GATHERING ELECTRIC DATA FOR IMPROVED EFFICIENCY IN POWER STATIONS

By Igal Kushnir

A DISSERTATION

Presented to the Department of
Information Technology
program at Selinus University

Faculty of Computer Science
in fulfillment of the requirements
for the degree of **Doctor of Philosophy**
in Information Technology

2021

CHAPTER ONE
INTRODUCTION

1.1.The Problem Statement

Today, the power plants are increasingly facing a challenge of meeting the ever-increasing demand of energy consumption in the world. These plants have complex systems that are characterized by reliability and maintenance policy. The degradation of the efficiency of the plant has been common in such systems and thus calling for efforts to enhance sustainable availability and reliability of energy (Melani et al., 2019). The most prevalent problems association with this situation has been the increased costs of generating electricity and a rise in the environmental effect. These issues have compelled most of the power plants to focus on continuous developments with the aim of improving their energy efficiency. This step has been necessary because increase in efficiency can tremendously bring about the reduction of the overall costs of electricity generation and the environmental effect caused by the power plants. There is a need to have a model can help in the analysis of the electric generation in the power plants, which must derive from the data that is gathered on electricity. Therefore, this thesis seeks to propose methods that can be used collect the data that can help in improving efficiency in power plants as a solution to the identified problems.

The shift to a new energy production and efficiency is crucial for Australia it intends to remain competitive in the global and regional energy industry. Currently, the existing power plants in Australia supply the total megawatt capacity that to cater for the peak demand in winter and summer in all the states, which forms the National Energy Market (NEM) and

Western Australia (Stewart, 2017). Australia's national transition plan has offered a vivid direction for the power plants in seeking to satisfy the high demand of energy consumption. The nation's government policy intervention to address the problem of energy demand has been stipulated in the national transition plan of Australia. However, the decision that governments, investors, and policymakers are encountering is when to replace such aged plants and the kind of technology that they need to use instead of the old one. The traditional technologies together with the plants of electricity production have implications exceeding the sector of electricity supply in Australia, and this issue could negatively affect the national energy productivity by dragging it. Since there lacks modest forecasted growth in Australia's electricity demand, no new generators have entered the market. This situation implied that the existing power plants need to be made efficient to serve the current energy demand of the Australians since no new plants are being commissioned.

1.2. Background of the Study

Power plants play a crucial in the production of the energy that meet the market demands in alignment with the energy regulatory requirements. Energy generation refers to the total amount primary energy generated in the country's economy, whose measurement is taken before the transformation or consumption. The discussions on the on global energy demand have now shifted from energy prices and legal requirements to the creation of energy efficiency in the power plants (Abele et al., 2015). In the recent years, there have been efforts made to maximize the entire process of electricity generation or single elements as electrical drives. Despite these efforts, there has been a significant potential for the measures of energy efficiency that are economically beneficial in power generating plants. Power companies around the world have started seeing the need to collect reliable and compressive data on different aspects of electricity to help in monitoring the efficiency and seek to improve this crucial parameter.

In a nation point of view, Australia is one of the world's region with power plants that have focused on efficient power production to meet the ever-increasing energy demand. The production of electricity in this country as part of the industry of energy supply has turned to be known by old power stations and aging technology. The accepted traditional fossil fuel fired power plant's economic life span is the age of about 40 years (Stewart, 2017). In Australia, it is estimated that around three quarters of the country's coal-fired power stations are working past their initial design life, while some have had refits of life extension. Most of the largest capacity power stations in the nation get power from brown and black coal, which shows the historical reliance of Australia on fossil fuels for the generation of electricity. The commissioning of many of these power plants was done in the 1970s and the 1980s. In effect, a large portion of these power plants are coming to an end of their commercial lives, with some of them already working extra miles ahead. Removing these plants again from the grid in fast succession would present the potential to Australia to see the capacity of roughly 7,654.3 MW being removed (Steward, 2017). However, the more recently commissioned power plants lack the capability of generating the same MW capacity as opposed to the older power plants.

In Australia, there are numerous wind and other renewable power plant projects aimed at improving the efficiency of the power generation in years to come, fueled by the scheme of Renewable Energy Target (RET). Fossil fuel power plants with large capacity of megawatts have also been proposed and additional capacity for the existing plants has been approved (Steward, 2017). It is, however, challenging to know if such projects are going to continue because the existing conditions of energy demand in the nation, together with the concerns of the community about the addition of further emissions from coal power plants. Thus, it is necessary for the policymakers to realize that a new approach would help the power plants make improvements in their efficiency of power production and distribution to the consumers in Australia. Moreover, any decision concerning the replacement of the old plants or technologies must be made in line with the international commitments of Australia to the reductions of emissions.

The Australian government developed a new standard of energy efficiency, known as generator efficiency standard (GES) to ensure that power plants operated at the desired efficiencies. The function of GES is to regulate greenhouse emissions from power stations, which covers all kinds of power plants that operate on diverse types of fossil fuels such as oil, coal, and gas (Mahlia et al., 2018). Audit or reviews are done at least one time in every five years to see to it that the desired performance in power plant efficiency is attained. Non-compliance power plants that take part in these standards attracts a penalty. The standard of emission efficiency is utilized in evaluating the efficiency of power stations in GES. The implementation of this standard aims at saving four mega tonnes of CO2, which would require AU$10 to save one tonne of CO2 given out as argued by Mahlia et al. (2018). In general, the execution of the standards of power plant efficiency has been one means of reducing the environmental effect.

In a local perspective, Victoria state, Australia has top renewable power station. The various sources of energy in this region include solar, wind, bioenergy, hydroelectricity, geothermal, and marine-wave and tidal resources. Hydroelectric power constitutes the biggest percentage, accounting for 56 percent, and wind is the second one with 34 percent of electricity production, followed by biomass (8 percent), and geothermal, which as accounts for 1.3 percent (Victoria, n.d.). Despite Victoria having many power plants producing electricity from the world-class renewable resources, it has several additional factors that determine their efficiency of their operations. One of the Such factors comprise access to as well as the capacity of the electricity grid. This factor indicates the need for enhanced efficiency of the power plants to ensure optimal efficiency and increased capacity and access to electricity by all the people in Victoria. The power plants in this state modelling their operations of electricity generation using the latest technologies of renewable energy under the expanded target of renewable energy. The schemes employed must consider the data the needs to be collected to help in the improvements of the efficiencies of such systems.

Data gathering is among the core prerequisites for production facilities of efficient energy in Australia. The new data collection technologies enable further analyses by the assessment of KPIs, condition or energy efficiency measures (Abele et al., 2015). This thesis provides a standardized method of collecting electricity relevant information or data of the machines of production and their parts on the PLCs. The collection of electricity data can help in direct monitoring of different parameters of the power generation system, such as energy efficiency, for standard interfaces, where the data is saved on a generic data server to monitor the power plant (Abele et al., 2015). Therefore, there is a need for the implementation of the electric data gathering technique on the electricity generation machines in power plants to help in monitoring the efficiency for future improvement.

1.3.The Aim and Objectives of the Study

The overall aim of this thesis is to investigate how to gather data on electricity to help in enhancing efficiency in power plants. The specific objectives that this study has to attain include:

1. To determine the different types of electric data that can be collected in the power plant.
2. To investigate the different methods that can be used to collect the electricity information to help in the improvement of the power plant system.
3. The recommend the best electric data collection technique for enhancing the efficiency in the power plant.

1.4.The Research Questions

a. What are the different types of electric data that can be collected in the power plant?
b. What are the different methods that can be used to collect the electricity information to help in the improvement of the power plant system?
c. Which electric data collection technique is considered to be the best for improving the efficiency in the power plant?

1.5.Research Approach

This study utilized descriptive research approach. The system review of the literature was employed discussing the several methods of electric data gathering in power plants for improved efficiency. This design concentrated on finding the connection between variables based on the study of Bryman and Bell (2011). In this study, the variables include the type of electric data and data collection system as independent variables and efficiency as a dependent variable. A descriptive design helped in obtaining the information that describes existing situation in the power plants regarding the different approaches used to collect the electric data. Creswell (2014) asserts that a descriptive study design offers facts concerning the nature and the status of a situation and how it takes place at the time of research. This research design was chosen based on the research problems in connection with the theoretical framework for this research. Therefore, this research was suitable since the intention was to collect detailed information concerning the different methods or systems of gathering electric data in power plants through descriptions that enabled the identification of the variables. The scientific databases for other literature reviews on the internet about the research topic were used in this study. Researcher was able to get the articles that attract his attention into specific measures of energy efficiency that could lead to the generation of data for this study. The available literature on the electric data collection methods and systems for enhanced energy efficiency in power plants was reviewed. However, the researcher got other reviews that offered data various topics, which failed cover power plants but in areas like residential buildings and others.

1.6. Study Significance

This thesis is significant because it provides insights on how to collect electricity data that can help power plants improve their energy efficiencies. It is important since the results can be extrapolated and replicated to other power plants, with the inferences made regarding efficiency of energy in power plants. The research findings and analysis can also be used by the power plant system designers and developers to identify which components need to be improved to obtain optimal energy efficiency. The paper can also be used by the researchers in their future studies as a source of information concerning the best method of gathering information on electricity for the enhancement of efficiency in power generation facilities.

1.7. Key Assumptions

There are some assumptions that the researcher considered in this thesis. The first assumption that was made in this study is that all the power plants operate on the same principles. The other assumption is that the selected power plant complies with the regulatory requirements of electricity generation. It was also assumed that the methods of electricity gathering do not increase the costs of energy and have no environmental impact.

1.8. Organization of the Thesis

This thesis has been organized in five main chapters. The first chapter has provided the statement of the problem that the study seeks to address, the background of the research, study aim and objectives, research questions, significance of the study, and the key assumptions that the researcher made during the study. In the second chapter, which is the literature review, the researcher has provided the theoretical and critical reviews of the past studies conducted on the subject of gathering electricity data to improve efficiency in power plants, while identifying the gaps in such studies. The third chapter entails the research methodology, which describes the research design, time and location where the study was carried out, the target population, sampling and sample size, the data collection methods and procedure, the data analysis and presentation, validity and reliability, and the ethical research considerations. The fourth chapter presents the results obtained from the research. In the fifth chapter, the researcher discusses the findings of the study, connecting it with the literature available on the topic of research. The paper ends with the conclusion, where the key findings are summarized, and recommendations for further research are given.

1.9. Limitations of the Study

This study had several limitations that the researcher needed the researcher to take into account. The first limitation was that the study relied on secondary data from the studies that were published in the past. The systematic review of the literature utilized in this research is only reliable because of the approaches utilized in the estimation of the impact in every primary studies. This issue implies that performing a systematic review could not address the issues that were prevalent in the research design as well as the implementation of the primary research as argued by Garg et al. (2008). The second limitation was that the research method used for this study could not rectify biases because of selective publication. This situation encouraged the studies that reported dramatic impacts to have higher chances of being spotted, summarized, and systematically reviewed as opposed to other publications that presented smaller sizes of the effects, leading to publication bias (Palma & Delgado-Rodriguez, 2005). The process of the review also focused on identifying the publications that are varied in their design, specific interventions, the participants used in the studies, and quality of the methodology. However, there seemed to be some subjectivity when the researcher was making a decision on the way related studies had to be prior to the review being deemed appropriate. It is likely that pooling poor quality studies together with those that were well-researched in the review would result in the worst findings of the study (Garg et al., 2008). This situation may end up making the readers of this study doubt the reliability and validity of the results obtained in this research. It is, therefore, needful for further study to be conducted taking into account these issues and using a study approach that is appropriate.

CHAPTER TWO:
LITERATURE REVIEW

2.1.Introduction

This chapter offers a theoretical and a critical review of the literature available on electric data gathering for improving efficiency in power plants. This section has been divided into four main subsections. The first part is the theoretical review, which focuses on different theories and concepts that have contributed to the exploration electric data gathering in power plants. In the second section, the author reviews the different literature available on the methods of sensing and collecting data. The third part of this chapter provides the critical review of the existing literature on the topic being examined.

2.2.Theoretical Review

2.2.1. The Concept of Efficiency

Energy efficiency plays a fundamental role in the reduction of the demand in the consumption of energy. The improvement in energy efficiency has been fueled by consumers who have made choices about the energy to use in their interests, instead of political leadership of a particular country. Engineers, especially in Australia, believe that the political concerns on elevated bills of energy focus on satisfying their personal interests in the power plants, neglecting the potential of energy efficiency for the reduction of such bills (Stewart, 2017). The choices of consumers have made substantial contributions to electricity consumption reduction because consumers are increasingly looking for products and appliances that are energy efficient. This condition is also believed to have been fueled by Commonwealth and the energy programs of state household such as ceiling insulation, home energy rating systems, and solar hot water. Steward (2017) holds that this trend can continue in decades to come since the behaviors and preferences of consumers keeps changing. Despite there being an expected increase in the electrical appliance use in the households due to the growth in the population, there is less energy being consumed because the energy efficiency of such products also keeps improving. Relatively less electricity is being used by mobile devices with Wi-Fi connections, which are gradually replacing conventional households that have stationary entertainment systems and home computers (Steward, 2017). The efficiency of energy is primarily attained at business and household user ends, though it also has ramifications for the management of the network. Engineers are increasingly employing the latest digital technologies to optimize the flows between electricity producing plants and users. Since some parts of electricity production appear to be monopolistic, especially in distribution and transmission, there is a necessity to have reforms that regulate such areas to accomplish a program of energy efficiency.

A host of studies of have tried to explore the efficiency concept, seeking to define it and discussion how it is measured. Forsstrom et al. (2011) view the efficiency of energy as a word applied in various meanings in diverse contexts. The authors continue to argue that there exists no single explicit quantitative measure of the efficiency of energy for all scenarios. They maintain that the measurement of the conversions of energy efficiency always occurs by taking energy output divided by the energy input of the process of energy conversion. It is further noted that in processes, it may be difficult to measure the output as energy, making the conversion process appear more complicated (Forsstrom et al., 2011). In such circumstances, it is needful to devise a measure for the output that can properly describe the need that bring about the demand for the energy, service, consumption, products, or process. This situation results in a demand indicator, which is used in place of energy input to determine the efficiency. Based on this explanation, energy efficiency is the ratio of the energy input to the demand indicator.

Some reviews on have connected the concept of energy efficiency to other areas other than power plants. For instance, Kaynakli (2012) reviewed the existing literature concerning the determination of the maximum thickness of the material of thermal insulation of a residential building together with its impact on the demand of energy. Moreover, Stevanović (2013) reviewed the available literature on various passive strategies of solar design, though the study focused on commercial and residential buildings. These studies have directed their efforts on examining the optimization of the efficiency of energy in residential buildings. Such studies have sought to develop diverse types of quantitative measures that can be used to provide a solution of energy efficiency improvement problem. Furthermore, the studies have investigated different areas of application where optimization of energy efficiency is required to help in addressing the issue effectively. The findings of these studies suggest the need for the gathering of the electricity data to be done at different fields of energy usage to help in developing solutions for the improvement of the efficiency of energy in power plants.

2.2.2. Measurement of the Efficiency of Energy

Measuring the efficiency of energy may not be a complicated a task as indicated in the prior research. According to Forsstrom et al. (2011), the measurement of energy efficiency greatly depends on the selection of indicators for output and input in the energy conversion process. In real sense, these two parameters can be determined in various ways, and thus selecting one method over the other always results in tradeoffs. The available literature presents four key categories of indicators for the measurement of the energy input and output used in the measurement of energy efficiency. The first group of indicators has been highlighted as thermodynamic indicators, which greatly depends on the thermodynamic quantities for both outputs and inputs. An example of such as an indicator is the thermal efficiency of a system of heating (Forsstrom et al., 2011). Economic signals form another group that is used in measuring both energy outputs and energy inputs in terms of money, for instance of energy expenditure per GDP unit. The measurement of the two parameters also relies on economic-thermodynamic aspects, which entails determining the delivered services or products (output) based on the market prices and energy in terms of thermodynamic units, for instance energy intensity of the GDP. Another indicator is physical-thermodynamic one, which comprises the measurement of energy inputs in terms thermodynamics with the outputs having physical units, such as the consumption of energy of a building for every square meter.

The existing literature on the measurement of energy efficiency suggests the improvement of efficiency being a crucial part of the power generating plant. Forsstrom et al. (2011) argue that the improvement of the efficiency of energy may not have an intrinsic worth, but it is a way of gaining other broader objectives of the power plant. The authors continue claiming that it is essential to consider other perspectives when measuring the efficiency of energy. One of the crucial points of view is eco-efficiency, which is described as the ecological efficiency whose aim is to determine how natural resources are used and the negative effects based on the data generated (Forsstrom et al.. 2011). This view relates to the concept of sustainability, which entails economic, social (institutional and cultural), environmental or ecological sustainability. For instance, the ecological sustainability is turning out to be more challenging for the firms that target sustainability. Thus, the information obtained on eco-efficiency can be utilized in determining the ecological sustainability of the power plant. Therefore, the efficiency of energy efficiency can be measured based on the information gained on eco-efficiency, particularly where non-renewable sources of energy are involved.

The measurement of energy efficiency also derives from the material efficiency perspective. The measurement of the efficiency of materials can be based on the fact that the drop in the consumption of energy can lead to an increase material consumption or the vice versa is also true. The concept of materials efficiency mostly relates to the energy embodied in materials of power plant structures. In power generating plants, the materials used to make the turbines may have an effect on the efficiency of energy being generated. Sritram et al. (2015) performed an experiment to determine the effect of different turbine materials (aluminum and steel) on the efficiency of power generation. The authors found that the efficiency of energy generation was higher for aluminum turbines as compared to steel turbines. The results imply that light materials are the most preferred for making turbines for the production of electricity for optimum efficiency. Researchers made this conclusion based on the data they collected on the speed of the turbine, values of power and torque, and total efficiency (Sritram et al., 2015). Therefore, the engineers designing the turbines used in production of electricity need to take into account the materials that can offer greater energy efficiency for the power plant.

2.3.Smart Grid

Smart grid (SG) presents a crucial area in the power plants that require the applications of data analytics. The past studies on mining electrical data have based on the issues of SG, involving power generation, distribution and transmission, and consumption. An extensive research has been conducted on the applications of big data in the planning of power generation, performance and efficiency of power storage systems, cost minimization and optimization of power grid, and economic load dispatch.

2.3.1. Management of Power Production

The literature indicates that the management of power generation is among the crucial application fields of big data analytics in SGs. The analytics of big data is applied in the planning and operation of power generation to aid the processes of decision-making (Ghorbanian et al., 2019). There are different ways in which power is generated. Some of the sources of energy include hydroelectric power plants, solar, wind, and coal, among others. These power plants have always sought to increase the efficiency in power generation to meet the ever-rising energy demand across the world (Zhang, 2020). In most studies, the authors have discussed the various techniques of improving the efficiency of thermal power production in the contemporary power plants. Zhang (2020) identifies some of these methods as raising the initial parameters of steam, lowering the final parameters of steam, use of combined cycle power generator, and water pre-heating. This article maintains that the improvement of the efficiency of a power plant can be through either conservation of energy within the system or expanding differences between starting and final parameters of steam. Overall, these parameters can only be measured if data collection methods are put in place by the power plants to help track the efficiency of the power within the system.

2.3.2. Transmission and Distribution of Power

The study also reveals the transmission and distribution of energy is a vital component of the power generation system. Kayastha et al. (2014) argue that the collection and sensing of data can be utilized in the distribution and transmission in monitoring transformers, underground lines, overhead lines, and substations. The authors hold that the use of data sensors to monitor real-time data about the current, voltage, voltage, frequency, and phase can result in the increased capacity of transmission up to between 10 and 15 percent as opposed to the models of capacity planning that employ static weather, temperature, and wind scenarios (Kayastha et al., 2014). The authors continue pointing out that sensors play a fundamental role reducing the loss of power transmission by offering automatic distribution capacitor working according to the transformer requirements. The information obtained from sensors can be used to compute thermal and mechanical efficiency of the power distribution and transmission. However, this article does not provide clear evidence if the obtained data from the sensor can be calculated to get the figures of energy efficiency in the power plant to warrant the need for improvement.

At the stage of power transmission and distribution, the monitoring, control, management, and protection of the power systems present further issues related to smart grid. In the SGs, the AMI represents the primary source of data (Ghorbanian et al., 2019). The other sources of data include PMUs, smart meters, and sensors, among other IEDs, which are known as control signals. Such data be utilized for diverse purposes, including improvement of power system efficiency, enhancement of customer service, predictive analytics, and fault diagnosis, outages prevention, real-time assessment of vulnerability, improvement of security and reliability, and safety. Monitoring of the power systems comprise the events as well as the state estimation depending on the smart devices and AMI, sensing of islanding, detection of oscillation, real-time monitoring of rotor angle, asset management, and extensive area situational awareness. All these aspects require the analytics of big data in SGs. Thus, big data analytics is needed and is used in diverse areas of power systems. Therefore, there is a need for an in-depth insight into the potential the big data analytics applications, which can offer more opportunities for the implementation and promotion of the SGs that are aided by big data in the offing.

The operation of distribution has also attracted interests among the researchers depending on the data that can be collected to assists in the improvement of efficiency. According to Kayastha et al. (2014), the generation and processing of a large amount of data in the SG can lead to the creation of a bottleneck by architecture of conventional centralized communication in the system. As an illustration, a PMU, which is form of a smart sensor for measuring the electrical waves on the power grid, can produce up to 50/60 phasor measurements in every second with the ability to easily yield big amounts of data within a short time (Kayastha et al., 2014). The transmission of all these data in the phasor in a centralized network leads to the creation of a congestion, and thus, requiring a control mechanism congestion. Another challenge of data collection arises from the sensing nodes that used in different locations of the power grid, where scalability is the greatest issue. This problem indicates a necessity for decentralized aggregation methods for the measurement of the sensor as well as the approaches of distributed computation for parameters of the actuator. Overall, there is a need for the distribution of the network of communication network with numerous controlling and processing units for the prevention of failure at a single point.

2.3.3. Power Consumption

The data on power consumption has also drawn interests among the researchers in the past studies. Kayastha et al. (2014) assert that data sensing in the premises of consumers brings about many opportunities to both public utility and consumers for maintaining a desired response of energy demand in the power grid, even in the cause peak hours. Electricity data collection in the premises of consumers can be carried out using smart meters. These gadgets offer data concerning real-time usage of electricity to help in the automatic control of electricity consumption. Researchers further maintain that the smart meter plays a role of a sensor node, whose work is to record the consumption of power in kilo watt hour (kWh) and usage time or time of use (Kayastha et al., 2014). Time of use has been defined as the activity of consumers at a certain time that is usually determined after every hour, day, or week according to the customer requirement. The real-time utilization of data helps the public utility to comprehend the real-time patterns and demand of energy consumption. The article highlights that such information can be utilized in increasing the efficiency of power generation by enabling better balancing of load or diversion of the flow from neighboring distributed energy resources (DERs). Therefore, the data on power consumption has proved to be necessary for the power plants whenever they need to make any improvements in the efficiency of the power generation system.

2.3.4. Storage of Power

The data on energy storage also constitutes a vital part of the SG for the efficient storage energy from energy sources, such as renewable energy sources. The output of these varies depending on the conditions of weather, for instance, solar panels produce energy only when the sun is shining. The literature proposes lead-acid battery as the ideal means of storing energy (Kayastha et al., 2014). This system requires the measurement and monitoring of temperature, current, and voltage. The most crucial parameter that requires to be measured, collected, and reported is state-of-charge (SOC). The measurement of the SOC is from the level and density of the acid. Such parameters change in the event the processes of charging and discharging takes place. The component of great importance in the storage system is the small ion exchange material, which builds a sensor to measure the concentration in a lead-acid battery. Another component that is required in this system is the temperature sensor for monitoring the operating point of concentration sensor and the battery’s reliability. The power storage can also be utilized in the vehicle, especially in a plug-in hybrid vehicle. Thus, there is a need to monitor and report the battery’s SOC in a vehicle to help the optimization of its efficiency. In general, the energy storage in power plants and other systems that use energy is essential since the data collected from them can assist in the decision-making processes concerning the improvement of the efficiency of power and other aspects.

2.4.State of Data Gathering

The collection of data has been cited in numerous studies to be an essential part of the power production. The power plants have numerous processes that produce information that can be used to make some decisions concerning the production of power. Most of the power plants depend on manual process of data collection to obtain the required information for the process of human decision-making. Such information constitutes the data source gathered manually by tasks that require extensive labor. Al Rashdan and Germain (2018) argue that the process of data collection is costly, though it is limited in applications. The power industry has, thus, realized the need to replace these labor-intensive tasks with automated systems of data collection that can help in deciding to make any changes within the power plants for optimum efficiency. The authors maintain that the collection, storage, architecture, communication, protection, and computational power of data form a section of the infrastructure of information in a power plant, which needs to be planned well in the efforts of data management. It has also been suggested that the best way of presenting the data for supported human decision-making in power plants is to include this process in the efforts of data visualization (Al Rashdan & Germain, 2018). The data and their pattern obtained can be interpreted for the decision-making concerning any improvements required, which has to be planned via numerous efforts in data analytics in years to come.

The prevailing state of data collection in the power plants has also drawn the attention of researchers in this area. In the current state of gathering data, most the utilities that have been utilized are known as “base state” (Al Rashdan & Germain, 2018). The existing attainable condition has been termed as the “modern state”, implying a condition that has either been attained by some utilities or other sectors. There is also another state of the future called the “state of the art” condition, which needs R&D (research and development) efforts to attain the desired data for sound decision-making. These states indicate that there is a potential for the power plants to automate the data gathering processes for the improvements of various aspects, including the efficiency of energy. In most of the studies, the findings indicated that manual and labor-demanding processes have been used in the power plants in the base state (Al Rashdan & Germain, 2018). The modern state of data collection is focusing on the reduction of manual processes using electronic together with semi-intelligent solutions or integrating sensors or fixed technologies that can assist in the reduction of labor demand in the data collection process. In the state-of-the-art era, it is anticipated that there will be more improved technologies for data sensing, collection, and monitoring for the improvements of efficiency in power plants. Such advancements entail improved techniques of data gathering, correlation of data, and machine learning activities. As a result, there will be increased fidelity and spatial distribution of data and merging of numerous sources of data, which will, in turn, lead to the utilization of sophisticated mobile gadgets, sensors, and machine comprehension techniques. The available literature on the data gathering and sensing techniques have targeted the methods of sensing data on one side and the data communication approaches on the other side.

2.4.1. Emerging Methods of Data Sensing

The previous studies have investigated several techniques of data sensing that aid in the gathering of data for various purposes. Kayastha et al. (2014) have studies three techniques under this category, which include units of phasor measurement, compressive sensing, and cooperative and decentralized data detecting. These techniques are as discussed below:

Compressive Sensing

Compressive sensing is another approach that that attracted the attention of researchers to determine how it relates to the methods of data collection in power plants. In the communication networks of the smart grid, the sampling requirement as well as the transmitting information from numerous sensors at the same time would pose a big challenge on the current sensor networks' state-of-the-art (Kayastha et al., 2014). Compressive sensing (CS) is a promising technique that can be used to overcome this challenge. This approach is a combination of the concepts of compression, data acquisition, optimization, and reduction of dimensionality. CS works under the principle of a K-sparse indicator $\mathbf{x}$ of N dimension being recovered from merely a few measurements that are not complete, $\mathbf{y} = \mathbf{Ax}$ of M dimension through an optimization of L1-minimization for matrix $\mathbf{A}$ (random projection) (Kayastha et al., 2014). This method takes fewer measurements as compared to traditional sensing techniques and does not require any extra compression. CS offers an effective method to manage the greater bandwidth of data and content of information.

Units of Phasor Measurement

The phasor measurement units, also known as synchrophasors, are used for measuring the electrical waves using a common source of time for synchronizing process. A phasor refers to a complex number standing for both the phase angle and the magnitude of the sinusoidal curves (sine waves). The measurements of the phasor that takes place simultaneously are known as synchrophasors, which can be supported by PMU (phasor measurement unit) devices. The sampling of PMUs in many applications is done from locations that are widely dispersed in the network of power system, which are then synchronized by the use of the common source of time of a clock of GPS radio (Kayastha et al., 2014). The technology of synchrophasor is useful for system planners and operators in measuring the state of the electrical system and in the management of the quality of power. The function of synchrophasors is to measure currents and voltages (which are essential in the calculation of power efficiency) at different areas on a power grid. The synchronization of these phasors makes it possible to draw synchronized comparison of two parameters or quantities in the actual time. As a result, such comparisons can help in the assessment of the conditions of the system.

Cooperative and Decentralized Data Sensing

The method of cooperative and decentralized data sensing has also been explored in some past studies. According to Kayastha et al. (2014), deregulations in electricity industry has resulted in the creation of numerous regional organizations of transmission within an extensive interlinked power system. The article maintains that since there are significantly many smart elements occurring in the power grid, there is a substantial need for more processing, communication, and computational resources. In effect, there is a growing necessity for more processing and control of distributed information for the operations of power systems. The study suggests a simple, though effective method for data decentralized and cooperative data detection, which is decomposition-and-merge (Kayastha et al., 2014). This approach is a two-degree hierarchical one, where lower level, every local area independently operates on its own estimator based on the measurements obtained locally. At the highest level, the results of the state estimation are received by the central coordinator from each area, alongside the boundary conditions when required, which are then merged to attain a solution for the entire power system.

2.4.2. Data Communication Approaches

Once data has been collected, it is necessary to communicate to the relevant personnel or department within a plant to help in decision-making. The discussion, in this part, focuses on diverse technologies used for communicating data in the power plant's smart grid. The section also presents a hierarchical architecture of the network that can be used in data communication and the protocols of networking for the AMI. The subsection also provides a review of the different requirements data communication in the smart grid. The challenges encountered in the previous studies concerning the use of smart grid technologies in data collection and sensing in power plants also form some good part of the discussion. Finally, the section discusses the technologies of middleware for the smart grid. The methods of sensor communications have attracted research attention in the field of power production.

Some of the techniques under this category include cooperative communications, cognitive ratio, and machine-to-machine communications, as discussed below:

Cooperative Communications

The collection of the data in power plants also derives from cooperative communications. Kayastha et al. (2014) define cooperative communications as the methods wherein numerous nodes are used to assist one another, especially in ad-hoc, sensor, and wireless mesh networks, for relaying or forwarding the packets of data to their endpoints.

One of the types of cooperative communications is WSN, which is applied to offer the transmission of data in urban scale environment of smart grid. Collaborative WSN has several ZigBee nodes that operate cooperatively in monitoring electrical substations in an urban environment (Kayastha et al., 2014). In this kind of network, there is one coordinator for controlling the transmission of data of all nodes to the network's concentrator. This technique involves the introduction of a reliable and secure scheme of cooperative communication scheme for metering infrastructure (AMI). It also has a multihop wireless network whose work is to link smart meters with AMI for transferring meter data to a collector in local networks. The method ensures data privacy, integrity, and trust services depending on mutual authentication. The development of these aspects occurs in three core processes, which inc

lude initializing, collection of meter-readings, and message distribution management. The process of initialization is carried out when a new smart meter makes a request to enter the AMI. After the smart meters joining the AMI, the process of data gathering is undertaken. The data generated from the meter is used together with the distribution process of management message for the adjustment of the production of electricity power, distribution, and transmission.

Cognitive Radio

Cognitive radio or CR methods have also been explored in numerous studies in the context of data measurement and collection in SGs. CR can be used to raise the efficiency and utilization of spectrum of radio frequency in wireless networks (Kayastha et al., 2014). CR can also lead to the spectrum assigned to the licensed users, as the primary users, to be dynamically and opportunistically accessed by those users who are not licensed (secondary users). In this technique, the implementation of intelligent algorithms is required by the unlicensed users to shun interruptions to those who are licensed, and thus enhancing the performance of transmission, and optimizing the utilization of radio resource. This approach can also be utilized for reliable communications of smart grid in cost-effectively. Most studies propose dual-radio architecture for CR in the applications of smart grid for the transmission of real-time data (Kayastha et al., 2014). Therefore, the data generated using these methods can help in the improvement of the performance of power plants, especially in energy efficiency because of the implementation of the system, tolerance of fault, and wide area of coverage.

Machine-to-machine Communications

The available literature also studies machine-to-machine communications as another crucial method of sensing communication in the context of power production. The machine-to-machine communication concept is among autonomous gadgets that has become a great supporter of pervasive communications. The most used technology is M2M communication, which has been utilized in the collection and transmission real-time data from different gadgets with no need for human interventions as argued by Kayastha et al. (2014). In the smart grid, this approach can be applied to attain a robust, efficient, and hierarchical solution of networking for diverse gadgets with significantly low capabilities of processing. M2M communication utilizes wired, hybrid, and wireless communications technologies. The literature also highlights the use of M2M communications in the HEMS (Kayastha et al., 2014). One of the network architectures is greater significance is power systems focuses on gathering of the data from the smart meter. The data sensed and collected from such meters is transmitted from houses to the smart grid’s concentrator, and then transferred to the base station of wireless area network (WAN). At the base station, the gathered data from the meter was transferred to the center of control. However, researchers point out that there has been an issue of maximum traffic concentration in the power grid that needs to be addressed (Kayastha et al., 2014). Moreover, the algorithm of optimal cluster formation can be applied to get the optimal position of traffic concentrator in an ad-hoc network.

2.4.3. Architecture of Communication Network

The literature available on smart grid present the architecture of the communication network as a crucial component when examining the methods data collection in power plants. The infrastructures of data communication network in the smart grid comprise AMI (advanced metering infrastructure), covering various systems. According to Kayastha et al. (2014), AMI plays a critical role in offering a system of automated communications between a utility company, as a control center, and smart meter. It can handle all the aspects of communications of the smart grid, providing a user-centric means of managing the usage of electricity in the power grid. Some smart grid functions can be implemented on AMI, such as the detection of consumer outage, measurement of customer voltage, reading of hourly remote meter, programming of remote meter, price signaling, and load control (Kayastha et al., 2014). The AMI has DAUs for the collection and relaying of the information from the side of the meter to MDMS. The work of the MDMS is to offer processing, storage, and management of meter data to ensure proper usage by applications and services of other power system. Moreover, a DAU links a premise of the customer to a network entity occurring at a higher rank in the power grid, for instance, the MDMS. In this case, the entity of data control entity including MDMS and DAU play a role of a data sink for gathering sensed information as well as a gateway for relaying the data from one network hierarch to the other. The systems and subsystems of AMI include sensor and actuator networks (SANETs), wide area measurement systems (WAMSs), NANs, WANs, and HANs.

SANETs

SANETS are formed by several actuators and sensors that have been installed in the power grid. SANET helps in monitoring of the operational behavior and characteristic of the smart grid gadgets that are distributed along the grid for the prevention of any disturbance or outage and provide effective DSM. The formation of SANET can be by use of a hierarchical topology, such as cluster tree topology, or a cooperative sensor network, or a linear topology (Kayastha et al., 2014). As an illustration, sensor network can be applied in the implementation of remote system monitoring, equipment fault diagnosis, and wireless automatic meter reading. Casey et al. (2011) introduces the design as well as the execution of a network of cross-platform sensor for monitoring of transmission line. Frolec and Husak (2010) present the sensor networks for monitoring the temperature of the conductor and overhead transmission line current. The monitored measurement can be utilized in optimizing the efficiency of power transmission and maximization of the capacity of amperes of the overhead line, fulfilling the requirement of reliability under weather conditions that are not certain. Li et al. (2010) proposes the link scheduling algorithm for the WSN using tree topology. This approach can be for monitoring of the transmission line. The study has particularly proposed the multichannel allocation algorithm for the network of multimode hierarchical type. The first step is to form the network in the hierarchical structure. The next step is preforming the channel assignment for the implementation of data forwarding based on the assigned channel. These steps aim at improving the throughput of the network. The ability of setting the actuator parameters based on the values observed using the sensors to attain a system-wide objective is of great significance in SANETs. Further, the installation of SANETs assists in the facilitation of direct link between the consumer and the public utility (Kayastha et al., 2014). The technologies like PLC, the Internet, BPL, WiMAX, and 3G/4G cellular technologies of wireless communication can be a suitable choice for meeting this interaction. The SANETs must optimized, in this case, for the optimization of the unique characteristics of communication, which may include severe environment of smart grid because of electromagnetic interference. The requirements of performance, which comprise limited resources of memory, time-sensitive latency requirements, power supply, and processing speed, as well as the need of variable QoS are great importance too.

WAMSs

WAMS has been known for offering real-time monitoring and control of the electrical power grid to prevent any future contingency. In power plants, the WAMS can be utilized for connecting PMUs with dynamic coverage and broad area. WAMS has one core advantage as compared to the conventional SCADA and EMS in the offering of only local monitoring and control is the greater coverage. It can also transfer sampled data of the data. The data of this nature is utilized in supporting several functions of the smart grid, which comprise instability prediction, state estimation, dynamic data logging and disturbance recording, real-time monitoring, damping or oscillation analysis, and power system dynamics analysis. A typical infrastructure of WAMS has multiple PMUs, mainly sensors, which are linked to a PDC (phasor data concentrator) via a regional network. The system of this nature is called the PMUs-PDC operating group. The role of a control center is to gather and aggregate phasor data from numerous working groups of PMUs-PDC via a WAN, such as synchronous digital hierarchy or synchronous optical network. WAMS is more preferred over SCADA since it can monitor the state of power system synchronously and nonstop. It also supports a high rate of sampling data, which makes WAMS the best choice for the failure recovery and dynamic event management.

HANs

The available literature on smart grid presents a HAN as the smallest subsystem in the response chain to energy demand of the power grid. It offers dedicated DSM with the utilization of smart meters (Ota, 2011). A HAN is made up of smart actuators and sensors for measuring different parameters such has the intensity of light and temperature; a smart meter for sensing and recording the usage of electricity, a HEMS for the optimization of the consumption of energy based on real-time information of pricing, and in-home display for the user interface (Kayastha et al., 2014). Moreover, the smart meter serves as a gateway to help the public utility in accessing and retrieving the data on electricity consumption for the response to energy demand. The most used wired technologies in HANs include PLC, KNX protocol, Wi-Fi, Z-Wave, ZigBee, and building automation and control network (BACnet). In general, SANETs in the premises of the customer form fundamental solutions for this subsystem.

NANs

Several studies have proposed the use of NANs in the literature on power grid systems. Kayastha et al. (2014) define a NAN as a collection of numerous HANs for the collection of sensed data to help in aggregation roles. In this system, all the data from diverse HANs are gathered at the control entity of data, that is, DAU. In this architecture, the related control center monitors the amount of electricity that is distributed to a certain neighborhood to offer effective response to energy demand. The control entity of data interacts with the smart meter using network technologies like fiber-optics, PLC, WiMAX, or satellite.

WANs

This architecture helps in linking numerous systems of distribution together. It is made up of various sensors, such as PMU, line temperature sensor, and voltage sag sensors, which cover nearly all the aspects of the lines of distribution and transmission, offering control and monitoring roles in the event there is outage or fault. Such systems of distribution can also comprise control entities of data, such as MDMS, which serve as hubs for gathering all the sensory data (Kayastha et al., 2014). The analysis of the gathered data is then done by a particular distributed control center that develops control algorithms. This infrastructure offers the backhaul communication for the connection of all the entities within the power grid. Several technologies of communication networking, including fiber-optic, PLC, cellular wireless, and many more can be utilized in linking all the functional subsystems in the WAN of the smart grid.

2.4.4. Protocols of Networking

Various networking protocols have been studied in numerous articles published on smart grid. Kayastha et al. (2014) suggest that it is essential for any network of communication for the smart grid to have outstanding features such as being secured, self-healing, efficient, scalable, and interoperable to be fit for a vast range of applications. The authors continue to point out that such a network needs to support the capabilities of open standard and plug-and-play ability. The IP-based network has been proposed as the effective solution to use since it is the backbone of the communications architecture of the smart grid (Kayastha et al., 2014). The previous studies have established the numerous merits associated with IP as compared to other current, making it wider applications in data networks. The authors continue to argue that most network technologies possess a convergence layer for IP (Kayastha et al., 2014). An IP-based network helps in bringing about multicast traffic together with the security that is reliable within the power transmission and distribution lines. This technology allows the integration of nearly all the network technology into an IP-based network. Researchers have identified IPv6 protocol as a suitable option as opposed to its predecessor IPv4 due to the new features it has, such as bigger address space, security services, and mobility (Kayastha et al., 2014). Nonetheless, the convergence with IPv6 cannot apply for all the existing network technologies. Thus, a combination of IPv4 and IPv6 would result in a more flexible approach. This greatest challenge with this approach is that it calls for proper translation of the boundaries of the network that could serve as the starting transition for the change to singe network of IPv6. Despite the old IP-based network being versatile, it only supports the delivery of the best-effort packet. This idea implies, it fails to bring about the QoS for networking of data in the smart grid. Nevertheless, the applications of the smart grid need some QoS guarantee in relation to the aspects of delay variation or jitter, latency, packet loss, as well as bandwidth for reliable and efficient working. Moreover, the networking of data needs to be characterized by proper protection of security in the smart grid.

The literature on smart grid indicates that QoS-based services can be incorporated into the networks of IP to make sure that there is quality data in the power systems. According to Ting et al. (2010), the Internet Engineering Task Force (IETF) has come up with numerous service models, including differentiated service or DiffServ, MPLS (multiprotocol label switching), as well as integrated services (InterServ) to fulfil IP QoS. MPLS and DiffServ are used for the provision of QoS of IP. Kayastha et al. (2014) present DiffServ as a computer networking infrastructure that can be utilized in providing QoS guarantee and low-latency for crucial network traffic, and at the same time offering simple best-effort service to applications that are not critical, which include file transfer or web traffic applications. The implementation of DiffServ can be done in IPv4 and IPv6 to offer the desired support of QoS.

Moreover, MPLS is viewed as a standard technology to speed up flow of network traffic that incorporates the information in Layer 2 concerning the connections of network, that is, latency and bandwidth, into IP (Layer 3) within a certain autonomous system for improving and simplifying QoS of the exchange of IP packet. Both these technologies are promising and can be used to provide necessary QoS support for smart grid. Ting et al. (2010) propose a means of providing a guarantee of QoS in IEC (International Electrotechnical Commission) automation protocol of 61850 substation using IPv6 DiffServ model. The integration of DiffServ and MPLS can also be utilized to fulfil the requirements of QoS of data that needs to be transmitted and distributed in the power grid.

The transmission control protocol (TCP) can be considered over IP as the protocol of transport layer for WAN in the power grid. The issue with TCP is that it incurs a substantial delay in the transmission of data and high overhead of network (Kayastha et al., 2014). TCP also has no mechanism for supporting real-time traffic, which makes it not to be appropriate for services in the power grid. The limitation is that it has no multicast ability, implying sending one packet to many destinations. This characteristic can be vital when transmitting many control signals to multiple systems and devices in the smart grid in case of emergencies. In the applications of this nature, the previous studies recommend the use of the protocols like real-time transport protocol that run over the UDP (user datagram protocol). Kayastha et al. (2014) describe UDP as a simple protocol of transmission generally utilized for the applications that are sensitive to time since it operates with no implicit retransmission and acknowledgment of messages. This protocol is made up of a smaller overhead in a message structure. Even though it also supports multicast, it has no mechanism for anchoring reliable delivery of packets in a smart grid. A combination of UDP and TCP, therefore, can be utilized to offer reliable and ordered packet service in smart grid.

Reliable delivery of data in a sensor networking environment, for instance in a HAN, remains to be an open research, and no standard transport layer protocol exists for that matter. Researchers express their notion of the inability of an end-to-end transport protocol like TCP to be used for wireless SANETs in the power grid (Kayastha et al., 2014). The problem with this protocol is that it has a mechanism of congestion control whose initial design was meant for wired networks and has extremely low rate of data of the sensors of the smart grid, reasons of energy efficiency, and elevated header overhead. The incorporation of HANs and SANETs in the smart grid requires the transport protocols to be light weight to provide support to low-end actuators and sensors. The authors have also proposed another protocol, known as the pump-slowly fetch quickly (PSFQ) protocol (Kayastha et al., 2014). This protocol has been developed for WSNs to offer reliable transmission in poor radio environment when applying a light weight signaling and a recovery mechanism of hop-by-hop error. PSFQ entails a retransmission of node requests from NANs after the detection of the packet loss. Another transport protocol for WSNs is a reliable multisegment transport. This protocol employs both end-to-end and hop-by-hop error recovery to ensure reliability in the transmission of data. The literature presents another protocol as event to-sink reliable transport, designed to attain reliable detection of events from every node of the sensor and congestion control, and at the same time reducing the consumption of energy.

In the smart grid, the transport layer protocol’s performance in the SANETs relies on the underlying MAC and network layer protocols. The previous research proposes a vast range of data-centric routing protocols for the WSNs can be utilized in the dissemination of data from the nodes of the sensor to the sink node in an energy efficient way, leading to the maximization of the network's lifetime. The different classes of sensor routing protocols include hierarchical routing, such as LEACH (Low-Energy Adaptive Clustering Hierarchy); location-based routing, such as GAF (geographic adaptive fidelity) protocols, and flat routing, comprising SPIN (sensor Protocols for Information via negotiation). Furthermore, the classification of protocols can be based on the operation of the protocols, giving the categories that include reactive, hybrid, and proactive protocols. Moreover, depending on the channel access as well as the mechanisms of collision resolution, the MAC various categories of protocols for traditional WSNs can be comprise contention-free protocols such as TRAMA, hybrid protocols like Z-MAC, and contention-based, such as S-MAC (Kayastha et al., 2014). Several protocols of this kind of MAC and routing can be employed in the smart grid environment for the detection and communication of data. Overall, the literature suggests the need to take into account the considerations of the traditional design, such as the constraints of energy and resources, and the problems like severe environmental conditions, scalability, adaptability, and stringent QoS requirements for the efficient and reliable MAC and routing protocols for the applications of smart grid.

2.4.5. Requirements of Data Collection

The data collection devise, smart grid, was made up of several systems and subsystems of different kinds with diverse requirements of data collection. The requirements for the power production and transmission data were different from the one for the premises of consumers (Kayastha et al., 2014). In this study, a large amount of data was collected, transformed, verified, and stored in almost real-time for decision support and intelligent response in the power plant's smart grid. The smart grid data sensing was scaled as expected to accommodate the desired levels of volume and complexity of data in the power generation system. As an illustration, a PMU (phasor measurement unit) was required to convey up to 30 reports per second to help in measuring power delivery quality (Kayastha et al., 2014). It was necessary to employ distributed computation methods for actuator parameters and decentralized aggregation methods for sensor measurement since the sensing nodes used in different locations of the power grid would cause challenges of scalability. The network's scalability was enhanced by proper organization of the sets of actuators and sensors into a graded topology through the formation of domains or groups. The domain controller was used to control these domains, which also helped in the aggregation of the detected data and setting the parameters of the actuator within the group. The other function of this gadget was to communicate with the domains that were higher in the category. Therefore, the general working of these actuators and sensors could be optimized using the top-level controller in the group. The data was collected at different locations of the power system as shown in Figure 1 below:

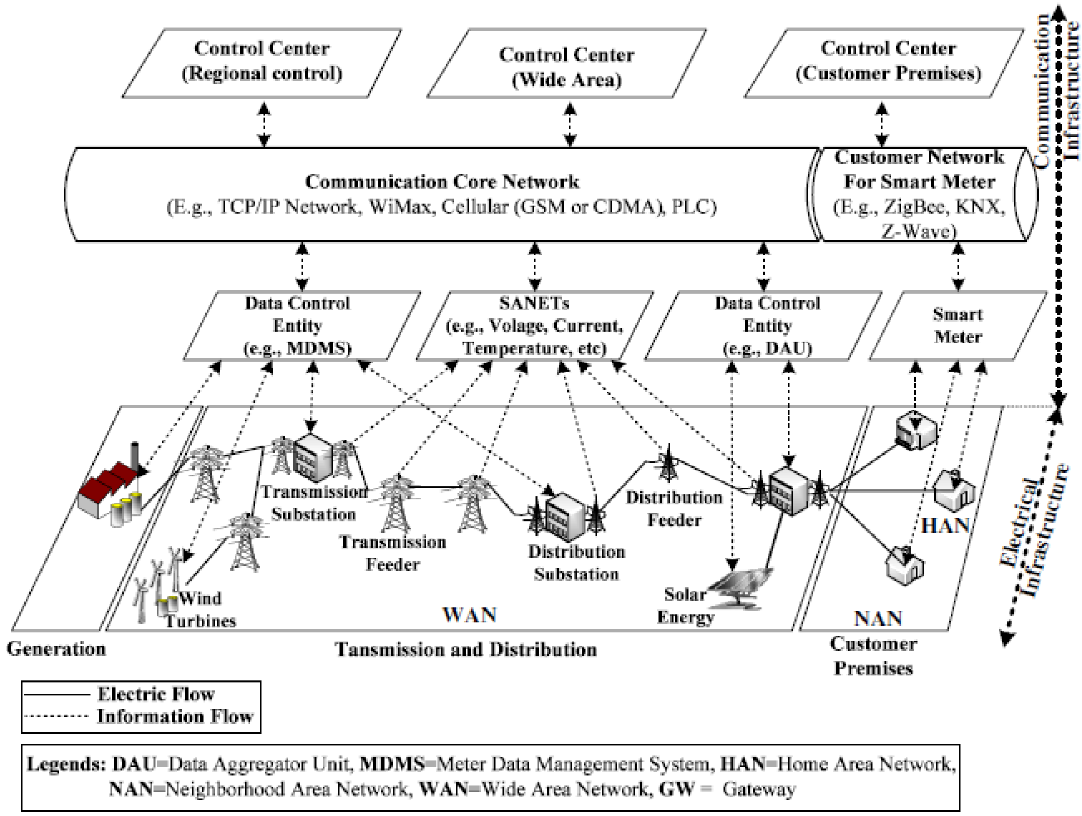

Figure 1: Data gathering in smart grid network

2.4.6. Challenges Related to Smart Grid

The several problems associated with smart grid in the collection and detection of data have need studied in the previous literature. Some of them include:

Bad data detection

There is a possibility of the attackers injecting bad measurement data which makes the smart grid vulnerable, therefore the need to detect and identify wrong connections, bad data resulting from drifts and large measurement bias in estimation of power system state.

Secure communications and data sensing.

Cyber threats and malicious attacks are some points of reference in the smart grid of data sensing and communications. The use of internet and cellular networks in public networks would lead to vulnerability and treats to many (Kayastha et al., 2014). Thus, a solution to securing the Smart grid is the use of private networks though the approach may not be economically efficient. Considerations of adopting hybrid approaches where smart grid is based on public networks like meter data collection whereas in the private network security sensitive applications like fault detection are can be adopted.

Service-focused networking:

In the smart grid, the design and implementation of sensing and communications of data can be based on a service-focused infrastructure. The resources including protocols, sensors, and actuators are viewed as a unified service that has a standard kind of interface (Kayastha et al., 2014). Such networking assists in reducing the system integration's complexity and encapsulation of the irrelevant information to the external entities such as a control center that does not require to know the physical features of a sensor. The challenge here is that there is a need to develop a service broker as a point where the consumer and the service supplier can exchange their services.

Cross-layer optimization of smart grid communication protocols

There is need to design robust communication protocols on the base of cross layer optimization which will satisfy the availability, reliability, and scalability requirements of data transmission in complex and lossy smart grid environments.

Cost optimization of network design

Performing a cost benefit analysis is needed before deploying network infrastructure for data sensing and communication to identify the risk and usefulness of technologies available. For an optimal design of data communication network, performance metrics should be analyzed for the development of data protocols and cost aware network infrastructure. For example, the use of CR technique could save cost for wireless bandwidth. However, it may result to delay and interference due to packet loss (Kayastha et al., 2014). Therefore, to minimize the total cost, a tradeoff of using dedicated wireless channel and CR must be investigated.

Improved device technologies

To improve the efficiency of energy and the devices lifetime, the performance effects of technologies and mechanism like power harvesting and low power consumption for actuator and sensor devices on networking and communication need to be analyzed.

Reliability analysis

There is need to adopt intrusion prevention methods and fault tolerant designs for improvement of communication and data sensing infrastructure. Investigation needs to be conducted on the component’s dependability used in communication and data sensing, considering the impact of different attacks caused by the adversaries.

Quality of service (QoS) framework

If the power supply is suboptimal the smart meter data is lost or the failure of power substation due to the delay in reporting a notification of cooling system fault. Therefore, QoS broker or QoS framework for communications and data sensing for the smart grid has to be developed for different smart grid data to be sensed and transferred without delay, loss and error. For different smart grid applications, the QoS metrics need to de defined accordingly to the service level agreement (Kayastha et al., 2014). Therefore, need for development of QoS service or QoS framework level agreement in the communication and sensing infrastructure. The mechanisms involve packet scheduling, error detection, multiple access, congestion control and recovery. Also, the need to perform optimization and analysis of the developed mechanisms.

Networking middleware

A significant problem of data in communication environment of heterogeneous smart grid is interoperability. This issue can be addressed by developing networking middleware to offer convergence among technologies that are heterogeneous (Kayastha et al., 2014). The networking middleware brings about an interface between communications protocols and applications, leading to the improvement of the system's flexibility and integration of network.

2.5.Networking Middleware

Some literature has been provided on the middleware of networking and the protocols of layer application. According to Kayastha et al. (2014), networking middleware provides an application interface of a greater level that boosts data communication's portability, flexibility, and interoperability. It also offers an interface between the requirements of a high degree application QoS together with management of low-level resource, which include the requirement of bandwidth and real-time monitoring of data. In general, middleware of networking can be applied in supporting the network protocol's interoperability to fulfil the smart grid's requirements.

A prime example of the networking middleware is GridStat. It is a simple publish–subscribe middleware applied in the acquisition of the status alerts and status variables in the power grid. The status variables refer to the periodic measurements of the settings of the control or status in the power grid. It makes it possible for such measurements to be managed, ensuring that it is updated. The status alerts, on the other hand, are utilized in notifying about the applications of any alarm condition or failures with high urgency. Moreover, the use of QoS brokers assists in the management of the assignment of resources and offers the required means of delivering messages with the desired requirements of QoS. Avista Utilities in Washington is implementing GridStat based on the SCADA system for acquiring the required data directly from the sensing devices (Kayastha et al., 2014). This middleware, thus, is among the old data gathering methods that have been used in power plants for the improvement of the plant's efficiency.

The Universal Plug and Play (UPnP) protocol™ is another type of middleware of networking used in the power grids. Kayastha et al. (2014) describe UPnP protocol as the gadget discovery protocol playing a role of supporting zero configuration, automatic discovery, and transparent networking, as well as controlling different types of home network gadgets. This protocol offers communication of data between two or more gadgets controlled by the point in the network or UPnP control device command. In this case, the idea is that the devices put to sleep state or shutdown whenever not in use for a particular time and woken up whenever there is need to do so (Kayastha et al., 2014). Every gadget that utilized the protocol of UPnP publicizes its power state and wake up status, among other general information occasionally or responds to point search messages of UPnP control. The work of the UPnP control point is to convey a multicast M-search message every moment it requires to access a certain gadget. The devices that active are accessed directly while the inactive ones or those in the sleep mode are activated using the proxy of power management. A test network shows that the average power consumption can be significantly minimized without compromising the usability by using the protocols of UPnP.

2.6.Energy Efficiency Indicators

The data collection can reveal the indicators of energy efficiency in the power plants. Forsstrom et al. (2011) have suggested a set of indicators with each indicator capturing various facets of energy efficiency better as compared to the others. The indicators studied seek to ensure that there is a balance between the diverse goals. One of such goals comprises to satisfactorily cover the diverse aspects of the efficiency of energy, which is vital to different stakeholders. The second one has been bringing about a precise description of the energy efficiency of the structures using energy, especially buildings. The other aim has been the ensure the indicators are applicable in the practical sense. Finally, the indicators covered seek to make sure that the required data is reasonably present to help in the determination of the efficiency of the power plants. The use of a blend of indicators may be needful if one must achieve the energy efficiency of structure, say a building. In general, the kind of indicators to apply rely on the objectives and situation of the analysis of the power plant.

2.6.1. Measurement of Consumption

The indicators revolving around the measurement of energy consumption revolve the quantity, *Q*. Q is described as the energy, in kWh, that is used every year or some other given time in the in a structure. *Q*, in practice, usually covers at least all the energy that is commercially delivered to the site, simply known as delivered energy (Forsstrom et al., 2011). There is a need to explicitly list what should and what should not be included in *Q* when using these kinds of indicators. Today, nearly all cases involve the flows of useful energy flows, which are inbound to the structure using the energy. Therefore, a vigorous analysis of energy usage needs to consider the energy that has been consumed as well as the one that has been saved in the all the activities at the consumer's end. At the present, common practice focusing on the consumption of energy by the buildings offer a sufficiently precise estimate of the building's total life-cycle energy usage since the significance of the other stages of the power system appears to be relatively small. Nevertheless, the energy efficiency of power plant can I prove as time advances, and thus it can be anticipated that the relative significance of life-cycle energy consumption rises to the place where ignoring it is not possible.

The available literature indicates that energy prices vary based on the different forms of energy. According to Forsstrom et al. (2011), carriers of high exergy energy have more uses as opposed to low exergy options, making them more valuable. Any flow of the delivered energy along the chain of energy is associated with the actual primary energy usage being numerous times higher as compared to the delivered energy usage. Thus, the energy values need to be considered, where with a factor p is applied to diverse flows of energy, such as $Q_{electricity}$, Q_{heat}, etc. In this case, the p value can indicate the exergetic, monetary, and primary energy, among other value of the different energy flows. It is essential to note that the energy flows having varying values should not added together before weighing them with the factor p. Therefore, the equation for Q is given as:

$$Q = P_1 Q_1 + P_2 Q_2 + \cdots \qquad (1)$$

In the above equation, the descriptions of the parameters used are as follows:

Q_1, Q_2, etc = The different flows of energy of diverse values, which comprise annualized embodied energies.

$P_1, P_2, etc.$ = the respective coefficient values.

2.6.2. Specific Energy Usage

The specific energy usage or consumption, or SEC is a formulated, general purpose, and easily interpreted energy efficiency measurement for buildings. It is determined by taking the energy delivered divided by the area under consideration in square metres. In this case, the area comprises either the building's heated area or the entire built area. The consideration of the heated area only has the advantage of giving a better description of the heating efficiency in the building. However, keeping certain areas of the building, including storage, with no heating is an energy efficient practice that leads to the measurement of energy efficiency. Currently, SEC is widely applied, which perhaps helps in comparing efficiencies at national and global level. Moreover, the data collection for this indicator is normally relatively easy. The other benefit that makes it widely used is that understanding SEC is easy. Despite having several weaknesses, for instance, the exclusion of economic factors and utilization rate, most researchers recommend it, combined with other indicators, to bring about regional and temporal data comparability.

2.6.3. Specific Consumption of Energy Modified for the Rate of Utilization

The building's energy consumption comprises the base consumption that occurs even though there is the actual utilization of the building and of the energy consumption of the user. The energy efficiency can be enhanced by raising the building's rate of utilization as the base usage, which includes minimum ventilation, heating, and other continuous services of energy, runs despite how much the building uses. The advantage of this indicator is that the need for more built space can be adequately addressed. The adjustment of SEC aims at allowing for different rates of utilization of the building. In this case, the useful energy outputs are measured. The modified SEC for utilization rate (UR) is determined by taking energy delivered divided by the product of the area and the utilization rate as shown in Equation 2 below:

$$SEC_{UR} = \frac{Q}{uA} \quad (2)$$

In the determination of this indicator, the most famous approach is to measure the ratio of actual daily consumption hours (T_{actual}) to the highest possible consumption hours T_{max}, given by the following equation:

$$u = \frac{T_{actual}}{T_{max}} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(3)}$$

T_{max} is highest at a limit of 24 hours each day, though there might other practical shortcomings in some cases. A prime illustration is the night hours being reasonably excluded for the space of offices together with many other public and commercial spaces. Thus, other more complicated means can be devised for measuring u. Such approaches may comprise the ratio of number of person hours, or the cumulative number of hours that various people in the building have spent, to the maximum possible number of hours per person.

2.6.4. Consumption Energy Intensity

This indicator has been devised to measure some of the aspects that SEC and modified SEC and cannot. It is the best option since it indicates the element of the consumption of energy per user on buildings, which is particularly better in the measurement of the efficient space use. Its comparability between diverse kinds of buildings is undesirable, with its impacts being the work that is tightly packed (Forsstrom et al., 2011). The existing literature, therefore, recommends the need to measure the quality of energy using an appropriate, though different indicators of the power system.

2.6.5. Economic Intensity of Energy

The previous studies have emphasized the need to understand the economic-thermodynamic indicators in power plants. Forsstrom et al. (2011) describes them as the as indicators where the output, which is the delivered services or services, is measured based on the energy in terms of the units of thermodynamics and market prices. The development of this kind of indicator for buildings relies on whether the worth of the building to those using has a market value that is monetarily measurable. Sometimes, defining the building's value may be challenging, though its estimation can be made through the comparison of the rents and sales prices of a similar building. Some types of buildings, especially some public buildings, are very seldom sold, and they are perhaps the trickiest to value. The previous studies suggest the definition of the building's economic energy intensity (EEI) to presented in Equation 4 below:

$$EEI = \frac{Q}{R} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

As seen in the equation, Q represents the energy usage for a particular time, while R stands for the rent in € that is paid for the same period.

Studying this indicator is essential in the determination of how efficient the energy reaches the users at a desired market price. According to Forsstrom et al. (2011), the rent aspect offers a minimum value to the building's utility to the tenant. In case the tenant feels he or she is paying more for the building as opposed to the benefits she or he gets from it, it is advisable for them to stay there. The EEI plays a critical role of giving the peculiar ability of comparing the usage of the energy of the utility obtained from it. An effect of the EEI definition is that higher rent that have same levels of energy usage implies greater efficiency of energy. In a sternly physical perspective, this consequence may appear to be making no sense. However, from an economic perspective, people must be motivated in some way to pay higher rent. This action may be attributed to the fact that they perceive a greater utility level, especially in the aspects of energy efficiency within the building in which they reside. Moreover, with the same level energy usage, people can be offered a more valuable and better output, implying, by definition, the improved efficiency of energy. In general, all these aspects depend on the kind of data that is collected at the point where the energy is consumed, especially in buildings.

Other studies have highlighted electric power transmission and distribution as an indicator for both economic and efficiency aspects. Cadena et al. (2009) defines electricity transmission as the transportation of power via a set of lines, which comprise their linking modules that run at voltages that are greater than or equal to 220 kV. The authors hold that there is a need for the power transmission lines to ensure the reliability of the network for improved efficiency in the power plants. The electric power transmission's core aim is to transport the electric power produced at the power plants, which is normally situated far from the centers of usage to the locations of the final energy consumption. The issue of transmission then occurs in the movement of electric power, P(t), units over a distance, L, units at a particular time (t). This task can be completed when an electric current, I(t), produced at the power plants is transported via a wire of an electric conductor of length (L) at a certain voltage (V). Moreover, this process takes place recurrently at distribution circuits in homes at the mains voltage that are low and in elevated and extremely high voltage distribution and transformers and transmission lines. The calculation of power transmitted P(t) is essential in the improvement of power efficiency. The following equation is used to determine the power being transmitted in the power grid:

$$P(t) = V \cdot I(t) \cdot \varphi \qquad (5)$$

The coefficient, φ, as indicated in the above equation is the variable number approaching 1.

In another economic lens, power losses in the lines of power transmission and transmission can lead to reduced efficiency of the power plant and increased costs of operation. The electric power transmission at a given distance, L, is associated with losses of power, which results from the dissipation of heat in the electric conducting material (Cadena et al., 2009). This scenario is represented by the equation below:

$$\Pi(t) = R.I(t)^2 \qquad (6)$$

The parameters used in the equation are described as follows:

$\Pi(t)$= Loss of power,

R= the wire's electrical resistance, and

I(t) = The current passing through the electrical conducting material.

The losses of electrical power have some economic benefits. Cadena et al. (2009) argue that they are significant since they form the electric power that the generators must produce along with electric power that is taken to consumers despite being dissipated as heat in the systems of transmission and distribution. Voltage is always raised such losses they increase with the square of transmission of the electric current. The objective of this activity is to minimize losses and electric current to economic levels in the event where big quantities of electric power is transmitted (Cadena et al., 2009). Therefore, proper data collection of these parameter is essential, implying that the power plants need to develop data collection system or methods that can help get the data on power losses for the improvement of power efficiency.

2.6.6. Performance Index of Energy

There is a need for benchmarking indicators for the assessment of the potentials of energy saving, which the other indicators have failed explicitly show. The energy performance index (EPI) is such an indicator that most studies have suggested to bridge this gap. Forsstrom et al. (2011) define EPI using the following relationship:

$$EPI = \frac{Q_{actual}}{Q_{BAT}} \qquad (7)$$

As indicated in the equation, Q_{actual} stands for the building's actual energy usage and $QBAT$ represents the energy usage of a building of the same kind with the best available technology (BAT) being applied. This indicator indicates the values bigger than one as technology makes it possible for the improvements of energy efficiency. If the value of the indicator is higher, the technical saving potential is also high. In this case, the understanding of BAT is that a technology that exists in the markets can reasonably be used with investments that are cost-effective (Forsstrom et al., 2011). Here, the technologies involved are net energy consumers on the entire building's level. When the BAT advances to the energy-plus houses level, Equation 5 yields negative values for actual buildings that is using the energy.

2.7. Methodologies of Benchmarking for the Measurement of Efficiency

The estimates of electricity production are viewed as an extension of the estimates of microeconomic output function in the power plant. The primary ideas stem from the consideration of the fact that the function of output stands for some "ideal" entity for optimizing the output for a certain input (Cadena et al., 2009). Reducing the inputs for a particular product level as well as the maximization of advantages are variations of a similar idea. At this point, the limits of productions are essential for the relative analysis of efficiency to enable the comparison between the performance of a specific unit and the ideal unit for the determination of the relative efficiency level. This idea illuminates the synonymity that occurs between benchmarks of production and efficiency in power plants.

The data collected on the efficiency can be by comparing every observation in relation to the frontier (best observed practice). This data can be utilized in the estimation of various quantities of efficiency that cover diverse aspects of efficiency, based on the argument of Cadena et al. (2009). Such concepts comprise allocative efficiency (AE), technical efficiency (TE), and productive efficiency (PE). The AE relates to the entity's capability to bring together the inputs to make the rate of marginal substitution equal to the input's relative price. The PE refers to the capability of producing at low or minimum cost. The TE relates to the ability of the to make use of the inputs such that the output is maximized.

Establishment of the benchmark is from the functions of electricity production, where the output is taken as an input function. The estimation of the benchmark also helps in informing about the plant’s technical inefficiency. It can also be built based on the functions of cost functions, where total cost is taken as a function of input cost and the level of the output, with its estimation indicating the productive inefficiency. The previous studies suggest the two primary methods that are used in the analysis of production frontier, which include SFA (Stochastic production Frontier models) and DEA (Data Envelopment Analysis) (Cadena et al., 2009). Moreover, the multivariate statistical approaches have also been highlighted, involving factor analysis and principal components. According to Cadena et al. (2009), the first two techniques were explicitly designed for an empirical research of benchmarks for the determination of a hierarchical organization of entities, firms, and individuals in the aspects of cost or output efficiencies. The other methods relate to the exploitation of primary characteristics of such techniques. Most of the nations around the world have executed the regulations on the incentives in the power sector and applied the methods of benchmarking in the assessment of relative efficiency. The core frontier-based benchmarking approaches that many studies have utilized and reported include SFA, Corrected Ordinary Least Square (COLS), and DEA. Some studies have been conducted in power transportation to measure efficiency in the utilities of distribution (Cadena et al., 2009; Forsstrom et al., 2011). Such studies view transmission as a natural monopoly, where the service is delivered by only one utility in most countries. However, there are 7 transmission utilities in Colombia (Cadena et al., 2009). Therefore, the benchmarking method like SFA and DEA are relevant for the purposes of regulation.

2.7.1. SFA

In the utilities of power transportation, the formulation of a function of output relates to the final product and various inputs. Cadena et al. (2009) and other researchers have extensively explored the primary notions revolving around the model of SFA. In these studies, the core idea has been to estimate a function of production represented by the following equation:

$y = X\beta + \epsilon$………………………………………………………… (8)

In this equation, y represents the natural logarithm of the output, X is the matrix that encloses the variables related to the inputs and to environmental aspects. ϵ stems from two elements as shown in this equation:

$\epsilon = v - u$ ……………………………………………………………. (9)

where the vector v relates to the residual, which is viewed as a symmetrical noise that is normally distributed, and u is the efficiency, which is non-negative in functions of output and non-symmetric. The presences of u make ϵ to be non-symmetric, having the expected value of

$E[\epsilon] = E[v - u] = -E[u] < 0.$ ………………………………………….. (10)

Therefore, seeking to finding normality or symmetry as the output functions are specified using ordinary least squares (OLS) or looking for the positive skewness means that the approach of SFA is not enough to describe the relationship between power input and output in the power plants.

In the procedure, the subsequent step is to derive the distribution of ϵ from the collaborative distribution of *u* and *v*, where *u* possesses a semi-normal distribution. It is then followed by the calculation of the function of likelihood for the distribution of ϵ together with the estimation of the parameters with a view to obtaining the TE for all the entities involved (Cadena et al., 2009). Finally, the calculation of TE is done from the conditional expectation of e^u when ϵ is provided. Therefore, the determination of these aspects is crucial for the power plants to come up with measures of improving their efficiencies.

2.7.2. DEA

The DEA model belongs to a group of non-parametric techniques that relate to optimization methods that give way to the calculation of the relative efficiency of various power firms. The power plants that have the best relative performance are believed to be those using a minimum number of resources for a particular result, which is input-focused, or the ones producing the best outcomes from a certain number of resources, which is output-centered (Cadena et al., 2009). Such power companies are deemed efficient, which are applied in the construction of the efficient frontier that forms the benchmark for all the other power firms. In the transportation of power, the supplied electricity is a variable that depends on the market, which the power firms have no control over. However, such electricity needs to be offered depending on the standards of reliability and quality. Thus, the efficiency of the companies dealing with electricity transportation informs of their ability to offer the power based on such standards and with number of resources that is minimal. The resources of this kind may include capital assets, operative, maintenance, and administrative expenses, and transportation losses. The studies that came up with DEA framework for the first aimed at assessing the relative efficiency of various schools under diverse programs. It was especially developed to address the units or companies with multiple outputs and inputs, and whose specification of the function of production is difficult.

The first model has evolved into more advanced approach of analyzing data. The special interest of most of the studies is the extensions for permitting variable returns to scale because the power companies being analyzed indicate diverse productive structures. Other crucial extensions comprise the likelihood of tackling non-discretionary outputs and inputs, the restrictions to the weight values provided to each output and input in the analysis of every power plant, as well as the inclusion of principal part in the frameworks (Cadena et al., 2009). Such attributes are relevant for this thesis since they comprise some conditions affecting the efficiency of the power plants, where some of them such as saline environments and geographic factors might be controlled by these firms.

The DEA model helps in the assessment of the efficiency of the power plants. The framework is, first, based on input function because it such an orientation makes it possible to take the measurements of the efficiency that relates to different aspects of the electricity generation, transmission and distribution, and consumption. The analysis considers the recourses involved in the power process to assist in the determination of the efficiency of the power plants in different dimensions. The quality of power being delivered to consumers is also great importance, with the electrical conducting material together with the environment factors being the determinants of how efficient a power plant is (Cadena et al., 2009). In general, the data obtained on these variables can be controlled by the power plants to make decisions on improving their efficiency.

2.8. Critical Review

The available literature on this subject indicates that smart grid depends on various existing and future networking technologies of wired and wireless communication, for instance, communication of power line. The studies indicate that most power plants have adopted a vast range of supporting technologies, which include sensing and measurement, integrated communications, advanced component, improved interface, decision support, and advanced control in the smart grid (Kayastha et al., 2014). The execution of integrated communications and sensor data gathering, and networking have been considered as the leading factor in the creation of a dynamic and interactive infrastructure of power grid. Such implementation aims at integrating all the downstream components, toward consumers, and upstream components, toward generator, to operate seamlessly. In this case, the network of data communication in the smart grid offers the required functionality of for linking diverse actuators and sensors for sensing and collecting data. The data sensing and gathering in the smart grid, however, still needs to combat several technical and procedural challenges and improve the efficiency of the power generation systems. Therefore, data collection is required since more fields of power generation, such as large solar and wind farms, are being used in supplying energy into transmission grid with the aim of balancing the ever-increasing demand of energy.

The existing research holds that WSN (wireless sensor network) is a technology that supports data collection and monitoring applications. Gungor et al. (2010) and Kayastha et al. (2014) give a reason for the uses of WSN as being associated with low cost and easy to install. A WSN consists of wireless sensors of Stanford make known as WiMMS units, which is commonly used in the structure of wind turbines to give information concerning the wind turbine's dynamic behaviour and loading response (Kayastha et al., 2014). The authors further explain that the components of such a wireless sensor include a computational core for storage and processing of data, a communication interface to offer connection to the core network, and a sensing interface to provide a conversion from analog to digital. Despite the advantages of WSN, authors claim that wireless communications are essentially limited by the available rates of transmission together with the bandwidth. The studies have, therefore, suggested that there is a need of offering distributed processing in such wireless sensors to enable local processing of the sensed data within the sensors of the power generation system. In this case, only a small amount of processed information is supposed to be transmitted. The gap in the research on WSN is that they fail to provide the empirical evidence on how this technology can used to collect electricity data to help in improving the efficiency in power plants.

The concept of data collection in power plants has not gained prominence because most of the plants are still using conventional power grid, which lack intelligence. Kayastha et al. (2014) define intelligence as the capability of controlling and monitoring the different functional components or industrial appliances for the provision of optimal generation and utilization of power. The SCADA (supervisory control and data acquisition) is one of such technologies that have been used in providing the infrastructure of communications for the traditional electrical power grids (Kayastha et al., 2014). This technology has been used to gather data from different sensors and transmit it to a central computer for storage and processing. Researchers argue that SCADA system lacks real-time capability of control for the network of power distribution and the control over its upstream function, especially from premises of the consumer towards the site of power generation, is limited. Thus, the current electrical power grid is being transformed into a more efficient, more intelligent, more environmentally friendly, and more responsive system. Kim et al. (2010) suggests the smart grid as the system that meets these requirements. Such an evolution stems from the rise in power demand, setting of distributed system, emergent generation of renewal energy, and the power flow that is not reliable. Therefore, addressing such challenges require a smart grid with efficient mechanisms, a dynamic architecture, and intelligent algorithms.

Some researchers have also developed interests in investigating the challenges that the current data collection systems, such as smart grid, are facing that have made it difficult to improve efficiency in power plants. Kayastha et al. (2014) establish several issues including secure sensing and communications of data, optimization of the cost of network design, poor detection of data, and lack of advanced technologies of data collection devices, among others. The lack of secure data detection and communications in the smart grid is believed to have subjected the power systems to malicious attacks and intrusions, among other cyber threats. This situation, thus, calls for an alternative approach or system to ensure that data detection and communication is secure, where the use of private communication networks can be the best appropriate means for power plants. In the context of optimization of the cost of network design, it appears that power plants have been experiencing increased costs when using the smart grid (Kayastha et al., 2014). This situation indicates that it is needful for the power generation plants to conduct cost-benefit analysis before installing the network infrastructure for the detection and communication of data. This step is necessary since it helps in quantifying the risk and usefulness of the technologies that are available for this purpose. The issue of bad sensing of data stems from wrong detections, bias in large measurements, and drifts. Bad measurement data are also injected by the attackers and thus subjecting the smart grid to danger. Some recent literature has studied several strategies from the perspective of defenders and attackers. Moreover, the conventional data collection, sensing, and communication systems consume a lot of power, and thus calling for advances device technologies with lower consumption of power. This strategy helps in improvement of the efficiency of energy in the power plant and the lifetime of such gadgets. Therefore, these issues indicate there is a need for current smart grid to integrate intelligence to obtain a desired electric data collection system for the improvement of electricity efficiency in the power plants.

CHAPTER THREE

METHODOLOGY

2.9.Introduction

This chapter highlights the methodology that was utilized in the collection and analysis of data for this study. The chapter was divided into the following sections: time and location, target population, sampling and sample size, data collection methods, research validity and reliability, analysis and presentation of data, requirements of data gathering, and ethical issues of the research.

2.10. Research Design

This study used a literature-based MRP (major research paper), in which field-study MRP was conducted. The MRP entailed a detailed research work utilizing a rigorous and appropriate study approach. The literature used in this research is from the previous studies that were published within the last ten years on the topic of data collection in power plants. The descriptive survey method was used, focusing on the determination of the relationship between variables or the frequency with which something takes place as indicated by Nassaji (2015, p. 129). A descriptive design focused on getting the information that describes existing phenomena by asking questions that relate to individual attitudes and perceptions concerning the subject being investigated. Nassaji (2015, p. 129) notes that a descriptive research design provides facts about nature as well as the status of a situation, the way it occurs at the time of research. The selection of this research design has been based on the research problem, in connection with the theoretical framework for this study. Therefore, this methodology was appropriate because the intention of the study is to gather comprehensive information via descriptions that were useful in the identification of the variables.

2.11. Time and Location of Research

Research was conducted in January 2021 in Victoria, Australia, on the electric data gathering in the Australian power plants. This place was selected because it had realized a drop in electricity production by 10 percent in a period between 2017 and 2018 as compared to other parts of Australia (Australian Energy Updates, 2019). This decrease stemmed from the closure of the Hazelwood brown coal power plant in Victoria in March 2017. The closure reflects the decrease in the efficiency in power generation, transmission, and distribution by the plant. The place presented a need for the collection of the data that was relevant in determining if decreased energy efficiency could have contributed to the closure of the power plants in Victoria.

2.12. Study and Target Population

The consideration of the target population was key to this study. A population as the total collection of elements concerning which the researcher wants to make inferences (Sekaran & Bougie, 2011). The researcher worked with a study population that was a subset of the entire research population. This study targeted the power plants located in Victoria, Australia. The articles written on data collection solution in the power plants in this region, as well as on the theories and concepts of power generation and efficiency were targeted in this research. The researcher found numerous documents and articles published about the way the power plants are using diverse data gathering solutions to improve different kinds of efficiencies. The researcher collected the documents randomly and collected secondary data from them. Most of the articles that were selected for this study sought examine the different forms of data collection methods and systems. The diverse types of database were examined to get the studies on the topic. The database utilized was accessible via the university library system of the researcher. The research articles used were published in English in peer-reviewed journal articles on data collection, power efficiency, and power plants as the key words. The data and information that was mined in the selected articles focuses on the concepts and themes of power generation, data collection methods and systems, and research field employed in investigating electric data gathering for improvement of efficiency in power stations. A total of 90 power plants within Victoria and outside the state were studied. Thus, the target population comprised those in the power plants such as design engineers, power station executive supervisors, power consumers, and data management officers who are basically the people who interact with electricity data daily. The target population from Victoria was 200 and was as represented in Table 3.1. The strata were attained by looking at the different participants who worked in areas where electric data management was critical.

Table 1: Target population

Type of population (Victoria and other states)	**Total Number**	**Percentage**
Design engineers	80	40%
Power station executive supervisors	35	17.5%
Data management officers	25	12.5%
Power consumers	60	30%
Total	**200**	**100%**

Source: Researcher 2021

2.13. Sample Size and Procedure of Sampling

The way the sampling and sample size was conducted was crucial for the research. Kumar (2011) refers sampling to as a deliberate selection of several individuals who need to give the information or data from which research made conclusions regarding some larger group whom such people represent. This part discusses the sampling size as well as the sampling procedures. The sample size is described as to a subset of the population that is deemed to be representative of the whole population. The sample size offers the researcher a vivid picture of the scope of his or her research based on the total number of the participants that were used in the study. The sample size of the study was determined by the size of the target population, the confidence level, the standard of deviation, and the margin of error or rather the confidence interval.

The sample size for this study will be obtained based on Fisher’s formula shown below:

$$\mathbf{n_0 = \frac{z^2 pq}{d^2}}$$

Where Z is the standard normal deviation of the required confidence level; p stands for the proportion of the target population that has traits that are being measured; q is equal to 1 – p; and d represents the level of the statistical significance set.

In this study, there is no estimate available for the target population proportion sharing the traits of interest, and thus 50 percent of the population should be used as suggested by Kumar (2011). Given z as1.96 and the required accuracy as 0.05 confidence level, the calculation to arrive at the sample size is as follows:

$$N = \frac{(1.96)^2\ (.50)^2}{(.05)^2} = 384$$

This calculation gave rise to a sample size of 384.

Yamane (1967) gave a simplified formula for proportions for calculating the sample size, with a confidence level of 95% and p = 0.5, and a 5% level of precision. The formula is as follows:

$$\boldsymbol{n = \frac{N}{1 + N(e)^2}}$$

Where n stands for the sample size, N represents the population size, and e is the level of precision. The target population of this study will be 200 participants, the sample size will be determined by substituting the values of N = 200, e = 0.05 in the equation to get:

$$\boldsymbol{n = \frac{200}{1 + 200(0.05)^2}}$$

n = 133. Therefore, the adjusted sample size for this study will be 133 participants.

A stratified proportionate random sampling method was used to select the respondents used in the study. Stratified random sampling is described as an unbiased sampling technique of putting together a heterogeneous population into homogenous subsets and choosing within the individual subset to make sure there is representativeness (Kothari, 2014). This form of sampling aimed at attaining the desired representation from different sub-groups in the population. In the sampling method, the selection of the subjects is done where the existing sub-groups in the population are less or more represented in the sample (Kothari, 2014). Therefore, this research employed a simple random sampling technique to select the respondents in every stratum.

2.14. Data Collection Method

In the descriptive design, where both qualitative and quantitative techniques of research were utilized to gather both secondary data. Kothari (2014) refers the secondary data as to the data that is gathered from published surveys and experiments that have been carried out by other researchers. In this thesis, the secondary data employed was chosen based on their relevance to the main questions of research. This kind of data was gathered from the company reports and magazines. The information about the methods of data collection to use in power plants was based on several factors such as the availability of data and financial and human resources. There was a vivid identification of what was to be gathered and what needed to be disseminated. Moreover, the review of the relevant methodological concepts and the consistency with national and global standards were crucial elements that helped in choosing the techniques of electric data collection methods that best fit power plants.

Several sources of data were also considered for this study. SCADA is the commonly used system for the purpose of sensing and communicating data (Kayastha et al., 2014). The other type of data sources in the selected power plant was modeling, which relates to estimation data item that cannot be determined directly but can be determined according to the observable and measured data. This source was considered because its advantages of being able to enable quantification of variables that cannot be directly observed or measured, quick results, and saving money and staff. The other method the researcher used to get the information about electricity in the power plant was in-situ measurement. It is defined as method for gathering comprehensive data on consumption based on a measuring gadget. Such a device is, for instance, installed at the final consumption point. The advantages of in-situ measurement that made it suitable for this study include high-quality results and provision of comprehensive data about individual appliances and patterns of equipment usage. WSN was another system that provided the information on electricity in the power plant. Its advantages were that it is easy to install and is cheap to acquire (Kayastha et al., 2014). The other method of electric data collection was an improved smart grid sensor, known as fiber-optic Bragg grating (FBG). This device has an ability to measure and collect the data of different parameters in the electricity generation plant. Therefore, the core intention of this study was to identify the best method for the collecting electric data to enhance efficiency in power plant.

The survey was employed to gather quantitative data in this study. Survey refers to the self-report tool for collecting data wherein the respondent fills in the questionnaires. The study used questionnaires as a research instrument to get the desired information about the feelings, attitudes, values, perceptions, thoughts, and behavioral intentions of the participants concerning the electric data gathering to improve efficiency in power plants. Every question on the questionnaire was developed to respond to the specific objectives of the research. The questionnaires were suitable for this research since they were practical, and it was possible to use them to gather data from many people within a short time and with low costs incurred.

2.15. Validity and Reliability

2.15.1. Validity

The instruments used to conduct this research needed to have a higher level of validity. Kothari (2014) defines validity as the quality in which an instrument or procedure used in the study is accurate, meaningful, true, and correct. The criterion validity was employed in this study. Taherdoost (2016) defines this form of validity as how far a measure relates to an outcome. It entails measuring how well the researcher can predict an outcome for the other measure. This test was necessary since the data on the performance of the electric power generation power plant in the past could help in improving the present and future efficiency of the plant. In this research, content validity was used as a measure of the degree to which the data gathered by use of questionnaires represents the study objectives. The validity of the instrument was verified by conducting a pilot survey on 10 participants who were not part of the sample size.

2.15.2. Reliability

It was also necessary to ensure that the research tool used for this research was reliable. According to Kothari (2014), reliability entails the estimates of the level to which an instrument of study generates consistent results after performing several trials. Checking for reliability was essential because it relates to the consistency in all the components of the measuring tool (Taherdoost, 2016). Reliability, in this research, was determined using a pilot survey where 10 subjects not included in the sample will be involved. After some time, the researcher used the same questionnaire to test the same population to examine the consistency of the responses. However, reliability could not be adequate without combining it with validity in the measurement of consistency of the research tools used for this study.

2.16. Data Analysis and Presentation

The analysis of the quantitative data was done using MS Excel computer package. The results were categorized, summarized, and presented using tables and graphs. Thematic analysis was used to analyze the qualitative data. Thematic analysis is described as a qualitative study approach that can be extensively utilized across various epistemologies and study questions (Nowell et al., 2017). This method assisted in the identification, organization, analysis, reporting, and description of the themes obtained within a set of data generated for this research. It was used because it did not need any comprehensive technological and theoretical of other qualitative methods but provided a more accessible kind of analysis of the previous studies that have been published on data gathering approaches in power plants.

Thematic analysis has also be helpful in summarizing core characteristics of a big set of data, compelling the researcher to come up with a well-structured strategy of handle the data from secondary source, and thus assisting in the production of an organized and vivid final report.

The establishment of the continual relationship between the dependent and independent variables utilized content analysis. Content analysis refers to a technique that is famously applied in the social sciences, and thus it is a viable selection for this research (White & Marsh, 2006) The process comprised investigating the comments or statements by the respondents to examine the themes identified in this study to support the arguments in the presentation of the quantitative results.

2.17. Ethical Issues

The researcher considered the ethical issues in this research. Monette et al. (2013) claim that it is necessary to consider the ethical issues right from the outset when the researcher formulates the title and determines the topics which need to be covered as the variables. Before conducting the actual research, the researcher sought permission from the relevant authorities since the human subjects were involved. The researcher also put in mind the subject of the research based on the potential significance or value, especially in relation to its potential for yielding new knowledge and responding the relevant needs. The researcher also ensured that there were no potential conflict of interest occurred in regard to the use of the findings for monetary benefits since it would result in the manipulation of the data to obtain a particular result. Therefore, upholding the ethical issues in this study aimed at protecting the validity, reliability, and significance of the research.

It was also necessary to protect the interests of the people who took part in this study. The researcher ensured this by considering the aspects of anonymity, confidentiality, and informed consent. According to Monette et al., (2013), it is essential to inform the participants, as well as the relevant authorities concerning the research one, is going to undertake. In this case, the researcher would inform power station authorities before performing the study to prevent suspicions as well as resistance from the members and the managers of power companies. The researcher also sought consent from the respondents who volunteered to take part in participation in this study.

The information which the respondents offered was to be treated with the utmost confidentiality and utilized only for the purpose of this study. The provision of confidentiality was also be upheld by information of the participants in a statement that was made in the introductory note. In this statement, the respondents were assured that the information that they were going to give for this study would not be publicized to the third party. If the confidentiality of a participant were violated, it is likely that they would be victimized by those who feel they are targeted in the power plants.

It was also be necessary to omit the names of the participants on the questionnaires to ensure anonymity. The aspect of anonymity relates to the respondents who are going to offer their responses after coming across the links that they get on their social media platforms. It was necessary for the respondents to remain anonymous to the researcher because the survey did not need them to give any information that will identify them. The respondents reached via mail also required to remain anonymous to ensure their answers cannot be traced back to the email accounts once they have accessed the survey link. Privacy was also be another crucial ethical issue to be considered in this research. The researcher assured the respondents that the information they were providing would be kept private. They were encouraged to be free not to answer any question that they think would be an intrusion into their privacy. Following this account, the researcher ensured that the names of the participants are not exposed, and codes were used instead.

CHAPTER FOUR
RESULTS

4.1. Introduction

This chapter offers the results obtained from this research, from the set of articles selected for analysis. The presentation of the chapter in in two parts, the specific results generated from the study being the first part while the empirical evidence from the previous empirical studies published data gathering in power plants being the second one.

4.2. Results of power plants based on the source of energy.

During the researcher, the data was collected from 90 power plants, whose sources of power comprised coal, natural gas, hydroelectric, renewable sources, landfills, and others. The results of the power plants based on their different sources of energy is as shown in Table 2 below:

Table 2: Results of the power plants based on the energy sources

Source of Power	**Number of Plants**
Coal	6
Natural Gas	16
Hydroelectric	32
Renewable	15
Fossil Fuels	10
Others	11

The data given in the table was used to construct the bar chart shown in Figure 2.

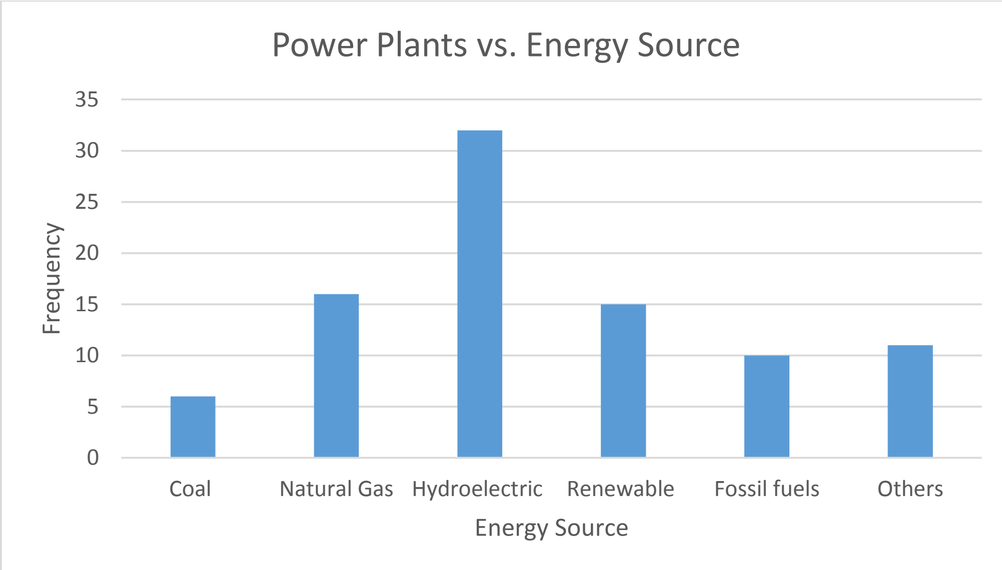

Figure 2: The bar chart showing the number of power plants by the source of energy

The results obtained in Table 2 and Figure 2 indicate that are more hydroelectric power plants as compared to the plants deriving energy from other sources. The hydroelectric power plants account for 35.56%, while those in the categories of renewable energy resources, natural gas, coal, fossil fuels, and others account for 16.67%, 17.78%, 6.67%, 11.11%, and 12.22%. The power plants getting their energy from coal are the least on the list.

4.3. Implementation of Data Collection Systems in Power Plants

The participants were asked questions related to the implementation of the systems and techniques of electric data collection in the power plants located in Victoria, Australia. The findings were as indicated in Table 3 below:

Table 3: Status of the implementation of data gathering systems and methods in the power plants

No.	Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly
1	The power plants have implemented various systems of data collection.	33.33%	44.44%	5.56%	8.89%	7.78%
2	The costs of the data gathering systems in the power plants have been high.	50.00%	27.78%	2.22%	11.11%	8.89%
3	Most of the data collection systems and methods are old.	36.67%	41.11%	1.11%	11.11%	10.00%
4	The level of satisfaction among different workers involved in collecting electric data is high.	20.00%	16.67%	2.22%	33.33%	27.78%
5	The implementation of the modern data collection technologies has been effective and efficient.	16.67%	17.78%	4.44%	31.11%	30.00%
6	The data collected on electricity have helped in improving the efficiency of power plants	36.67%	46.67%	1.11%	10.00%	5.56%

The responses on the question of whether the power plants have implemented various systems of data collection or not has shown the highest number of the respondents agreeing (44.44%) and strongly agreeing (33.33%) to this idea. Those who disagreed and strongly disagreed accounted for 8.89% and 7.78%. The responses in the category of “neutral” constituted only 5.56%.

The greatest number of the respondents strongly agreed to the fact that the costs of the data gathering systems in the power plants have been high, with 50.00% response rate. This figure was followed by those who agreed, accounting for 27.78%. Those in the groups of neutral, disagree, and strongly disagree constituted 2.22%, 11.11%, and 8.89%, respectively.

Regarding the question of most of the data collection systems and methods being old, many respondents agreed to this idea, with those in the category of “agree” accounting for 41.11%. Moreover, those in the groups of “strongly agree”, “neutral”, “disagree”, and “strongly disagree” were 36.67%, 1.11%, 11.11%, and 10.00%.

The responses concerning whether the level of satisfaction among different workers involved in collecting electric data is high indicate a high rate among those who felt this was not the case. The results for those who disagreed was 33.33% and 27.78% for those who strongly disagreed. Those who agreed and strongly disagreed accounted for 16.67% and 20.00% respectively. Those who did not show whether this idea was true or not (neutral) were 2.22%.

The question on whether the implementation of the modern data collection technologies has been effective and efficient also generated mixed responses among the participants. The leading response rate was for those who disagreed (31.11%), followed by those who strongly disagreed to the idea, accounting for 30.00%. 17.78% agreed to the idea while 16.67% of the respondents expressed their strong disagreement to the idea. Furthermore, 4.44% were unsure of whether the execution of the systems and methods of data gathering in power plants was efficient and effective.

Concerning the query of if the data collected on electricity have helped in enhancing the efficiency of power plants, most of the respondents (46.67%) agreed to this idea. 36.67% of the participants strongly agreed that the data gathered has been used to make decisions on improving the efficiency of power plants in Victoria, Australia. Those who disagreed and strongly disagreed were 10.00% and 5.56%, respectively. Finally, those who were not sure about where the data gathered led to the improvement of efficiency in power plants were 1.11%.

4.4. Empirical Results

This section presents the results that were obtained from the empirical studies that were performed in the past by other researchers. The results have been grouped into the components of power systems and determination of power efficiency.

4.4.1. Results based on the components of the system used for collecting data

Table 4: Results of data collection devices and methods for different components of power plant

Components of Power System	**Data Collected**	**Method/Device of Data Collection**
Power Generation	Electricity generated, load factors, thermal efficiency, maximum load, fuel used, electricity consumed, net power supplied, capacity, and sales.	Meter Data Management System (MDMS) and modeling method.
Transmission and distribution		
• Wind Turbines	Electricity produced, power used, power supplied, capacity of equipment, etc.	Meter Data Management System (MDMS)
• Transmission substation and transmission feeder	Voltage, current, temperature, transmission losses.	Sensor and actuator networks (SANETs)
• Distribution substation and distribution feeder	Temperature, current, voltage, losses, and quantity of electricity supplied.	Sensor and actuator networks (SANETs)
Customer Premises		
• Home area network (HAN)	Electricity used, conditions of peak load, temperature, etc.	In-measurement and smart meter such as FBG, PLC, WiMAX, or satellite and modeling method.
• Neighborhood area network (NAN)	Power billing and electricity distributed to the neighborhood.	Smart meter such as FBG, PLC, WiMAX, or satellite and modeling method.

The results in Table 4 show different methods and devices of data gathering at different points of the power system. The results were categorized according to the location where the data was collected. In the power generation area, the MDMS and modeling technique were used to collect the data on electricity produced, thermal efficiency, load factors, maximum load, electricity consumed, sales, net power supplied, fuel used, and capacity of the equipment. At this point, WSN was also used for sensing and monitoring of electric data at the power generation point.

At the power transmission and distribution point of the smart grid, the data collected to monitor substations. The areas where data was gathered include wind turbines, transmission feeder, transmission substation, distribution substation, and consumer premises. The data collected at the locations comprise current, voltage, temperature, transmission loss, and quantity of electricity supplied as indicated in the table. The systems used to collect the data were SANETs. The sensors and actuators were among the devices that sensed the variations in the current and voltage, which were useful in the determination of thermal efficiency of the power system.

Collecting the data in consumer premises was also essential in determining the parameters that contribute to the efficiency of the power plant. At this point, the areas of concern were HAN and NAN. The data gathered at these points encompassed Electricity used, conditions of peak load, temperature, power billing, and electricity distributed to the neighborhood. The device used in the area was smart meter such as optic FBG, WiMAX, PLC, or satellite, while the data collection technique was in-measurement. The measurement action in Australia was found to be part of the survey on consumption of natural gas and electricity of households, though the greatest concentration was on electricity only. It is only in Australia, unlike other nations, where the measurement of power consumption was performed by the households (Australian Energy Updates, 2019). It was found that most of the consumer premises had smart meters, which were provided by the power plants to help in monitoring the consumption of power by homes. The information observed and collected from the smart meter was as shown in Figure 2. The metering of the consumption of power was to be done for more than 1 hour for computers and TVs, 24 hours for freezers, and coolers, or by use of dishwashers and washing machines.

Function	Description
Start function: Time	
Function 1	Voltage (V) and frequency (Hz)
Function 2	Current (A) and power factor (Cos Φ)
Function 3	**Power** (W)
Function 4	**Measurement time**
Function 5	**Consumption** (KWh)
Function 6	Costs in Euro
Function 7	Price by kWh.

Figure 3: General information on the smart meter

Modeling method was also used to collect the data on electricity consumption by households. The information obtained using this technique was validated based on the different modeling parameters that were used. The data was used to compute the final energy consumption and annual power demand variation as shown in Figure 3. The variation in demand was measured based on the peak load per hour (Forsstrom et al., 2011).

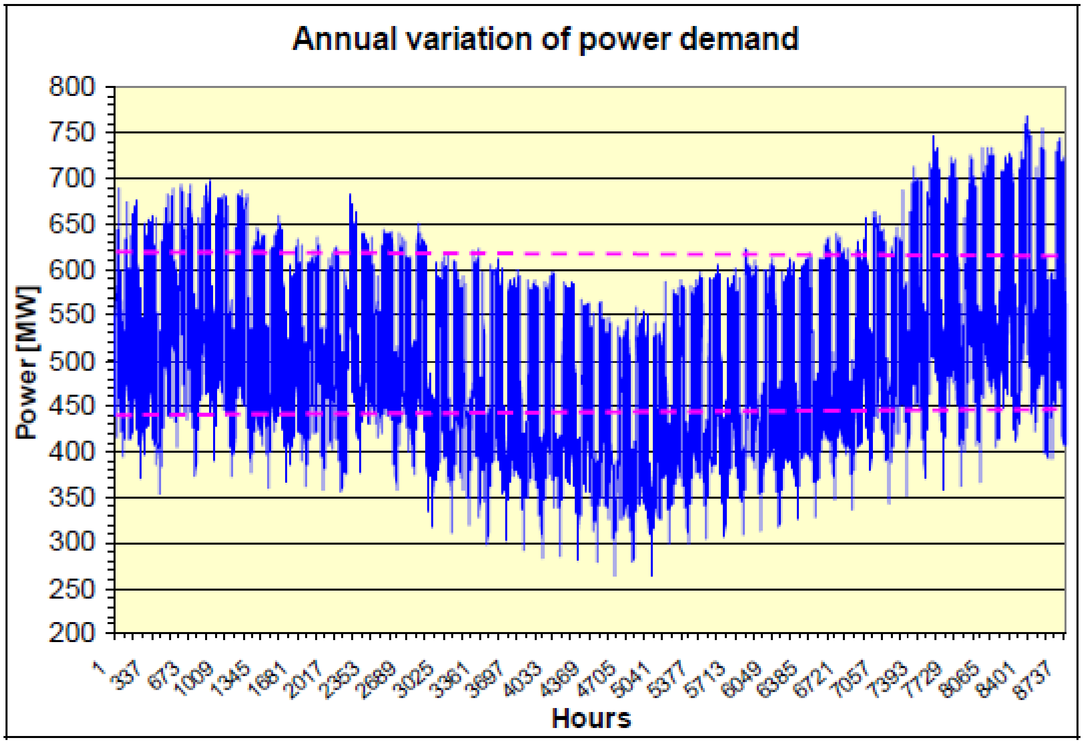

Figure 4: Yearly variations of the demand of power on a seasonal and hourly basis (Forsstrom et al., 2011)

The figure indicates that power demand varied based on the heating requirements of the consumers and the season. This information helped in the determination of efficiency of the power plants during different seasons.

4.5. Efficiency Determination

The performance efficiency was determined based on the data of different parameters of electricity. The parts of performance that were measured include energy efficiency, operational efficiency, and economic efficiency. The measurement of energy efficiency was by taking the output power (watts) or energy output in joules divided the input power (watts) or energy input in joules, expressed in the form of a percentage. The measurements that were collected for this purpose comprised voltage (V) volts, current (I) in amperes, and time (t) in seconds. The formula used to get energy efficiency using these parameters was as follows:

Energy, E = Power (P) x time (t) Joules

Where power in Watts = Voltage (V) x Current (I)

$\eta = \frac{Energy_{out}}{Energy_{in}} \times 100\%$... (1)

Where η stands for the efficiency in percentage, E_{in} is the energy input in joules (J).

E_{out} is the energy output in joules (J).

The determination of economic efficiency was based on the following formula:

$\varphi_{ee} = C / E$ …………………………………………. (2)

where φ_{ee} = economic efficiency (Euro/kW)

C = costs of production for a stated period (Euro)

E = energy output from the power plant in a given period (joules or kWh)

Operational efficiency was determined using the following relationship:

$\mu_{oe} = \frac{(100)E}{E_{100\%}}$... (3)

Where μ_{oe} is the operational efficiency in percentage, E = energy output from the power plant in within a given period (joules or kWh), and $E_{100\%}$ it the potential energy output from the power plant that is run at 100% within a particular period (kWh or joules).

CHAPTER FIVE
DISCUSION

5.1. Introduction

This section of the thesis discusses the findings obtained from both the actual survey and empirical research. The areas of concern related to data gathering to improve the efficiency of power plants include status of the implementation of such the systems and techniques of data collection, type of data collected, and various methods and systems that have been proposed for this purpose.

5.1.1. The Status of the Implementation of Data Collection Systems and Methods

The findings on the implementation of the systems and techniques of electric data collection in the power plants located in Victoria, Australia has been helpful for this these. The results have been in determining the existence of such data collection methods, level of satisfaction among those working in the power plants, challenges of SG, and the data that has been collected to help in the improvement of power plants' efficiency. The findings indicated of the existence of the data collection solutions the power plants in Victoria, though most power plants seemed to be still using the old technologies. Coal power plants uses the old methods and systems of collecting and recording data in the power generation. Even though the figures obtained from the research present coal as the last source of energy for the power plants in Victoria, coal remains to be the second biggest fuel being used in Australia (Australian Energy Statistics, 2019). Most of these plants are scheduled to close in few years to come because of their reduced efficiencies in operations and meeting the demands of the energy consumers, as well as the emergency of new forms of energy that are hitting the coal industry hard. Some of the new technologies comprise the renewable energy, which have several advantages over the coal fired plants. The literature present some of the advantages such as being cleaner, eco-friendly, and cost-effective among other benefits (Arangarajan et al., 2014). Some of the renewable sources of energy used in Victoria include solar and wind. The solar power plants comprised Wemen Solar Farm, Swan Hill Solar Farm, Bannerton Solar Farm, etc. In Victoria, some of the coal-fired power plants comprise Yallourn, Loy Yang A, and Loy Yang B. Gas turbines have also been used in the region, which has also indicated the use of both modern and old system of collecting data. Some of the plants deriving their energy from gas include Longford, Valley Power, Mortlake, and Bairnsdale, among others. Hydroelectric power plants dominate the area, and thus implying that they need to have up-to-date data collection system to help in enhancing their power efficiency to address the issue of increased energy demand. Moreover, natural gas constituted 25% of the energy being consumed in Australia as reported between 2017 and 2018 (Australian Energy Statistics, 2019). Fossil oils have also been used in some plants to generate electricity, it is not clear how they collect the data to help in improving their efficiency in power generation and operations. Some plants were also found to be using other sources of energy such as landfills and biomass, among others.

The research findings also indicated the relationship between the level of satisfaction among the workers in the power plants and the data collection solutions employed. Those who felt that they are satisfied reveal that any system they were using worked well with them.

For instance, it is possible that the young workers working on the power system did not like the old data collection methods, where some needed to take the readings manually to determine the parameters they required for the calculation of efficiencies. This kind of workers require advanced devices that would sense and collect the data and then communicated to computers where the efficiencies can be determined automatically in the SG. Such systems can also detect energy wastages and give the required computations using the various indicators of efficiencies in the power system to help in any improvements needed (Forsstrom et al., 2011). Therefore, better and latest data collection solutions are required for power plants to achieve this goal.

The issue of costs has been a limiting factor for the implementation of the latest data collection solutions in the power plants. This finding resonates with the existing literature in the sense that the requirements of data collection systems are extensive. The AMI architecture is composed of many components, such as SANETs, WANs, HANs, WAMSs, and NANs, among others (Kayastha et al., 2014). Acquiring and managing these components is costly since technology is always expensive. It would also require the power plants to hire IT personnel who are highly experienced. This situation implies that they have to be ready to pay high salaries as compared to the resources they have been using on the traditional methods of collecting data in the power systems.

The status of implementation of the data collection technologies indicates the need have modern solutions that are to effective and efficient. The methods that are currently in use such as SCADA, among others are not effective. This technique can only offer the infrastructure of communications for the traditional power grids. The available literature has suggested the use of smart technologies because of their varied characteristics and benefits. Kayastha et al. (2014) supported this idea by highlighting some crucial characteristics of SGs, which include high power quality and reliability; self-healing; cyber-attacks resistance; assets and storage optimization; consumer-friendly; etc. The latest solutions of data collection also have a vast range of supporting technologies, including advanced components; enhanced interface, improved measurement sensing; and integrated communications. They are also helpful in supporting the decision-making process, which can be adopted for the AMI in SG. The implementation of sensor data collection and networking together with integrated communications can help in the creation of an interactive and dynamic power grid architecture. This step can assist in the integration of all the downstream (toward consumers) and upstream (toward generator) components to operate seamlessly. The implementation of data collection technologies in the SG will offer the necessary data and functionality for linking diverse actuators and sensors for sensing, including meter reading, collecting real-time data from different locations of the SG for the improvement of efficiency in the power grid (Kayastha et al., 2014). The systems will also play a critical role in controlling the data collection process to bring about a reliable, resilient, and optimal operation of the power plants. A combination of actuators and sensors with advanced techniques of data communication will also assist in meeting the increasing via the integration of demand from consumers. Achieving this goal may requires the integration of renewable resources of energy, isolation and restoration of power outage, optimization of energy flow, empowerment of the consumer with instruments to optimize their consumption of energy and offering electricity power quality. The data collection of sensors can be utilized in the SG to get and convert data into information for the improved performance and efficiency of the power grid. A good illustration is a smart meter that can serve as a sensor node for measuring, prediction, and reporting of the usage and demand of power to the consumers and to the public utility for the provision of information that can help in responding to energy demand. Therefore, sensors can be installed in the power grid for the detection of any parameters that can help the management of the power plants to take the required action to improve their efficiency.

5.1.2. Types of Data Collected

The data collected on electricity can help in improving the efficiency of power plants. The data gathered indicated the indicators of energy efficiency of the power system. However,

indicators that this study suggests not final since they are only the current perspectives of the authors. The indicators rely on the purpose of the measurement. One of the measurements was the quantity of energy, denoted by Q. It refers to the energy, in kWh, that consumers use yearly or for a certain period. This measurement was used in place of the energy consumption, which was used in the computation of energy efficiency of the power plant. Thus, a vigorous analysis of energy usage needs to consider the energy consumed and saved at the end-user stage. The energy input is also another measurement that was taken at the power plants. The energy input was based on various sources of energy supply to the plant, such as coal, fossil fuel, natural gas, and renewable energy resources, among others. Another crucial parameter to measure in the power grid was specific energy consumption (SEC), as suggested by Forsstrom et al. (2011). SEC has been commonly utilized in making comparisons of the energy consumption at national and international levels. The data on SEC is always easy to collect and it is also to understand.

Economic energy intensity was another indicator for the energy efficiency. In this case, the data collected focused on output and services or product delivered. The parameters are measured based on the prices of energy in the electricity market and energy demands depending on the thermodynamic units (Forsstrom et al., 2011). This measurement can be taken in residentials areas, where the rate of consumption of energy is high. The power plants, thus, need to take data in these areas to ensure that it is economical to deliver power to these places at affordable rates. This activity helps the power generation firms to determine their efficiency in operations and services to customers.

5.1.3. Different Methods and Systems of Data Collection

This study examined also different methods and devices that can be used to collect data on electricity to help the power plants improve their efficiency. The study identified different types of data that can be collected in the power system to help in the determination of power efficiency. The available literature explains how the efficiency (both energy and performance) of power plant is measured. The data collected by all the different types and methods of data gathering presented the efficiency indicators of the power plant. The indicators were found in this study and resonate with those published in the previous literature include energy, power, voltage, current, and time (Forsstrom et al., 2011). However, there was no indicator that could help in determining all the aspects of energy and power efficiencies. Thus, a set of parameters were found to be affecting a particular aspect of energy or power efficiency. For instance, a set of measurements for determining energy efficiency comprised current, voltage, and time. Therefore, the measurements of these parameters need to be synchronized to the power system of the plant to help in the identification of power consumption for every part of the plant and consumers.

The findings indicate that there are numerous methods and devices for this purpose indicate that the power stations in Australia were aware of the need for gathering data of different parameters within their power system. The electric data collection technique proposed for power plants is an improved smart grid sensor, known as fiber-optic Bragg grating (FBG). This system is used for the construction of network of smart transmission. It proved to be effective for this purpose because of its ability to measure and collect the data of different parameters in the electricity generation plant. FBGs are usually utilized within a packaging of loose tubes in which the shielding of FBGs from mechanical stress is done using a protecting capillary. This method has proved to be appropriate for sensing and collection of electric data in power plants because of its significant improvements in the recent years. The consideration of this system is line with the theory of in-FBG sensors for industrial applications. The optical FBG sensor networks can be integrated with standard electronic controllers in the power by the use of an optoelectronic interface.

5.1.4. Stages of Power Grid

The data at each stage of the power plant was found to be essential to identify the need for efficiency improvement at that point and overall plant. At the power generation stage, the sources of power production determined the type of data to be gathered to help in the determination of power or energy efficiency (ABB Inc., n.d.). As an illustration, the hydro power was determined by the voltage that was supplied and the current in the power line in a given period. Moreover, the fuel used and chemical reactions that occurred during the generation of electric energy also formed a source of electric that were necessary for the measurement of the operational, energy, and cost efficiencies of the power plant. The efficiency of generation which appeared to be varied based on the technology or system used was essential for power plants. Most of the technologies employed in power generation result in 60 percent efficiency (ABB Inc., n.d.). However, it was realized that a significant amount of energy remains in the process of electricity generation even where there was an efficiency of 60 percent. This data indicated the possibility of the production cost being high for the plant together with a tremendous wastage of the resources that are limited in nature. Therefore, the information could only be found to improve the energy, cost, and operational efficiencies of the power plant when better data gathering method and system was used in the power production area.

The transmission and distribution stage were also another point were efficiency mattered. At this point, the energy is transported in large amount from where it is generated to where it is consumed via a process called transmission. Power is then separated from distribution, meaning the process of electric energy delivery from the transmission grid of high voltage to specific places like a commercial park or residential street. The literature indicate that distribution consists of the feeder lines and substations that take power from the grid of high voltage and gradually steps it down to a lower level (120v) that can be used in residential locations (ABB Inc., n.d.). During power transmission and distribution, there is the loss of some energy as a result of resistance of the equipment and wires through which the electric passes. The past studies indicate that power loss accounts for between 6 and 8 percent (ABB Inc., n.d.). The data on the losses collected show that the cost of inefficiency is high. Therefore, there is a need for the power plants to collect and monitor the data on efficiencies to help in coming up with ways to improve the system. The components that need to be installed with data collection systems are the transformers, which are the most standardized parts of the electrical power equipment that present a prime improvement target in the power plant.

CONCLUSION AND RECOMMENDATIONS

The collection of electric data presents a prime target for improvements of efficiency in a power plant. There are several sections of the plant where data gathering system and methods need to be employed, which comprise power generation, transmission and distribution, and consumption points. Each of these areas requires the collection of a set of data on electricity to help in the determination of efficiency. The three types of efficiencies that need to be improved in a power system include energy or power efficiency, cost efficiency, and operational efficiency. The data collected for this purpose indicate the need for the power plants to have better and advanced methods and systems for gathering the data to help in spotting where the efficiency is poor within the power plant and seek to improve it. However, this study has a limitation that it relied on the secondary data. Therefore, there is a need for further researcher where primary data has to be collected using a survey method to make the study results valid and reliable.

REFERENCES

ABB Inc. (n.d.). Energy efficiency in the power grid. Retrieved from http://www.altomelbilen.dk/rapporter/Energy%20efficiency%20in%20the%20power%20grid.pdf.

Al Rashdan, A.Y. & Germain, S.W. (2018). Automation of data collection methods for online monitoring of nuclear power plants. United States. https://doi.org/10.2172/1475451

Arangarajan, V. Than, A.M., Shafiullah, G.M., & Stojcevski, A. (2014). Potentialities of Renewable Energy in Victoria, Australia. *Journal of Energy and Power Engineering*, 8, 990-1001.

Australian Energy Statistics. (2019). Australian energy update 2019. *Australian Government*. Retrieved from https://www.energy.gov.au/sites/default/files/australian_energy_statistics_2019_energy_update_report_september.pdf.

Bryman, A., & Bell, E. (2011). *Business research methods*. Oxford University Press, USA.

Cadena, A., Marcucci, A., Perez, J.F., Duran, H., Mutis, H., & Palacious, F. (2009). Efficiency analysis in electricity transmission utilities. *AIMS'Journals, x*(x), 1-22.

Casey, P., Jaber, N., Tepe, K. (2011). Design and implementation of a cross-platform sensor network for smart grid transmission line monitoring, In *Proceedings of Smart Grid Communications Conference (SmartGridComm)*, 303–308.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Sage Publications.

Forsstrom, J. Lahti, P., Pursiheimo, E., Rama, M., Shemeikka, J., Sipila, K., Tuominen, P., & Wahlgren, I. (2011). Measuring energy efficiency: Indicators and potentials in buildings, communities and energy systems. VTT-Research Notes 2581.

Frolec J, Husak M. Wireless sensor system for overhead line ampacity monitoring, In *Proceedings of International Conference on Advanced Semiconductor Devices & Microsystems (ASDAM)*, 211–214.

Ghorbanian, M., Dolatabadi, S.H., & Siano, P. (2019). Big data issues in smart grids: A survey. IEEE Systems Journal, 99, 1-12. https://doi.org/10.1109/JSYST.2019.2931879.

Gungor, V.C., Lu, B., & Hancke, G.P. (2010). Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics, 57*(10), 3557–3564.

Kayastha, N., Niyato, D., Hossain, E., & Han, Z. (2014). Smart grid sensor data collection, communication, and networking: a tutorial. *Wireless Communications and Mobile Computing,* 14, 1055-1087. https://doi.org/10.1002/wcm.2258.

Kaynakli, O. (2012). A review of the economical and optimum thermal insulation thickness for building applications. *Renewable and Sustainable Energy Reviews*, 16, 415- 425.

Kim, Y.J., Thottan, M., Kolesnikov, V., & Lee, W. (2010). A secure decentralized data-centric information infrastructure for smart grid. *IEEE Communications Magazine, 48*(11), 58–65.

Kothari, C. R. (2014). *Research Methodology: Methods and Techniques*. New Delhi: New Age International Ltd.

Kumar, S. (2011). *Methods for Community Participation: Complete Guide for Practitioners*, Vistar Publications, New Delhi India.

Li, L., Zhu, Y., & Yu. Y. (2010). Link scheduling and data forwarding in wireless sensor networks of long chains tree topology, In *Proceedings of IET International Conference on Wireless Sensor Network (IET-WSN),* November 2010; 297–300.

Mahlia, T.M.I., Lim, J.Y., Aditya, L., Riayatsyah, T. M. I., Pg Abas, A.E., & Nasruddin. (2017). Methodology for implementing power plant efciency standards for power generation: potential emission reduction. *Clean Technologies and Environmental Policy*, 20:309–327

Melani, A.H.A., Murad, C.A., Netto, A.C., Souza, G.F.M., & Nabeta, S.I. (2019). Maintenance strategy optimization of a coal-fired power plant cooling tower through generalized stochastic petri nets. Energies, 12(1951), 1-28.

Monette, D. R., Sullivan, T. J., & DeJong, C. R. (2013). *Applied Social Research: A Tool for the Human Services*. New York: Cangage Laerning.

Nassaji, H. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*, *19*(2), 129-132.

Nowell, L.S., Norris, J.M., White, D.E., & Moules, N.J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16, 1-13.

Ota, N. (2011). The home area network: architectural considerations for rapid innovation, White Paper for
Trilliant Inc. Retrieved from: http://www.trilliantinc.com/libraryfiles/white-papers/HAN_white-paper.pdf.

Sekaran, U. & Bougie, R. (2011). *Research Methods for Business: A Skill Building Approach* (4thed). New York: John-Wiley.

Sritram, P., Treedet, W., & Suntivarakorn, R. (2015). Effect of turbine materials on power generation efficiency from free water vortex hydro power plant. *Materials Science and Engineering* 103, 1-7. https://doi.org/10.1088/1757-899X/103/1/012018.

Stevanović, S. (2013). Optimization of passive solar design strategies: A review. *Renewable and Sustainable Energy Reviews*, 25, 177-196.

Stewart, Mark. (2017). The future of Australian electricity generation. *The Institution of Engineers Australia.* Retrieved from https://www.engineersaustralia.org.au/sites/default/files/resources/Public%20Affairs/The%20Future%20of%20Australian%20Electricity.pdf.

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management,5*(3), 28-36.

Ting, Y., Zhidong, Z., Jiaowen, W., & Ang, L. (2010). Research on transmission data system of smart grid based on IPv6 DiffServ model, In *Proceedings of Power and Energy Engineering Conference (APPEEC)*, 1–4.

Victoria. (n.d.). Chapter 2: Victoria’s renewable energy resources. Retrieved from https://www.parliament.vic.gov.au/images/stories/committees/enrc/renewable_energy/report/Chapter_2.pdf.

White, M.D. & Marsh, E. (2006). Content analysis: A flexible methodology. Library Trends, 55(1), 22-45.

Zhang, T. (2020). Methods of improving the efficiency of thermal power plants. Journal of Physics: Conference Series, 1449(012001), 1-5. https://doi.org/10.1088/1742-6596/1449/1/012001.