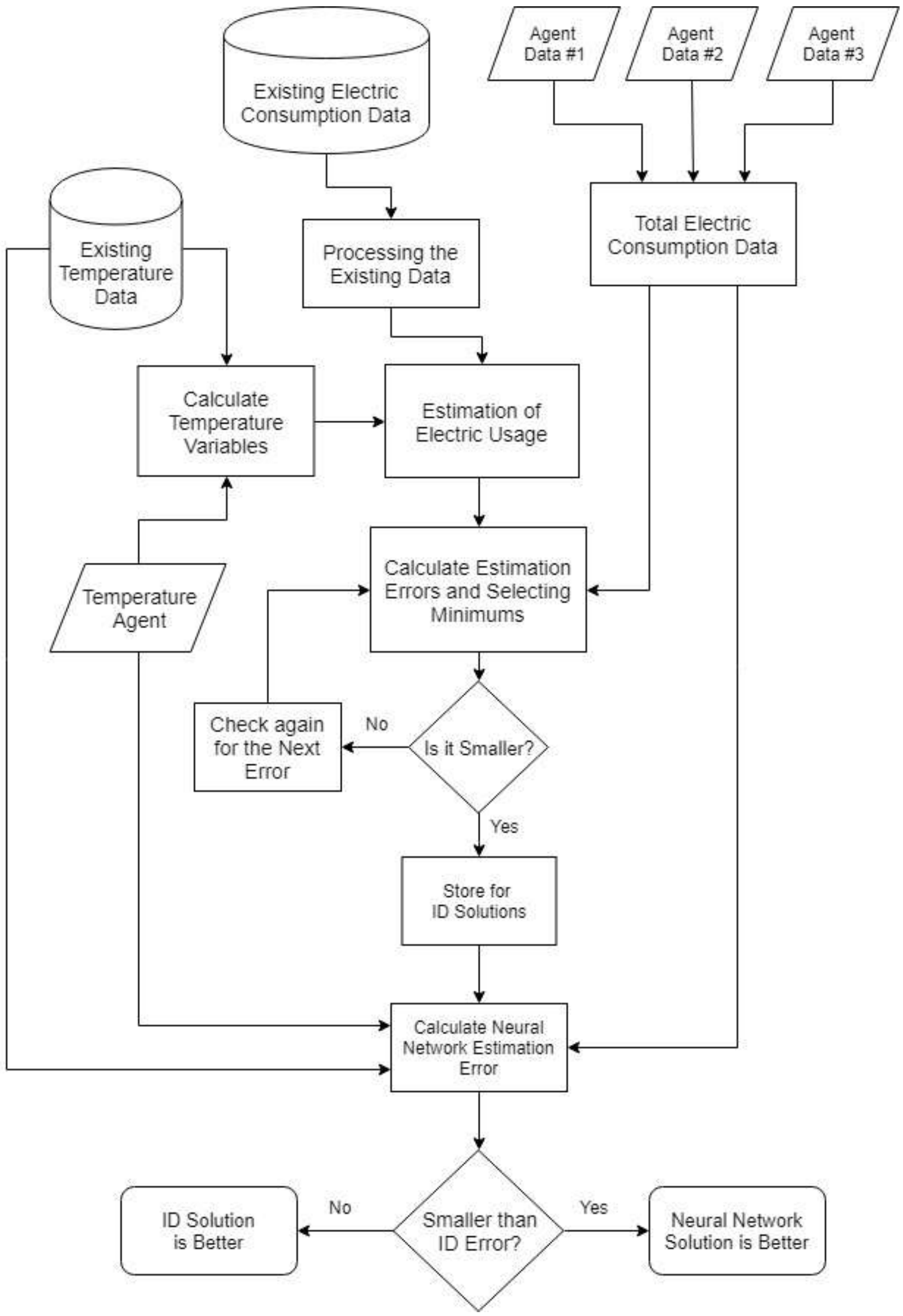

Figure 3.1: Flowchart of the Research Approach

The flowchart depicted in Figure 3.1, introduces the process in which the research will follow through to be able to create a framework for future households. It will consist of making estimations based on the existing data sets and measured electric consumption data, which are explained in the following sub section, and a number of temperature variables which will be explained in detail in their respective experiment sub section in section 4. The precision of these estimations will be measured by the errors calculated from actual results obtained through experimentations versus the estimations made. Finally, through comparing the different experimentation results, best solutions achieved with the minimal estimation errors will be chosen as the basis for the possible framework.

3.3) Research Data and Collection

Two sets of data will be used in order to be able to generate cost optimization which is the main goal of this research project. In this section, these two data sets will be identified and the acquisition ways for both will be explained briefly.

3.3.1) Data Set #1

The primary research data set obtained for the research was provided by the supervisor of this research. This data provides recorded logs of kWh usage of over three thousand apartment flats in Cork, Ireland and was recorded across a period of almost a year and a half. The kWh data in these logs will be analysed with regards to the date and time they were observed.

3.3.2) Data Set #2

The second data set will be composed of data that was gathered from an agent based simulation test run. Thus, this data set was not available at the start of the research as opposed to the primary data set. The quantitative method of gathering the necessary data will involve several sensors that will be set up in a smart home configuration. Each of these sensors that are included in the Z-wave mesh network will be acting as independent agents in the system.

Due to both set of data being gathered by quantitative research methods, which use standardised measurements to observe and record research data, both of these data sets are going to be considered primary data type. (Research-Methodolgy, n.d., c) Although the first data set was readily available, which is a common property of secondary data, the data was collected through specific time intervals and the electrical usage data for each household categorized. Thus, it falls under the category of primary data.

3.4) Methods of Data Analysis

The research aims to form the basis of a conceptual framework, from which innovative solutions can be derived for building smart homes and other building applications and if possible other fields within the scope of Industry 4.0. In order to reach such a framework, the data sets that are involved and used for this project will have to be analysed. The research requires a quantitative approach regarding both of the data types that were used. The reasoning behind this approach is that the raw data observed from the readily available data set #1 and the sensor data that is going to generate data set #2 needs to be transformed into an output form that can be utilized to form correlations and as a result constitute the conceptual framework. (Research-Methodolgy, n.d., d) This means that numerical outputs will be will be gained after the analysis of both data sets to reach a conclusion.

The primary data set that was explained in the previous sub section, was received in large .txt files, and in order to be able to read and process the data set, an appropriate platform was chosen. For this reason, the coding language Python was used on Eclipse IDE. The mechanics behind the analysis will be explained in detail in the next section 4.1.2. Through this analysis the numerical data of the electrical usage in various households across will be stored and become available for estimation operations for the chosen benchmark time period.

As mentioned in the previous sub section, section 3.3.2, the secondary data set is obtained through a simulation by using a multi agent Z-wave hub. As it was explained in the respective Literature Review sub section, section 2.2, Z-wave networks are specifically designed for home automation systems, hence the reason, Z-wave adaptable sensors were used to gather and later analyse the sensor data. The analysis of the data set #2 is done by monitoring and categorizing the sensor data obtained by the home automation system. The output from the system will be used as the measured data to calculate the estimation errors from the estimations made with the existing data as observed in the research flowchart.

3.5) Research Objectives

The objectives of the research can be broken down into several milestones. These are as follows;

By categorizing and analysing the primary data, the behavioural pattern of the numerous households that were included in the data collection will become available to observe. The analysis of the data set will provide the necessary means to make estimations from the available data. The successful analysis and display of these results will act as the first milestone.

After successfully implementing the smart home concept through the utilization of Z-Wave and multiple sensors mentioned afore, in the test room, the sensor data will be gathered. The analysis will depict the usage of kWh over the test period. After obtaining secondary data from the smart home configuration, a benchmark model for the individual electrical usage will be obtained. Which will be the data obtained from the room between the start and the end time of the simulation. By achieving this, how the electricity was used in the experimental household will be displayed similarly to the analysis of the primary data set. The successful implementation and the demonstration of the smart home and the acquisition of the results will be the second milestone.

After the calculations of the temperature variables, which are necessary for the estimation process, the estimation errors will be found for the benchmark data. This will be followed by finding the best possible options from the available households by correlating the minimum estimation errors of data set #1 and the results obtained from the previous milestone. Finally, a list of household IDs will be presented as the most fitting IDs that can be used for future use for smart homes. The correlation process and the ID solutions will be the third milestone.

After acquiring the results mentioned above, the final milestone that will be reached by comparing the estimation errors with another method for creating estimations. Through the usage of artificial neural networks, a set of estimations will be generated for the experimental room. The errors obtained by implementing this concept, will be compared with the estimation errors of the IDs that were chosen as the best solutions from the previous milestone and the results that provide the minimum of the errors will be chosen for the creation of the framework, thus concluding the objectives of this research and act as the final milestone.

4) Experimentation

In this section the processes that were used to analyse the necessary data for the research and the systems that were used will be explained thoroughly. Since there are two types of data as explained in section 3.3, explanation of the first two experiments will be covered in two sub sections. For both data types, first the inputs that went into the system will be explained, followed by how they were used and analysed by their respective experimental system and finally the results obtained from both of the experiments. Finally, the last experimentation that was conducted, will be explained with regards to its purpose and how it used the data obtained from the previous experiments in the same manner with regards to the first two experiments.

4.1) Primary Experiment

The primary experiment is the analysis of the primary data described in section 3.3.1. The data was received in large .txt files, and they were read and processed in Python. The code that was generated will also be explained concisely as possible, to be able to illustrate the analysis process. In addition to the explanations, the additional libraries will be explained briefly as well to consolidate and reinforce the analysis done by the code.

4.1.1) Inputs

The data files that contain the kWh logs were used as the input for this experiment. In this sub section the format of the logs will be briefly described. As it can be seen from the table below, the log data are placed in order with ID, 5-digit decimal time format and kWh value respectively. Unfortunately, the data were randomly placed and were not in order with regards to any variable.

Table 4.1: A Fragment of the Raw Input Data from the Existing kWh Data

4033	20045	0.943
4033	20046	0.611
4033	20047	1.558
4033	20048	1.334
4033	19546	0.325
4033	19547	0.314
4033	19548	0.317
4033	19601	0.304
4033	19602	0.299
4033	19603	0.216
4033	19604	0.217
4033	19605	0.22
4033	19606	0.22

As mentioned before in section 3.3.1, there are approximately 3500 individual apartments involved across a time period of 1.5 years, approximately. For each day there are 48 logs, meaning that kWh data were measured every 30 minutes and displayed in the last 2 digits of the time variable. The first 3 digits represent the date in *dd/mm/yy* format. The first of this 5-digit variable being 19501 and representing *01/01/2009 0:00:00 – 0:29:59* and the last one, 71648, representing *06/06/2010 23:30:00-23:59:59*

4.1.2) Analysis Process

First, the folders holding the texts of kWh usage with their IDs and dates are extracted into new directories by using the built-in libraries *os, zipfile and shutil*. Next, the extracted text files are read, stored and sorted with regards to their ID and their dates, since the logs are placed randomly, as depicted in the previous section.

A benchmark version of data is needed in order to be able to generate a framework for further electrical usage. The benchmark data must overlay the data that will be gained from the secondary experiment, which will be obtained by a test run from the simulation room, which will be explained in the analysis sub section of the secondary experiment, section 4.2.2.

The benchmark data was obtained by separating the necessary kWh data for the time period used for the test run across every ID. In addition, the decimal format of the time variable for each kWh data was transformed into *dd/mm/yy* format as well, by a method that used the built-in library *datetime* of Python. These were later written in CSV format by implementing *csv* library and with respect to their IDs for later use in the following sections. The results from this experiment can be observed from the table below.

Table 4.2: A Fragment of the Analysed kWh Values for Household #4000, from 12:00 to 18:30 in 26th August 2009

The results for 4000s values are;			
Values of 4000 " : Mean = " 0.305			
43225	26/08/2009 12:00:00-12:29:59	4000	1.354
43226	26/08/2009 12:30:00-12:59:59	4000	0.993
43227	26/08/2009 13:00:00-13:29:59	4000	0.447
43228	26/08/2009 13:30:00-13:59:59	4000	0.354
43229	26/08/2009 14:00:00-14:29:59	4000	0.284
43230	26/08/2009 14:30:00-14:59:59	4000	0.72
43231	26/08/2009 15:00:00-15:29:59	4000	0.67
43232	26/08/2009 15:30:00-15:59:59	4000	2.695
43233	26/08/2009 16:00:00-16:29:59	4000	1.695
43234	26/08/2009 16:30:00-16:59:59	4000	0.646
43235	26/08/2009 17:00:00-17:29:59	4000	0.456
43236	26/08/2009 17:30:00-17:59:59	4000	0.469
43237	26/08/2009 18:00:00-18:29:59	4000	1.403

4.2) Secondary Experiment

The secondary experiment will follow the same format as the former experiment and analyse the respective data set. As mentioned in the regarding sub section in Methodology, section 3.3.2, the data was gathered through multiple sensors in a smart home configuration by using a mesh network made available by using Z-wave technology, and in this experiment it was implemented by using an Aotec Z-Stick Gen2 acting as the hub, and all sensors were connected to this device. A Raspberry Pi Model 3 was used to monitor the data. The RPi was chosen for this purpose due to, it being a compact minicomputer and available to everyday users throughout the world.

Similar to the analysis of the first experiment, first the inputs will be explained, in this case the sensors that gathered the data and the data they were obtaining. This will be followed by the explanation on how the sensor data were processed and finally the outputs that were generated from this experiment.

4.2.1) Inputs

There are 5 sensors in total that were used for data gathering. These include 3 voltage sensors, 2 4-in-1 sensors. The sensors that were used to gather the necessary data are listed in the following table;

Table 4.3: Sensor List Used in the Experimental Room

Sensor Name	Data Type	Output Type or Unit
(TZ69E Plus, Socket with Power Meter) x3	Power, Voltage, Ampere	W, V, A
Fibaro Motion Sensor	Motion, Illuminance Temperature, Vibration	Alarm, Lux, °C, Alarm
Hauppauge mySmartHome 4-in-1 Sensor: 01559	Motion, Illuminance Temperature, Accelerometer	Alarm, Lux, °C, Alarm

4.2.1.1) Depictions of the System

The smart home configuration that was setup in the experimental room, will now be depicted in two different ways. The first illustration will simply depict the system as it is and will depict the electronics as it is seen in Figure 4.4. The second illustration will depict the system in the agent based simulation configuration and this can be seen in Figure 4.5.

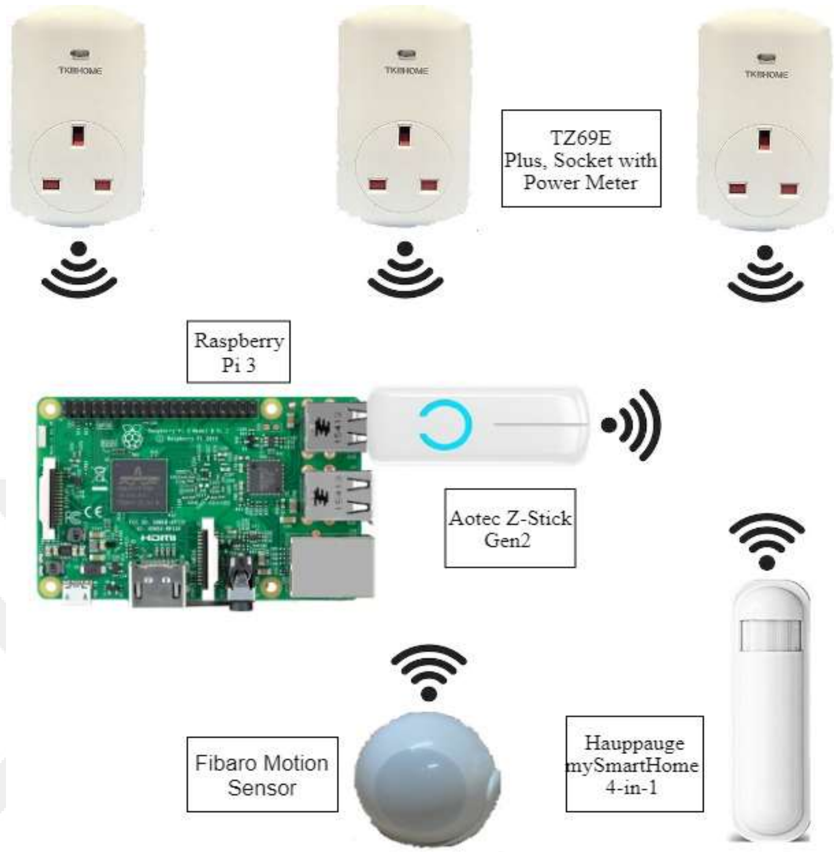

Figure 4.4: Electronic Configuration of the Experimental Room

As mentioned in the previous section and as seen in Figure 4.4, the sensor agents are connected wirelessly via Z-Wave network to the Z-Stick, which is operated by the USB connection to the Raspberry Pi 3, where the Domoticz server is running continuously while the Raspberry Pi is powered.

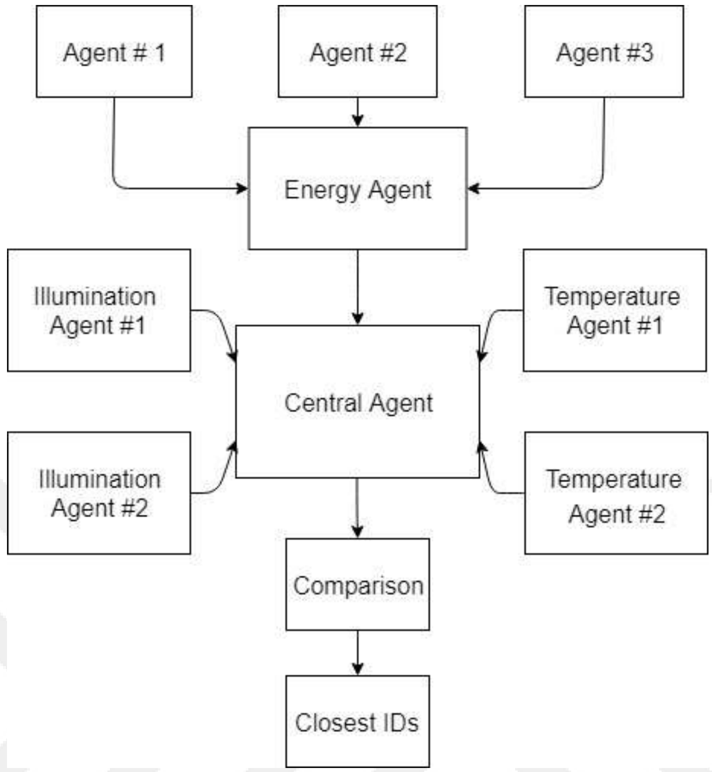

Figure 4.5: Depiction of the Model Room with regards to Agent Based Modelling.

As explained in section 2.3, agent based simulations can be beneficial when modelling systems. For this system, every sensor and device that constituted the model were considered as an agent in the system that acted independently. In addition, since the 4-in-1 sensors also gathered the temperature and the illumination data separately, these were depicted as independent agents in the system as well. The Central Agent is the hub in the room, where data is gathered through the sensors and monitored, in this case it depicts both the Raspberry Pi 3 and the Z-Stick since both depend on each other with regards to their functionality to the experimentation. The blocks Comparison and Closest IDs will be explained in the Analysis section as those are the parts of the system that produce the output for the secondary experiment as explained in section 3.5.

4.2.2) Analysis Process

Unlike the first experiment, there was no coding involved in the analysis of the input data of this simulation. After setting up the sensors in the smart home configuration using the Z-wave mesh network, the software application Domoticz was used to store and log the necessary sensory data. To summarize, the software allows users to control and automate their own home system and monitor the data gathered by the sensors included in the Z-wave network.

Idx	Hardware	ID	Unit	Name	Type	SubType	Data	Last Seen
59	Zwave-USB-HUB	0002[redacted]00	1	Sensor	Light/Switch	Switch	Off	2019-08-28 12:29:34
60	Zwave-USB-HUB	0002[redacted]35	1	Home Security	Light/Switch	Switch	Off	2019-08-28 12:29:34
67	Zwave-USB-HUB	0002[redacted]35	7	Alarm Type: Home Security 7 (0x07)	General	Alarm	Event: 0x00 (0)	2019-08-28 12:29:34
62	Zwave-USB-HUB	002[redacted]03	255	Illuminance	Lux	Lux	172 Lux	2019-08-28 12:13:32
39	Zwave-USB-HUB	0002[redacted]01	1	kWh Meter	General	kWh	0.193 kWh	2019-08-28 12:13:27
40	Zwave-USB-HUB	002[redacted]01	2	Unknown	Usage	Electric	3.9 Watt	2019-08-28 12:13:27
41	Zwave-USB-HUB	0002[redacted]01	1	Voltage	General	Voltage	227.9 V	2019-08-28 12:13:27
42	Zwave-USB-HUB	[redacted]01	0	Unknown	Current	CM113, Electrisave	0.0 A	2019-08-28 12:13:27
43	Zwave-USB-HUB	0002[redacted]01	1	Percentage	General	Percentage	0.47%	2019-08-28 12:13:27
44	Zwave-USB-HUB	0002[redacted]01	1	kWh Meter	General	kWh	0.015 kWh	2019-08-28 12:13:27
45	Zwave-USB-HUB	002[redacted]01	2	Unknown	Usage	Electric	0 Watt	2019-08-28 12:13:27
46	Zwave-USB-HUB	0002[redacted]01	1	Voltage	General	Voltage	227.2 V	2019-08-28 12:13:27

Figure 4.6: Devices Included in the Hub, Displayed by Domoticz

After Domoticz stores the data of the sensors, every possible output from the sensors can be viewed and many can be plotted, such as; illumination, temperature, Voltage and power. Several output plots from the sensors are displayed below.

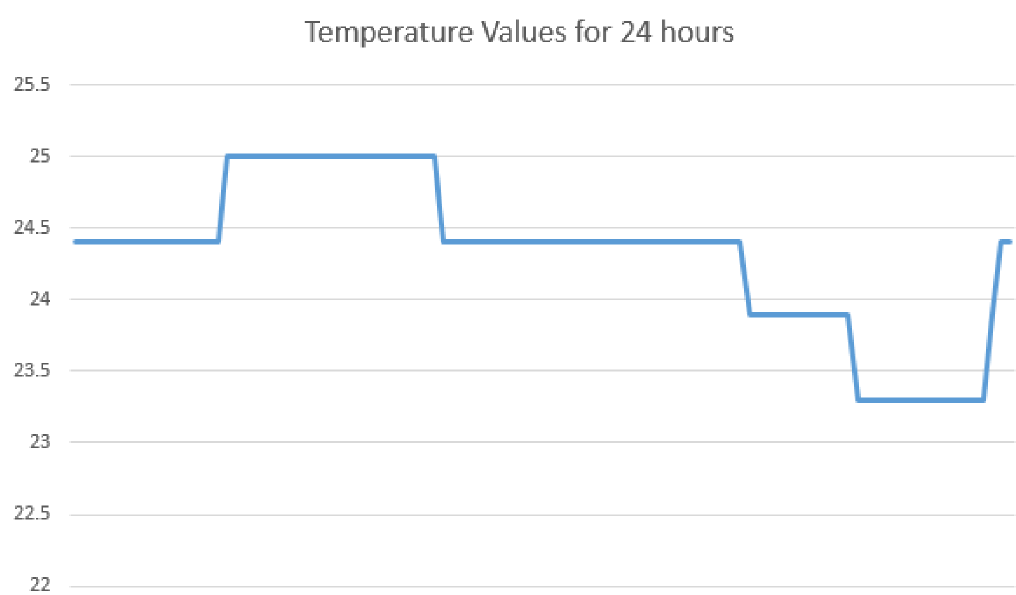

Figure 4.7: Temperature Plot of the Experimental Room, 27-28 August 2019

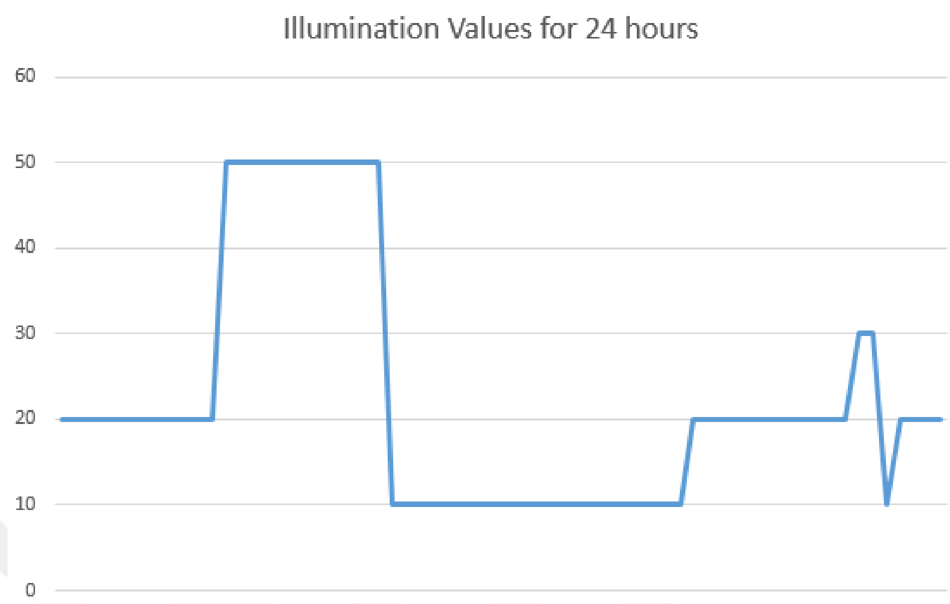

Figure 4.8: Lumination Plot of the Experimental Room, 27 to 28 August 2019

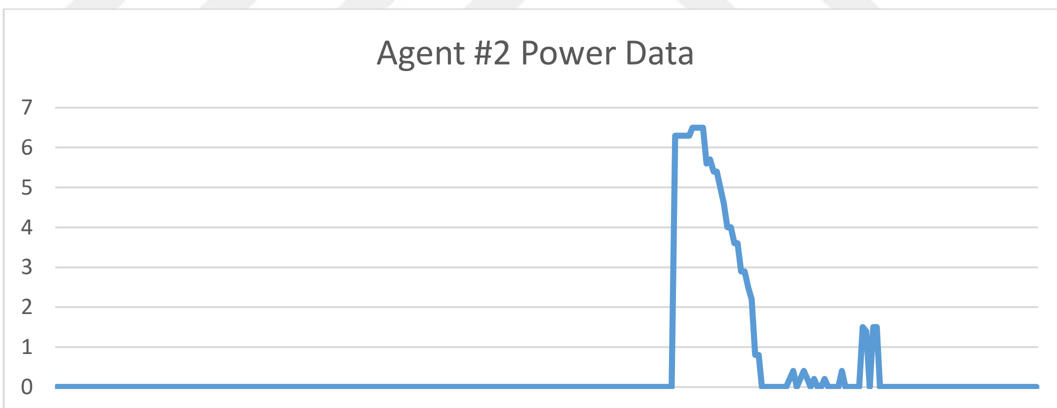

Figure 4.9: Power Data of Agent #2, 27 to 28 August 2019

In addition to the plots, Domoticz can also process and store the data in CSV format as well. When CSV form is used, the recorded sensory data is displayed in 5 minute intervals.

Although there are several types of sensory data that can be observed and can be commented on, for this research's objective, the data recorded from the voltage sensors was used to draw correlations. The data will be obtained in the CSV format as mentioned before.

In this experiment, a test run that started from August 26th 2019 12:00 and ended on August 28th 2019 12:00 was used to gather the output of this experiment. The power data in CSV format generated by Domoticz by the test run of 48 hours, was used as a benchmark in order to draw an estimation correlation.

In order to be able to derive estimations from existing data from Cork for the electrical usage models in Exeter student accommodations, a relation had to be found between the consumption data and an another variable. For this purpose, the temperature data from both cities was used as this variable. As Cruz Rios states (2017), high electricity consumption is observed when the temperature either keeps getting decreased below a certain point or increased above a threshold as it can be observed in the figure below.

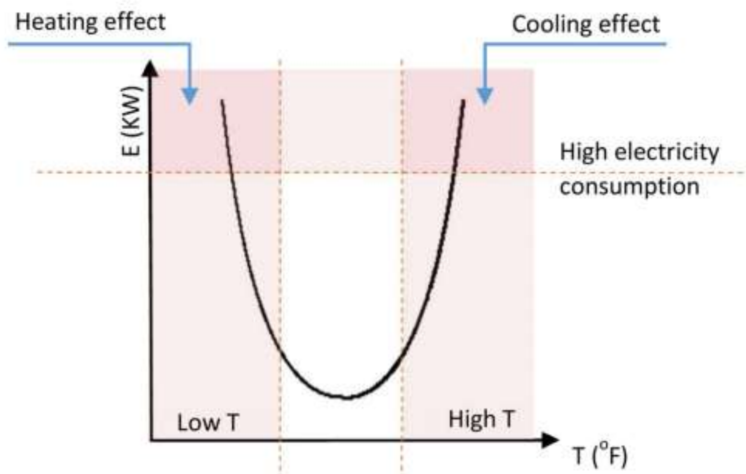

Figure 4.10: Relation Between Temperature and Electric Consumption (Cruz Rios et al., 2017)

First the relation between the two sets of temperature data was found by due to the relation observed in Figure 4.10 due to the cooling effect since the benchmark time period was during Summer;

$$T_1 = \alpha \times T_2$$

Where T_2 is the temperature data from Cork and T_1 is the temperature data from Exeter. By using the data electrical consumption data from Cork and the α value, an estimation value will be achieved by;

$$Exeter_{Estimation} = \alpha \times Cork_{Usage}$$

However, directly creating estimations directly from 10-year-old temperature data may not be so accurate. This inconsistency is mainly due to the reality of climate change. It is already estimated that the average global temperature will increase between 2.5 °C and 10 °C in the next 100 years, which will have everlasting effects, potential sociological and economical harm and severe conditions to live in. Furthermore, according to IPCC, the Intergovernmental Panel on Climate Change, the increase in the will vary from region to region across the globe. (Jackson, 2019)

It has already been reported by the Met Office that the UK has faced its warmest 10 years since 2002 and the highest temperature record was broken with 38.7 °C, in Cambridge on 29th of July 2019. (McGrath, 2019) With regards to the regions around Exeter and Cork, it is estimated that an increase around 2.7 °C and 2.5 °C respectively, in the next 60 years approximately, as demonstrated in the figure below. (Metlink | Weather & Climate Teaching Resources, n.d.)

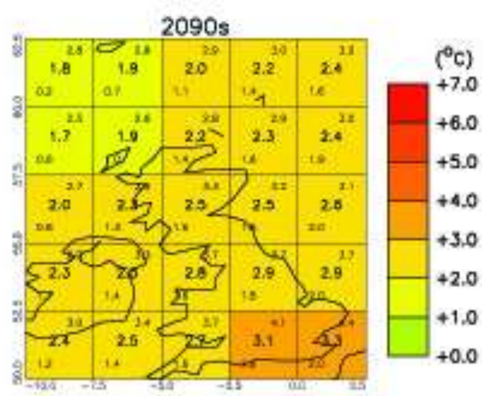

Figure 4.11: Estimated Temperature Increase in the UK in 2090s (Metlink | Weather & Climate Teaching Resources, n.d.)

As explained above, due to the ongoing effects of climate change caused by global warming, average temperatures across the globe are expected to see an increase. Therefore, a new with an additional variable must be introduced to demonstrate the ratio of change of the average temperature from 2009 to 2019 in the cities of Cork and Exeter.

$$T_{Exeter,2009} = \alpha_1 \times T_{Cork,2009}$$

$$T_{Exeter,2019} = \alpha_2 \times T_{Cork,2019}$$

$$T_{Cork,2019} = \beta_1 \times T_{Cork,2009}$$

$$T_{Exeter,2019} = \beta_2 \times T_{Exeter,2009}$$

Variables β_1 and β_2 are the change variables representing the change in temperature for the 10-year period. Thus for the most accurate estimation values, a new formula must be derived;

$$T_{Exeter,2019} = (?) \times T_{Cork,2009}$$

The following formula was derived from the four equations depicted above;

$$T_{Exeter,2019} = \frac{\alpha_2}{\beta_1} \times T_{Cork,2009}$$

For simplicity, the symbol γ will be used for $\frac{\alpha_2}{\beta_1}$.

By using the γ value, expected values for electric usage are calculated by multiplying with the analysed mean kWh data of the 3444 households, obtained from Cork. Thus;

$$Exeter_{Estimation} = \gamma \times Cork_{Usage}$$

The hourly temperature data from Cork, Ireland was obtained from the historical data base of Met Eireann and the Cork Airport's Station was chosen due to its closeness to the city compared to Moore's Park and Roches Point Lighthouse stations in order to get the most accurate temperature results. The air temperature data was archived hourly, so for the research purposes the half an hour mark will be taken as the mean of temperature data of the hour and the following hour.

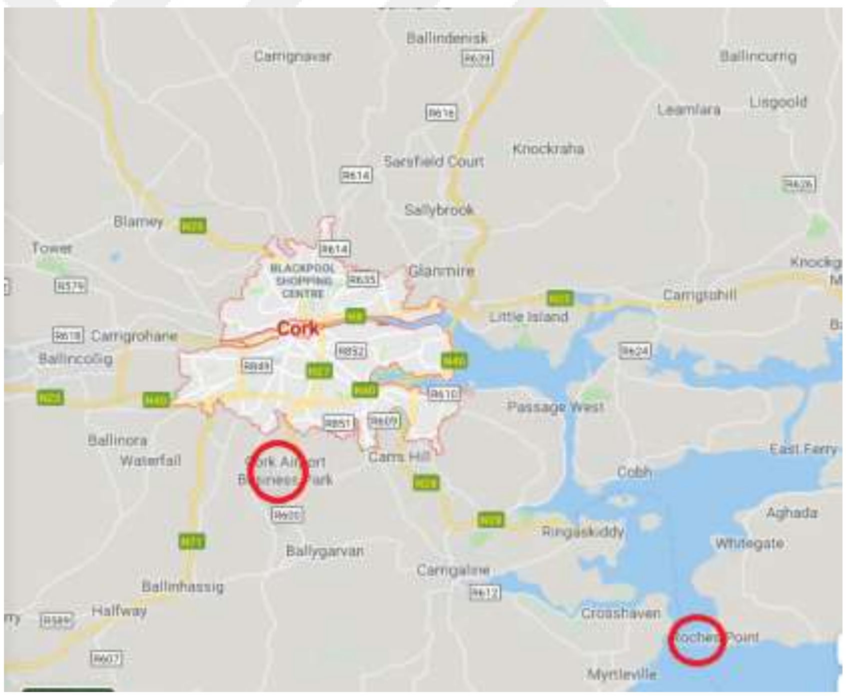

Figure 4.12: Locations of the Stations Within the Map of Cork, Ireland.

Figure 4.13: Location of Moore Park with Regards to Cork, Ireland

The hourly weather data for Exeter during the simulation period was also obtained for the time period between 26th August 2009 12:00 and 28 August 2009 12:00 as well as the data for 26th August 2019 12:00 and 28 August 2019 12:00. The temperature data wasn't obtained from the sensors that were set up inside the experimental room. The reason for this, is that both sensors were placed in different parts of the room and the measured data obtained from those sensors were recording room temperatures. Since the data obtained from the Cork Airport weather station is the air temperature, air temperature data was obtained from an official air temperature source as well in order to maintain consistency of the experiment. The Exeter Airport weather station was used to obtain the air temperature similar to the data obtained from Cork Airport weather station.

After the gathering of the four sets of temperature data, the following plots were generated to depict the temperatures of the necessary time periods;

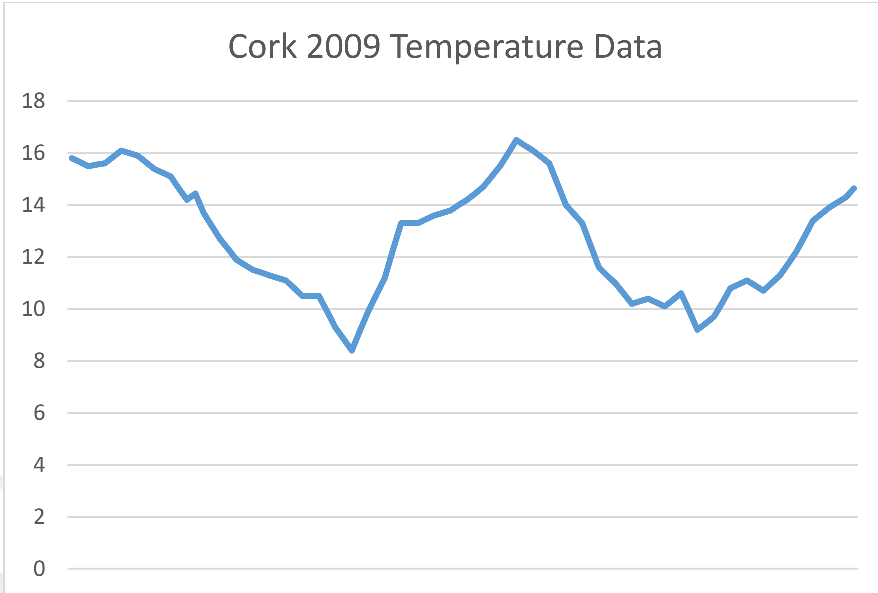

Figure 4.14: Temperature Plot of Cork, between 26th August 2009 12:00 and 28 August 2009 12:00 (Met.ie, n.d.)

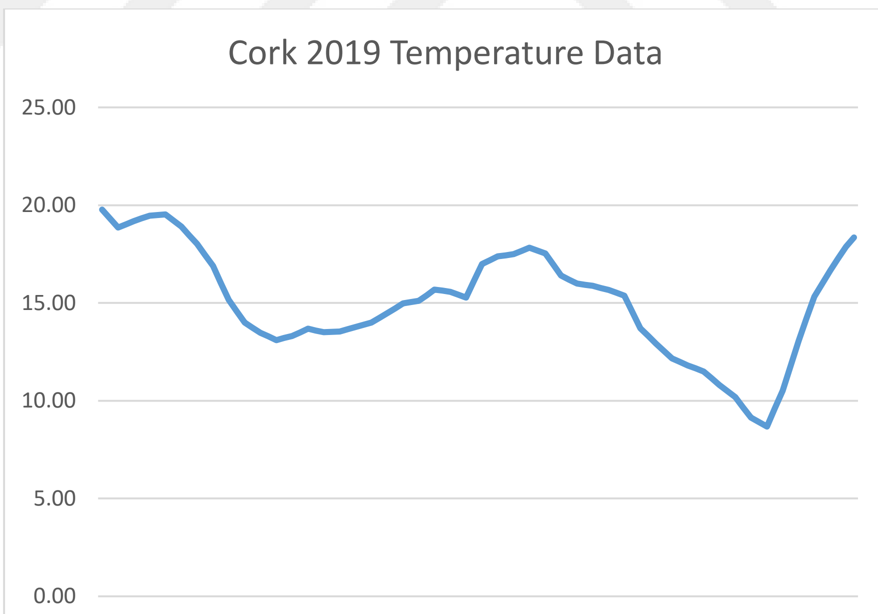

Figure 4.15: Temperature Plot of Cork, between 26th August 2019 12:00 and 28 August 2019 12:00 (meteoblue, n.d., a)

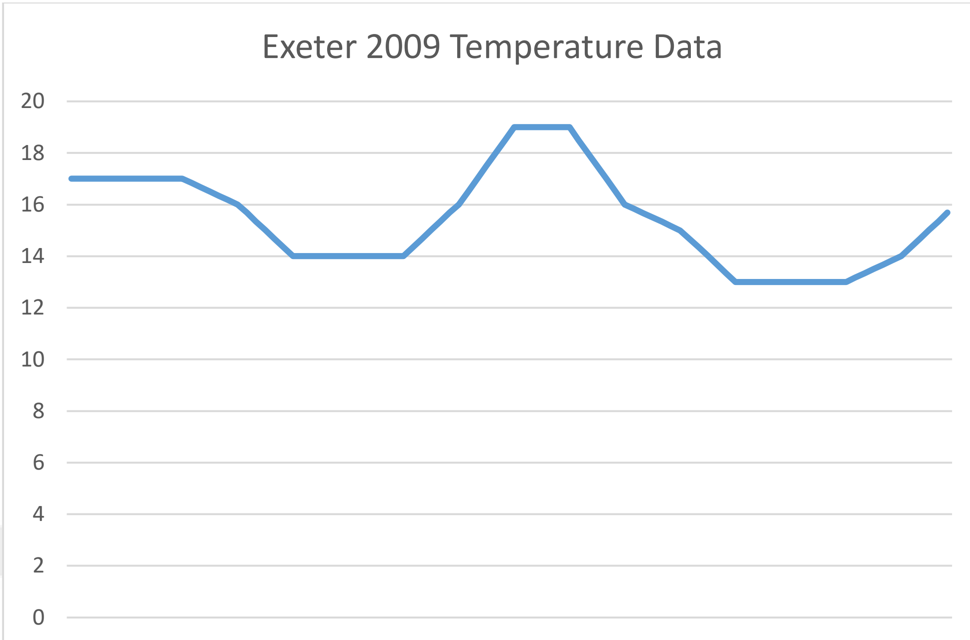

Figure 4.16: Temperature Plot of Exeter, between 26th August 2009 12:00 and 28 August 2009 12:00 (WorldWeatherOnline.com, n.d.)

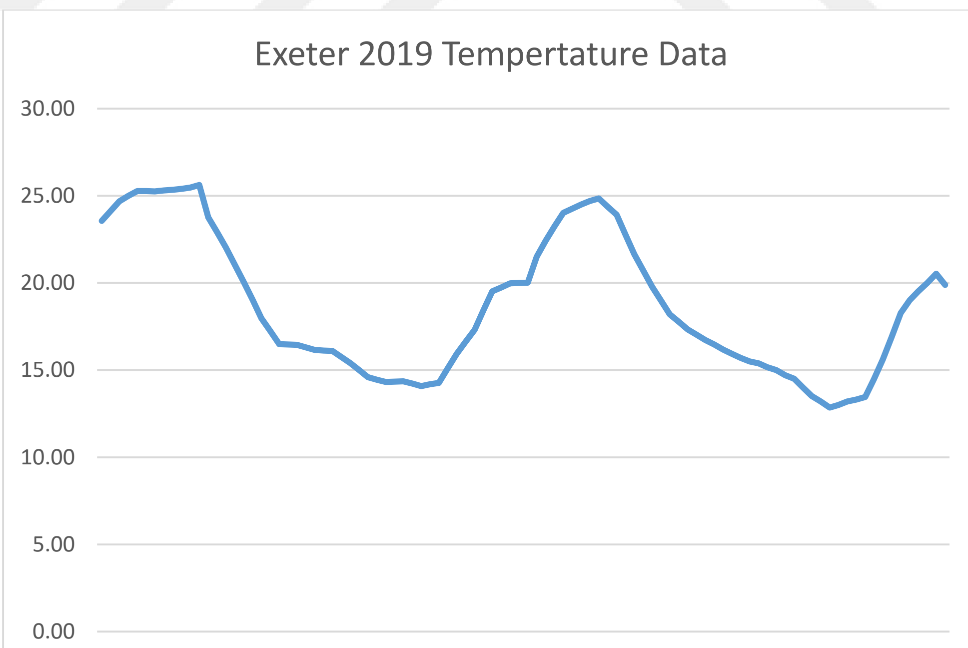

Figure 4.17: Temperature Plot of Exeter, between 26th August 2019 12:00 and 28 August 2019 12:00 (meteoblue, n.d.,b)

From the temperature data obtained and as explained and depicted in the plots above, the necessary variables for the estimation calculations were found as seen in the following table;

Table 4.4: α_1, β_1 , α_2 ,β_2 , γ Examples Between 12:00 and 18:00 for 26th August

Time and Date	α1	β1	α2	β2	γ
26th August 12:00	1.076	1.251	1.191	1.385	0.952
26th August 12:30	1.086	1.234	1.249	1.419	1.012
26th August 13:00	1.097	1.216	1.309	1.452	1.077
26th August 13:30	1.093	1.223	1.313	1.469	1.074
26th August 14:00	1.090	1.230	1.317	1.486	1.070
26th August 14:30	1.073	1.220	1.307	1.486	1.072
26th August 15:00	1.056	1.209	1.297	1.485	1.072
26th August 15:30	1.063	1.218	1.298	1.488	1.065
26th August 16:00	1.069	1.228	1.298	1.491	1.057
26th August 16:30	1.086	1.227	1.322	1.494	1.077
26th August 17:00	1.104	1.227	1.347	1.498	1.098
26th August 17:30	1.115	1.210	1.387	1.506	1.146
26th August 18:00	1.126	1.193	1.319	1.398	1.105

As depicted in section 4.2.1.1, there are 3 individual socket agents working within the simulation model that gather the electricity consumption data. These agents measure the Raspberry Pi, smart phone charger and the kettle power data in 5 minute intervals. In order to use the simulation data appropriately, the following steps were applied;

Since the logs of data displayed in the previous parts were the total electric usage data of a single household, the same must apply to the experimental room. So, the power data obtained was summed to depict the total usage inside the experimental room;

$$W_t = W_1 + W_2 + W_3$$

Furthermore, since the data logs from Cork were measured and recorded in 30 minute intervals, the experimental data must be adopted to this standard as well for the sake of consistency. So, the data was averaged as seen in the following equation;

$$W_{30min} = (W_{5min,t1} + W_{5min,t2} + W_{5min,t3} + W_{5min,t4} + W_{5min,t5} + W_{5min,t6}) \div 6$$

Finally, since the data recorded from the experimental room is for power consumption (W) and not the energy usage (kWh), the following conversion was also applied to the total power usage in the following manner;

$$kWh_{30min} = (W_{30min}) \times (0.001) \times (0.5)$$

Where (0.001) is for the scientific notation of the symbol k, and (0.5) is due to the 30 minutes being half of an hour. With the actual results obtained, the error for estimations can now be calculated by the following equation;

$$Estimation\ Error = Actual\ Result - Estimation\ Value$$

Table 4.5: Mean Data of Cork, Expected Values, Observed Data and Estimation Error Between 12:00 and 18:00 for 26th August

Time and Date	Mean of Cork Values(kWh)	Expected Values	Actual Results	Estimation Error
26th August 12:00	0.779	0.742	0.00195	-0.73965
26th August 12:30	0.781	0.791	0.00193	-0.78871
26th August 13:00	0.757	0.815	0.00190	-0.81309
26th August 13:30	0.719	0.772	0.00190	-0.77012
26th August 14:00	0.704	0.754	0.00190	-0.75172
26th August 14:30	0.71	0.761	0.00190	-0.75888
26th August 15:00	0.736	0.789	0.00190	-0.78738
26th August 15:30	0.753	0.802	0.00193	-0.80010
26th August 16:00	0.79	0.835	0.00195	-0.83341
26th August 16:30	0.854	0.920	0.00198	-0.91795
26th August 17:00	0.916	1.005	0.00205	-1.00338
26th August 17:30	0.917	1.051	0.00203	-1.04893
26th August 18:00	0.884	0.977	0.00194	-0.97478

By using the data obtained from the secondary experiment, an additional module was created with Python to observe the best possible IDs that could be used. First by using the γ values for each timeframe estimations were made for each ID and later individual ID estimation errors were calculated by subtracting the actual kWh value measured by the agents in the experimental room for each time period on the benchmark.

Later on, the best solutions were found by gathering the 100 IDs that provided the maximum number of minimum estimation error values across each benchmark time interval. The IDs that didn't existed or recorded only 0 were excluded since they can't be used for potential frameworks. These IDs were then written in CSV format and provided as the output of this experiment. In Table 4.6, solutions from different IDs are presented, sorted in numerical order from the lowest to the highest and the kWh data from 12:00 to 18:29.

Table 4.6: Various ID Solutions from the Output of the Experiment

4004	0.016	0.016	0.035	0.124	0.111	0.013	0.29	0.19	0.467	0.555	0.202	0.675	1.223
4019	0.096	0.102	0.353	0.071	0.138	0.021	0.009	0.128	0.203	0.323	0.009	0.009	0.009
4219	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
4233	0.144	0.194	0.085	0.062	0.067	0.01	0.01	0.01	0.041	0.067	0.01	0.01	0.009
4263	0.289	0.551	0.157	0.575	0.107	0.133	0.401	0.084	0.069	0.076	0.026	0	0
4310	0.092	0	0.054	0.044	0.004	0.046	0.008	0.038	0.045	0	0.045	0.04	0.003
4376	0	0	0	0	0	0	0	0	0	0	0	0	0
4410	0.013	0.013	0.013	0.021	0.536	0.109	0.032	0.072	0.366	0.074	0.075	0.2	0.203
4465	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.113	0.191	0.19

4.3) Tertiary Experiment

The last experiment utilizes artificial neural networks, ANN, to create estimations for electrical usage. This experiment was done by using the MatLab platform. The concepts behind the experiments were briefly explained in the regarding Literature Review sub section of 2.2.1 The data for this experiment will be gathered from the previous experiment which will be explained in the inputs sub section below in more detail.

The results obtained from this experiment will be used to compare the estimation error results and to come to a conclusion to observe whether using a neural network trained with certain inputs or the estimated outputs would provide lesser errors. In a similar manner with the previous experiments, first the inputs will be explained followed by how the input data was processed to obtain results.

4.3.1) Inputs

The inputs that were used by this experiment are; the air temperature of Exeter of the test period which was obtained for the purposes of the last experiment, the room temperature of the experimental room which was also obtained during the last experiment through the implementation of a 4-in-1 Z-Wave sensor in the experimental room, the measured kWh data of the experimental room which will be used for two variables and the reason will be explained in the analysis process for this experiment.

Although the benchmark data from the previous experiments consists of 96 different values composed from the 30 minute intervals, for this experiment only the first 95 values can be used.

This is because estimations cannot be made from the last benchmark data since then next output would be out of the time frame of the benchmark.

Table 4.7: A Fragment from the 5x95 Input Matrix

23.55	24.12	24.68	24.98	25.27	25.26	25.25	25.3
27.3	27.3	27.3	27.3	27.3	27.3	27.3	27.3
0.00195	0.00193	0.0019	0.0019	0.0019	0.0019	0.0019	0.00193
12	12	13	13	14	14	15	15
0	30	0	30	0	30	0	30

4.3.2) Analysis Process

The inputs were processed through the additional toolbox of MatLab, *nntool*, which provides a GUI for the training and simulating for artificial neural networks with ease as seen in Figure 4.18.

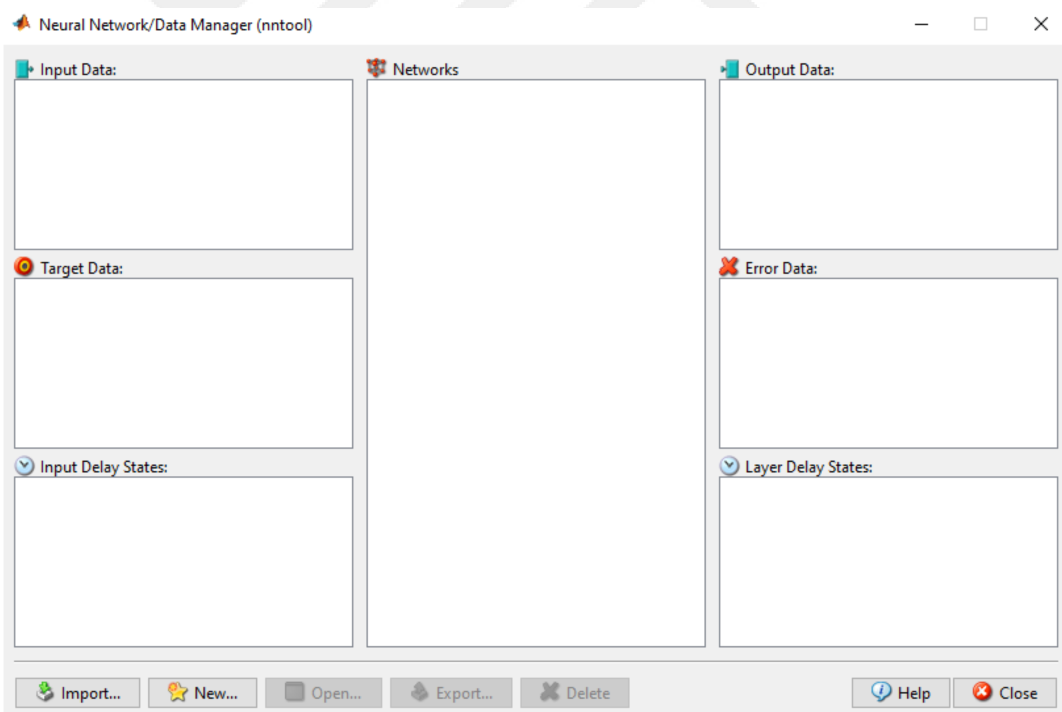

Figure 4.18: GUI of *nntool*

The inputs were introduced in double matrices as well as the output which is the target data. In this experiment the input values formed a 5x95 matrix and the output was a 1x95 matrix. If the dimension variable of the matrices do not match the training of the network can't proceed.

The depiction of the created network by the GUI can be seen in Figure 4.19.

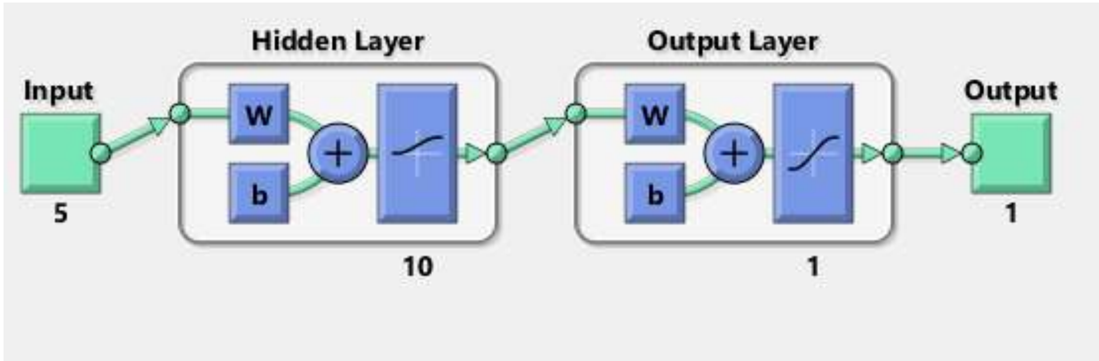

Figure 4.19: Depiction of the Created Network

From Figure 4.19, it can be observed that there are two layers with different functions that are applied to the input. The first layer, *Hidden Layer*, applies LOGSIG transfer function to the input and this process is followed by the application of *Output Layer* transfer function, TANSIG. For this experiment, training function TRAINLM was used, which applies Levenberg-Marquardt optimization and is the fastest compared to other functions of the toolbox. (Uk.mathworks.com, n.d.) This optimization was chosen because it was designated to be the best option in an another energy management application of artificial neural network. (Kowell et al., 2019) In addition, the adaption learning function LEARNGD was used.

The network type is configured as feed forward backdrop. In addition, 10 and 100 neurons were used and trained and the results are as follows. In addition to the variety of neuron numbers two different gradient values were tried during the training of the neural network. The quality of the training is measured by MSE, which is mean square error of the errors generated from the output of the training. Ideally, this value should be as close to 0, so training results with lower MSE mean closer output values to the expected output. In this experiment the training goal was chosen as 0.001. The training tests that failed to meet the performance quota are discarded.

The best outcome from the four cases, 10 neurons 10e-100 minimum gradient, 10 neurons 10e-1000 minimum gradient, 100 neurons 10e-10 minimum gradient and 100 neurons 10e-1000 minimum gradient, was chosen as the output that will be used. The best result was obtained with 100 neurons and 10e-1000 gradient values since the training with these parameters provided the lowest estimation errors compared to the other 3 cases.

Table 4.8: A Fragment of Outputs and Errors for the Four Cases

estimated output		0.00193	0.0019	0.0019	0.0019	0.0019	0.0019
10neurons with 10e-100 gradients							
	outputs	0.004374	0.0018004	0.0031	0.001833	0.0023	0.001842
	errors	-0.002444	9.96E-05	-0.0012	6.72E-05	-4E-04	5.77E-05
10neurons with 10e-1000 gradients							
	outputs	0.028941	0.0368787	0.03284	0.019203	0.0225	0.010011
	errors	-0.027011	-0.034979	-0.0309	-0.0173	-0.021	-0.00811
100neurons with 10e-100 gradients							
	outputs	0.003603	0.0036027	0.00446	0.003392	0.0043	0.00311
	errors	-0.001673	-0.003409	-0.0026	-0.00149	-0.002	-0.00121
100neurons with 10e-1000 gradients							
	outputs	0.002358	0.0016059	0.00226	0.001597	0.0021	0.001595
	errors	-0.000428	0.0002941	-0.0004	0.000303	-2E-04	0.000305

5) Discussions, Conclusions and Future Research

5.1) Discussions

In this sub section, the results obtained from the three experiments will be analysed and compared, and the implications of the results will be discussed in detail. By reaching a conclusion after the comparison of the results from the experiments, by utilizing the error values, the final step of the research flowchart will be completed as seen in Figure 3.1.

The best solutions provided from the second experiment and the best solution of the third experiment will now be compared, in order to find which solution should be used for a future framework.

For comparison, the minimums of the errors for each of the best solutions obtained from the second experiment were found. Later on, these results were compared with the error results from the best solution from the third experiment. The comparisons were done in Excel by using MINIMUM and AVERAGE functions.

Table 5.1 A Fragment of Error Comparison

0.00095	2E-05	0.0011	0.002	6E-05	0.00105	0.002	0.001	0.002	0.00159	0.002	0.002	0.002
0.000428	0.00043	0.0004	4E-04	4E-04	0.00043	4E-04	4E-04	4E-04	0.00043	4E-04	4E-04	4E-04
0.000428	2E-05	0.0004	4E-04	6E-05	0.00043	4E-04	4E-04	4E-04	0.00043	4E-04	4E-04	4E-04

Overall in only 16 cases the solutions from the second experiment were able provide results with smaller errors out of the 95 time frames.

The time periods when the ID solutions proved better were;

12:30, 13:30, 15:00, 16:30, 18:00, for 26th of August,

5:30, 7:30, 16:00, 19:00, 19:30,21:30 for 27th of August,

0:00, 4:00, 7:00, 8:00 and 9:30 for 28th of August,

However, just looking at single minimum error values is not enough. Overall, the ID solution mean errors are far greater than the errors seen from the neural network result as seen in the table below.

Table 5.2: Comparison of Mean ID Errors and ANN Errors

ID Solution	0.343626	0.20949	0.164552	0.1455	0.122136	0.151126	0.135204	0.16065	0.1865
ANN Solution	0.000428	0.00043	0.000428	0.0004	0.000428	0.000428	0.000428	0.00043	0.0004

When the minimums of the average errors of the ID solutions are compared with the errors of the neural network, the minimum errors are always from the results of the third experiment.

From this it can be concluded that in general utilization of artificial neural networks provide better estimation results compared to the developed ID solution method for this research. With regards to the decision making at the last stage of the flowchart, ANN method was chosen, based on the comparison demonstrated above.

It should be also considered that the network models developed for this research purposes were not trained multiple times. Therefore, their iteration numbers were low. For even better output with more minimalized errors, the network can be trained more as well.

5.2) Conclusions

As the demand for more energy and resources increases in the developed world and the developing countries, generating and implementing optimization based solutions will become a necessity. One branch of such solutions will be applied in smart homes of the future. For this purpose, the goal of creating a basis for a conceptual framework based on previous electrical usage and their behaviour was chosen for this research.

In order to develop such a framework, an appropriate methodology was adopted and the course of action was designated by a flowchart observable in Figure 3.1. For the duration and the completion of the research this flowchart was used as the modus operandi.

First, by analysing existing data across thousands of households and over a year long period, enough consumption data was processed to generate estimations. These estimations were done by considering the effect of temperature on the household usage of electricity. Furthermore, the effects of ongoing crisis' such as the climate change was taken into consideration and the estimation process was enriched. The results of these estimations were compared with the measured data obtained through an experimental room. The room was set up so that, through the implementation of smart home concept and necessary sensors, the usage of electricity for household items were monitored for a time period that created a benchmark model. The existing and processed data of the various households were adapted to the time frame of the benchmark, and the respective estimation errors were found for each household. Through the analysis of these errors, the best IDs were chosen as solutions with regards to the similarity of the behavioural patterns of their respective errors and the measured electrical usage of the experimental room to be used as the behavioural models that can be used for future modelling for the experimental room.

For the final step of the research purposes, an artificial neural network was also implemented to generate estimations, which was trained with the information of air temperature, room temperature, hour and time, the measured consumption data as input as well as the output of the system, where the measured data of the next interval was designated as the target data.

Through analysing the output and error values from the training tests with different parameters the model with the minimum errors was chosen. Later on, the results of the artificial neural network model were compared with the ID solutions' results that were chosen in the previous step.

It was found out that using ANN provides results with lesser errors compared to the results obtained from the ID solution. Thus, it was decided that for the purposes of future estimations and framework building for new buildings, artificial neural networks should be applied instead of the ID solution for better results.

5.3) Future Research

In this sub section, other possible research and project ideas that can be looked into and can be implemented in the future with the technologies that were used for this research's purposes will be discussed as well as the other technologies discussed in the Literature Review.

With regards to the Z-Wave home automation systems or smart home applications in general, there are various future works that can be realized as viable projects. Firstly, with software applications such as Domoticz, that enables direct control over electronic components, that are Z-wave compatible inside one's household, individuals can monitor and customize their own household and device usage. Although such technologies already exist and many software counterparts of Domoticz are available on the market, such as Home-Assistant.io or ioBroker, it is still relatively new, considering that home automation is a relatively young concept and has potential for more innovation.

Besides personal usage, the behavioural patterns of usage from individual households can be collected and studied for commercial benefit for companies as well. By gathering and analysing the patterns in to a Machine Learning algorithm, a system could be built to operate based on the learned data. One possible commercial product could be a device regulating heating output of industrial or domestic buildings.

An alternative to the data that was used for this research can be considered as well for the same goal of developing a framework. Since, the benchmark that was used for this research was set during the summer time, an alternative time period can be chosen in the winter time. As Cruz Rios states (2017), and as it was depicted in Figure 4.10 the consumption of electricity increases with the increase of temperature due to the cooling effect, and also increases with decreasing temperatures with the heating effect.

Since the experimental room that was setup for the purposes of this research was a student accommodation room, the sockets from which the electrical consumption data could be measured and monitored were limited. If the same setup could be applied to a regular home, where more sensors could take measurements from more devices, a more appropriate model of an individual's electric consumption behaviour could be monitored and can be used as a benchmark.

In addition to the correlations between electric usage and certain time periods, similar correlations and possible estimations can be made between water usage and a number of different variables as well. One such possibility could be the estimations regarding water usage during summer time versus winter time in individual households. For such a research project, electronic components such as LinkTap may be considered, as it is also configurable with Z-wave and can be easily integrated into a smart home system. Furthermore, as it was displayed in Figure 4.7 and Figure 4.8, illumination and temperature data can be gathered and analysed to be used in such research topics as well. Furthermore, such studies can be expanded by utilizing the aforementioned variables in order to develop estimation algorithms by using Machine Learning. In that sense, similar to the first data set of this research, with enough data collection this may be achieved in a future study.

However, it should also be noted that home automation systems are not infallible when it comes to cyber security and privacy. With regards to Domoticz, there have been various remarks on the official forum online, about attacks and disruptions caused in the home automation system and to the electronics that are actively operating and connected to the system, from foreign IP's on individual's home systems. Attacks like these can result set off alarms, manipulate the lights and so on. (Domoticz.com, 2016) Considering the loss of privacy that may arise from such attacks and from the knowledge that usage is monitored by home automation, it can be said that the recorded personal usage logs of data can be spied on by malware, in a way similar to cookies that are encountered when browsing the Internet. For an example, by observing the electricity data of the lamps and the temperature sensors connected to the smart home product, intruders may understand when you are awake or perhaps out of the house. Similarly, these attacks and the evident vulnerability of Domoticz also imply that other available products on the market could be susceptible to such attacks as well.

In order to counter and prevent loss of privacy and security, firewall scripts could be implemented on the platform that Domoticz or the counterpart software is set up, or if possible products with existing security measures for home automation systems.

Besides the possibility of spyware or individual monitoring of a household or possible estimation algorithms, these technologies can be implemented on a far greater scale as well. These could be smart cities and smart manufactories which are also within the scope of the Industrial Revolution 4.0. A mass implementation of mesh networks with sensors could be useful for detecting the patterns in a manufactory plants in order to make estimations for the future costs that are spent on energy costs.

In addition, since in the second experiment conducted for the research, just as sensors were considered the agents for the simulation, a home could be considered an agent for a smart city implementations regarding automation and reduction of costs. Or in a similar sense, a scenario could be considered where sub plants in a manufacturing system could act as agents.

As well as proving beneficial to reduce production and energy costs in a manufacturing plant, these technologies can also further progress the rate of complete mechanization of factories which will result in the lights out phenomenon, by the ever increasing rate of communication and connection between the devices as seen in Figure 2.6. (Essentracomponents.com, n.d.)

6) **References**

12CAD.com. (2018). *CAD/CAM in Industry 4.0*. [online] Available at: https://www.computeraideddesignguide.com/cad-cam-industry-4-0/

Balatsky, A., Balatsky, G. and Borysov, S. 2015. Resource Demand Growth and Sustainability Due to Increased World Consumption. *Sustainability* 7(3), pp. 3430–3440. Available at: http://dx.doi.org/10.3390/su7033430

BBC Bitesize. (2019). *Reasons for increase in demand for energy - Revision 4 - Higher Geography - BBC Bitesize*. [online] Available at: https://www.bbc.com/bitesize/guides/zpmmmp3/revision/4

Business News Daily. (2019). *Industry 4.0: How Technology is Revolutionizing the Manufacturing Industry*. [online] Available at: https://www.businessnewsdaily.com/10156-industry-manufacturing-iot.html.

Cruz Rios, F., Naganathan, H., Chong, W., Lee, S. and Alves, A. (2017). Analyzing the Impact of Outside Temperature on Energy Consumption and Production Patterns in High-Performance Research Buildings in Arizona. *Journal of Architectural Engineering*, 23(3), p.C4017002.

Cybersecurity for Industry 4.0 Cybersecurity implications for government, industry and homeland security. (2018). [PDF] Available at: https://www.ey.com/Publication/vwLUAssets/ey-cybersecurity-for-industry-4-0/$File/ey-cybersecurity-for-industry-4-0.pdf

Darby, S. (2017). Smart technology in the home: time for more clarity. *Building Research & Information*, 46(1), pp.140-147.

Domoticz.com. (2016). *IoT and security: Shodan - Domoticz*. [online] Available at: https://www.domoticz.com/forum/viewtopic.php?t=11575

Ec.europa.eu. (2019). *Energy consumption and use by households*. [online] Available at: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20190620-1

Eisenstein, P. (2013). *Overhauling industry: Ford celebrates moving assembly line centennial*. [online] NBC News. Available at: https://www.nbcnews.com/businessmain/overhauling-industry-ford-celebrates-moving-assembly-line-centennial-8C11339779

Eliwhitney.org. (2019). *The Cotton Gin | The Eli Whitney Museum and Workshop*. [online] Available at: https://www.eliwhitney.org/museum/eli-whitney/cotton-gin

Ervural, Beyzanur & Ervural, Bilal. (2018). Overview of Cyber Security in the Industry 4.0 Era. 10.1007/978-3-319-57870-5_16. Available at : https://www.researchgate.net/publication/319861803_Overview_of_Cyber_Security_in_the_Industry_40_Era

Essentracomponents.com. (n.d.). *What exactly is lights-out manufacturing? | Global Manufacturer & Distributor of Component Solutions — Essentra Components*. [online] Available at: https://www.essentracomponents.com/en-gb/news/product-resources/what-exactly-is-lights-out-manufacturing

Festo.com. (n.d.). *Industry 4.0 | Festo*. [online] Available at: https://www.festo.com/cms/nl-be_be/56644.htm

Fritz, R. (2019). *Home Automation Mesh Networks Explained*. [online] Lifewire. Available at: https://www.lifewire.com/mesh-networks-for-home-automation-817660

Greenfield, D. (2016). *Siemens Positions Itself as an Industry 4.0 Example | Automation World*. [online] Automationworld.com. Available at: https://www.automationworld.com/siemens-positions-itself-industry-40-example

Hofmann, E. and Rüsch, M. (2017). Industry 4.0 and the current status as well as future prospects on logistics. *Computers in Industry*, 89, pp.23-34.

Horst, Diogo & Duvoisin, Charles & Vieira, Rogerio. (2018). Additive Manufacturing at Industry 4.0: a Review. International Journal of Engineering and Technical Research. 8. 3-8. Available at : https://www.researchgate.net/publication/327701729_Additive_Manufacturing_at_Industry_40_a_Review

Jackson, R. (2019). *The Effects of Climate Change*. [online] Climate Change: Vital Signs of the Planet. Available at: https://climate.nasa.gov/effects/ [

Kamarul Bahrin, M., Othman, M., Nor Azli, N. and Talib, M. (2016). INDUSTRY 4.0: A REVIEW ON INDUSTRIAL AUTOMATION AND ROBOTIC. *Jurnal Teknologi*, 78(6-13).

Kowell, S., Wicaksono, H., Yuce, B., McGlinn, K. and Rezgui, Y. (2019). User Centered Neuro-Fuzzy Energy Management Through Semantic-Based Optimization. *IEEE TRANSACTIONS ON CYBERNETICS*, [online] 49(9), pp.3278-3298. Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8412214&tag=1

Lamkin, P. (2019). *Z-Wave explained: What is Z-Wave and why is it important for your smart home?*. [online] The Ambient. Available at: https://www.the-ambient.com/guides/zwave-z-wave-smart-home-guide-281

Macal, C. and North, M. (2005). TUTORIAL ON AGENT-BASED MODELING AND SIMULATION. [online] Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1574234

Manufacturing-operations-management.com. (n.d.). *Industrial-Revolution-Industry-4-timeline*. [online] Available at: https://www.manufacturing-operations-management.com/.a/6a00e551030252883401b8d138b7c7970c-popup

Masood, T. and Egger, J. (2019). Augmented reality in support of Industry 4.0—Implementation challenges and success factors. *Robotics and Computer-Integrated Manufacturing*, 58, pp.181-195.

McGrath, M. (2019). *UK's 10 warmest years all occurred since 2002*. [online] BBC News. Available at: https://www.bbc.co.uk/news/science-environment-49167797

Met.ie. (n.d.). *Historical Data - Met Éireann - The Irish Meteorological Service*. [online] Available at: https://www.met.ie/climate/available-data/historical-data Exeter_2009 https://www.worldweatheronline.com/exeter-weather-history/devon/gb.aspx

meteoblue. (n.d., a). *Weather history download Cork*. [online] Available at: https://www.meteoblue.com/en/products/historyplus/download/cork_ireland_2965140
meteoblue. (n.d., b). *Weather history download Exeter*. [online] Available at: https://www.meteoblue.com/en/products/historyplus/download/exeter_united-kingdom_2649808

Metlink | Weather & Climate Teaching Resources. (n.d.). *How will the UK's Annual Temperature Change? - Metlink | Weather & Climate Teaching Resources*. [online] Available at: https://www.metlink.org/climate/will-uks-annual-temperature-change/

Museum of Obsolete Media. (n.d.). *CD-ROM (1985 –) | Museum of Obsolete Media*. [online] Available at: https://obsoletemedia.org/cd-rom/

Newworldencyclopedia.org. (2019). *Industrial Revolution - New World Encyclopedia*. [online] Available at: http://www.newworldencyclopedia.org/entry/Industrial_Revolution

Niiler, E. (2019). *How the Second Industrial Revolution Changed Americans' Lives*. [online] History.com. Available at: https://www.history.com/news/second-industrial-revolution-advances

Olson-Raymer, G. (n.d.). *Political & Social - Manifest Destiny*. [online] Gorhistory.com. Available at: http://gorhistory.com/hist110/unit3/political.html

Reiher, T. and Koch, R. (n.d.). *PRODUCT OPTIMIZATION WITH AND FOR ADDITIVE MANUFACTURING*. [PDF] Paderborn: C.I.K./DMRC, University of Paderborn, pp.1092-1103. Available at: https://pdfs.semanticscholar.org/d826/ee370c7698f458abc7aed131d5a0c24b1dfd.pdf

Research-Methodology. (n.d., a). *Research Philosophy - Research-Methodology*. [online] Available at: https://research-methodology.net/research-philosophy/pragmatism-research-philosophy/

Research-Methodology. (n.d., b). *Research Approach - Research-Methodology*. [online] Available at: https://research-methodology.net/research-methodology/research-approach/

Research-Methodology. (n.d., c). *Data Collection Methods - Research-Methodology*. [online] Available at: https://research-methodology.net/research-methods/data-collection/

Research-Methodology. (n.d., d). *Qualitative Data Analysis - Research-Methodology*. [online] Available at: https://research-methodology.net/research-methods/data-analysis/qualitative-data-analysis/

Rojko, A. (2017). Industry 4.0 Concept: Background and Overview. *International Journal of Interactive Mobile Technologies (iJIM)*, 11(5), p.77.

Ross, I. (1998). The Invention of the Transistor. *PROCEEDINGS OF THE IEEE*, [online] 86(1), pp.7-28. Available at: https://pdfs.semanticscholar.org/1bb8/72a1f4d27b457621c44ed937ac42cf37097a.pdf

Rüssmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P. and Harnisch, M. (2015). *Industry 4.0 Future of Productivityadn Growth in Manufacturing Industries*. [PDF] Available at: https://www.zvw.de/media.media.72e472fb-1698-4a15-8858-344351c8902f.original.pdf

Sanayidegelecek.com. (n.d.). *History – Future of Industry*. [online] Available at: https://www.sanayidegelecek.com/en/sanayi-4-0/tarihsel-gelisim/ .

Smith, R. (1998). *ModelBenders - Encyclopedia*. [online] Modelbenders.com. Available at: http://www.modelbenders.com/encyclopedia/encyclopedia.html

Uk.mathworks.com. (n.d.). *Levenberg-Marquardt backpropagation - MATLAB trainlm-MathWorks United Kingdom*. [online] Available at: https://uk.mathworks.com/help/deeplearning/ref/trainlm.html

Vaidya, S., Ambad, P. and Bhosle, S. (2018). Industry 4.0 – A Glimpse. *Procedia Manufacturing*, 20, pp.233-238. doi: https://doi.org/10.1016/j.promfg.2018.02.034

WorldWeatherOnline.com. (n.d.). *Exeter Historical Weather*. [online] Available at: https://www.worldweatheronline.com/exeter-weather-history/devon/gb.aspx

Zhong, R., Xu, X., Klotz, E. and Newman, S. (2017). Intelligent Manufacturing in the Context of Industry 4.0: A Review. *Engineering*, 3(5), pp.616-630.

Appendix

Form 2 - Project Proposal (ECMM164) Student Name: Umut Makascioglu

MSc Project Proposal
ECMM164

Title: Advanced intelligent system applications towards Industry4.0 framework

Student Name: Umut Makascioglu
Programme: Postgraduate Taught
Student number: 680054221
Candidate number: 031198
Student email: um214@exeter.ac.uk
MSc Programme Name:
MSc Engineering Business Management

Supervisor: Dr. Baris Yuce

Project proposal approval			
I confirm that I have read the MSc project proposal and the student has incorporated my suggested changes to my satisfaction. The proposal has well defined aim, reasonably detailed objectives, methodology and work plan.			
Supervisor Signature:	Baris Yuce		
Supervisor Name:	***Dr. Baris Yuce***	***Date:***	*06/03/2019*

(It is students' responsibility to get the project proposal approved from their supervisors and submit a signed copy of the front page to a.m.macgregor@exeter.ac.uk by 8th March 2019)