**SAFEPASS:**

**AN IMPLEMENTATION OF BOX-JENKINS AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) FOR CROWD FORECASTING APPLIED IN QUARANTINE PASS**

A Capstone Project Presented to the Graduate Program

College of Engineering and Technology

Pamantasan ng Lungsod ng Maynila

In Partial Fulfillment of the Requirements for the Degree

Master’s in information technology

­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

By

Joane Marie F. Llamera

Dr. Khatalyn E. Mata

Thesis Adviser

August 2021

**APPROVAL SHEET**

The capstone project hereto titled

**SAFEPASS:**

**AN IMPLEMENTATION OF BOX-JENKINS AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) FOR CROWD FORECASTING APPLIED IN QUARANTINE PASS**

prepared and submitted by Joane Marie F. Llamera in partial fulfilment of the requirements for the degree of Master’s in Information Technology has been examined and is recommended for acceptance and approval for **Oral Examination**.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**DR. KHATALYN E. MATA**

Adviser

PANEL OF EXAMINERS

Approved by the Committee on Oral Examination

with a grade of \_\_\_\_\_\_\_\_\_ on \_\_\_\_\_\_\_.

**PROF. MANUEL L. OCAMPO**

Panel Chair

Chairman

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Panel Member Panel Member

Member Member

Accepted and approved in partial fulfilment of the requirements for the degree of

Master’s in Information Technology.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Dr. Denvert C. Pangayao Dr. Clydelle M. Rondaris

Director Dean

Graduate Program College of Engineering and Technology

**ABSTRACT**

The maximum length of the abstract is one to two pages. Use 12-point Times New Roman and 1.5-line space.

An abstract should contain the summary of the study including significant findings, information, data and analysis. Also, problems and objectives are solved and addressed.

**ACKNOWLEDGEMENTS**

The acknowledgements should also be limited to one page but can exceed if necessary.

The CET-Graduate Program would like to thank Ms. Diana Jane Saya for helping us format this thesis manuscript for the undergraduate and graduate program of the College of Engineering and Technology.

The pattern of the manuscript was from the Chemical Engineering Department of De La Salle University.

**TABLE OF CONTENTS**

*(Note: Table of Contents, List of Tables and List of Figures are single spaced.)*

[APPROVAL SHEET i](#_Toc305755519)

[ABSTRACT ii](#_Toc305755520)

[ACKNOWLEDGEMENTS **Error! Bookmark not defined.**](#_Toc305755521)

[TABLE OF CONTENTS ii](#_Toc305755522)

[LIST OF FIGURES ii](#_Toc305755523)

[LIST OF TABLES ii](#_Toc305755524)

[INTRODUCTION 2](#_Toc305755525)

[1.1 Background 2](#_Toc305755526)

[1.2 Fonts 2](#_Toc305755527)

[1.3 Paragraphs and Line Spacing 2](#_Toc305755528)

[1.3.1 Section numbering 2](#_Toc305755529)

[1.3.2 Lists 2](#_Toc305755530)

[1.4 Significance of the Study 2](#_Toc305755531)

[1.5 Scope and Limitations 2](#_Toc305755532)

[REVIEW OF RELATED LITERATURE 2](#_Toc305755533)

[2.1 Referencing within the Text 2](#_Toc305755534)

[THEORETICAL FRAMEWORK 2](#_Toc305755535)

[3.1 Equations **Error! Bookmark not defined.**](#_Toc305755536)

[MATERIALS AND METHODOLOGY 2](#_Toc305755537)

[4.1 SI Units **Error! Bookmark not defined.**](#_Toc305755538)

[4.2 Figures **Error! Bookmark not defined.**](#_Toc305755539)

[RESULTS AND DISCUSSION 2](#_Toc305755540)

[5.1. Tables 2](#_Toc305755541)

[CONCLUSIONS AND RECOMMENDATIONS 2](#_Toc305755542)

[LIST OF REFERENCES 2](#_Toc305755543)

[APPENDIX A: GANTT CHART 2](#_Toc305755544)

[APPENDIX B: MATERIAL SAFETY DATA SHEETS (MSDS) 2](#_Toc305755545)

[A.1 Requirements for MSDS **Error! Bookmark not defined.**](#_Toc305755546)

**LIST OF FIGURES**

Figure 2.1

**LIST OF TABLES**

Table 2.1

**Chapter One**

**INTRODUCTION**

* 1. **Background of the Study**

It has been more than a year since the novel coronavirus disease (COVID-19) outbreak was first reported in Wuhan, China, yet it remains a global issue. Although several vaccines are already available to help prevent people from getting seriously ill or dying from the disease, most countries are still struggling to battle the pandemic as the virus mutations continue to arise.

In the study entitled "Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020" by Cai, Jing et al. (2020), data analysis performed has indicated an indirect transmission of the Covid-19 virus resulting from contamination of common objects, virus aerosolization in confined spaces, or spread of virus from asymptomatic infected individual. With this, the World Health Organization have recommended maintaining an inter-personal distance of 1.5 or 2 m (about 6 feet) from each other to minimize the risk of contagion through the droplets that an individual usually disseminate through their nose and mouth.

In another study entitled "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study" by Badr et al. (2020), results had shown a strong and statistically significant correlation between social distancing, quantified by mobility patterns, and reduction of COVID-19 case growth. The researchers have emphasized the role of social distancing as an effective way to mitigate COVID-19 transmission and should remain part of personal and institutional response to the pandemic.

Forecasting plays an important role during the pandemic and is highly important for effective governmental decision making, for managing supply chain resources, and for informing very difficult political decisions as, for example, imposing a lockdown or curfews.

The Inter-Agency Task Force for the Management of Emerging Infectious Diseases (IATF – EID) acts as the Philippine government's instrument to assess, monitor, contain, control, and prevent the spread and local transmission of COVID-19. They perform risk-level classifications of Provinces, Highly Urbanized Cities (HUCs), and Independent Component Cities (ICCs) upon which they implement localized lockdowns for identified critical areas.

Enhanced Community Quarantine is the most restrictive quarantine classification where strict home quarantine shall be observed in all households. People mobility will be limited to accessing essential goods, services and for work in offices or industries permitted to operate such as public and private hospitals, health, emergency and frontline services, essential goods manufacturers, and the likes. Localities under an Enhanced Community Quarantine (ECQ) are issued with quarantine pass allowing one person per household to buy essential goods or services.

The current implementation of quarantine pass is done manually on which several issues had been observed. An advisory from DILG Region V has been released on March 20, 2020, advising Provincial Directors to instruct Local Officials that the issuance of quarantine pass slips must be distributed on house-to-house basis in their areas of responsibility as releasing of Barangay pass in Barangay Halls defeats the intent of the Enhance Community Quarantine and Social Distancing. On a news article from Inquirer.net released on 23 March 2020, a barangay captain in Lanao del Sur has been arrested for “selling” passes that would allow people to leave their homes during the quarantine period.

By developing a mobile responsive application, these issues and the manual effort spent in generating and validating quarantine passes can be resolved. Augmented with decision support system, the proposed application will enable the government and residents to reduce make informed decisions aiming to reduce the transmission of COVID-19 virus.

* 1. **Statement of the Problem**

The current process of generating, issuing, and validating quarantine passes mostly involves face to face interaction defeating the purpose of the Enhance Community Quarantine. Not only that it requires a lot of time and effort, but it is also prone to exploitations by those who wanted to get pass the quarantine guidelines.

The manual process does not have proper means to validate and authenticate the quarantine pass, thus, authorities are unable to apprehend any offenders. These loopholes may eventually result to an increase in the number of unauthorized persons outside of residence that may forfeits the effectiveness of the quarantine guideline.

The policies alone set by IATF during Enhanced Community Quarantine become less efficient if not implemented with appropriate tools and remain unmonitored.

**1.3 Objective of the Study**

**1.3.1 General Objective**

The main objective of this capstone project is to develop a mobile responsive web application that would automate the process of generation and validation of quarantine pass and include crowd forecasting for decision support system.

**1.3.2 Specific Objectives**

Specifically, this capstone project seeks to achieve the following objectives:

1. To provide automated quarantine passes for residents that will be validated at essential establishments by using Quick Response Code scheme.
2. To provide information to residents about the current crowd count of essential establishments.
3. To provide users with crowd forecast of essential establishments for upcoming days by using Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model on the mobility data of quarantine pass holders.
   1. **Significance of the Study**

Results obtained from this capstone project will benefit the following stakeholders:

**Local Government Officials.** This tool will remove the manual process of generating, issuing, and validating of quarantine passes during the implementation of Enhance Community Quarantine, enabling them to focus on other important activities. The application can act as an additional support system in ensuring that the guidelines of Enhanced Community Quarantine are implemented accordingly and by providing crowd forecasts for decision support system.

**Residents.**  By having a mobile accessible application, individuals can easily obtain a quarantine pass without requiring face to face interaction and allow them to decide on when is the best time to procure essentials goods based on the information provided by the system.

**Future Proponents.**  This capstone project could serve as a good reference material for students who are to conduct study of the same nature.

**1.5 Scope and Limitations**

To set boundaries on this capstone project, the proponent would focus on developing a mobile responsive web application which would automate the manual process of generation and validation of barangay quarantine pass embedded with crowd forecasting.

This capstone project will use a dummy data for residents, barangay, and essential establishments information to emphasize on how the automated system can provide usable crowd forecasts for decision making.

To limit the scope, the system to be developed will not cover other types of quarantine passes such as Special Quarantine Pass, Self-Employed Pass, 1 Week Transit Pass, 1 Day Transit Pass and the likes but only on Quarantine Pass intended for access to essential establishments.

**1.6 Definition of Terms**

**Forecasting**

Forecasting is about predicting the future as accurately as possible, given all the information available, including historical data and knowledge of any future events that might impact the forecasts.

Forecasting should be an integral part of the decision-making activities of management, as it can play an important role in many areas of a company. Modern organizations require short-term, medium-term, and long-term forecasts, depending on the specific application.

**Mean**

The arithmetic mean, more commonly known as “the average,” is the sum of a list of numbers divided by the number of items on the list. The mean is useful in determining the overall trend of a data set or providing a rapid snapshot of your data. Another advantage of the mean is that it’s very easy and quick to calculate.

**Regression**

Regression models the relationships between dependent and explanatory variables, which are usually charted on a scatterplot. The regression line also designates whether those relationships are strong or weak. Regression is commonly taught in high school or college statistics courses with applications for science or business in determining trends over time.

**Chapter Two**

**REVIEW OF RELATED LITERATURE**

*(Each chapter should begin in a new page.)*

**2.1 Referencing within the Text**

*(Note: In the manuscript, the section and sub-section titles are given by the researchers.)*

Use the author-year style of referencing within the text. For example, “Bioleaching can also be called bioextraction, biorecovery, biosolubilization or biohydrometallurgy (Bosecker, 1997).”

If there are two authors, both should be mentioned. For example, “… the processes for ammonia synthesis and the hydrogen from synthesis gas via the steam reforming route both release a large quantity of greenhouse gases into the air (Wood and Cowie, 1998).”

If there are three or more authors, use the first author followed by “et al.”. For example, “Fly ash is used for blended cement, raw material for concrete production, construction materials and embankments, mining applications, water stabilization and fillers and aggregates (Meawad et al., 2010)”. Please take note of the punctuation.

An author may also be cited in the following manner: “In the study conducted by Allorde et al. (2006), the Philippines coal has one of the lowest quality and is very undesirable as raw material for combustion with 14% ash content.”

**Chapter Three**

**THEORETICAL FRAMEWORK**

**3.1 Quarantine Pass Manual Process**

Upon issuance of IATF resolution declaring high-risk Provinces, Highly Urbanized Cities (HUCs), and Independent Component Cities (ICCs) under localized Enhanced Community Quarantine, the Local Government Units (LGUs) are advised to implement lockdown procedures limiting people mobility for essential activities only such as buying of food, groceries, medicines etc.

To ensure the public’s access to essential goods quarantine passes are issued by respective LGUs to residents within their jurisdiction. Only one quarantine pass will be given per household.

The generation and issuance of quarantine pass is purely manual process done by LGU officials. Quarantine pass templates are printed out on a piece of paper and distributed either in the Barangay halls or delivered on house-to-house basis in their areas of responsibility.

At checkpoints/essential establishment, the quarantine pass will need to be presented to the checkpoint/barangay officials for manual verification.

**3.2 Quick Response Code**

QR Code is a two-dimensional symbol. It was invented in 1994 by Denso, one of major Toyota group companies, and approved as an ISO international standard (ISO/IEC18004) in June 2000. This two-dimensional symbol was initially intended for use in production control of automotive parts, but it has become widespread in other fields. Now QR Code is seen and used every day everywhere in Japan for the following reasons: • Several characteristics superior to linear bar codes: much higher data density, support Kanji/Chinese character, etc. • It can be used by anybody free of charge as Denso has released the patent into the public domain. • Data structure standard is not prerequisite for current usages. • Most mobile phones in Japan equipped with cameras that enable reading of QR Codes can access Internet addresses automatically by simply reading a URL encoded in the QR Code.

**3.3 Autoregressive Integrated Moving Average (ARIMA) Model**

Autoregressive Integrated Moving Average (ARIMA) model uses time-series data and statistical analysis to interpret the data and make future predictions. The ARIMA model aims to explain data by using time series data on its past values and uses linear regression to make predictions.

An ARIMA model can be understood by outlining each of its components as follows:

* Autoregression (AR): refers to a model that shows a changing variable that regresses on its own lagged, or prior, values.
* Integrated (I): represents the differencing of raw observations to allow for the time series to become stationary (i.e., data values are replaced by the difference between the data values and the previous values).
* Moving average (MA): incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

**ARIMA Parameters**

Each component in ARIMA functions as a parameter with a standard notation. For ARIMA models, a standard notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to indicate the type of ARIMA model used. The parameters can be defined as:

* **p**: the number of lag observations in the model; also known as the lag order.
* **d**: the number of times that the raw observations are differenced; also known as the degree of differencing.
* **q**: the size of the moving average window; also known as the order of the moving average.

The parameters take the value of integers and must be defined for the model to work. They can also take a value of 0, implying that they will not be used in the model. In such a way, the ARIMA model can be turned into:

* ARMA model (no stationary data, d = 0)
* AR model (no moving averages or stationary data, just an autoregression on past values, d = 0, q = 0)
* MA model (a moving average model with no autoregression or stationary data, p = 0, d = 0)

Therefore, ARIMA models may be defined as:

* ARIMA(1, 0, 0) – known as the first-order autoregressive model
* ARIMA(0, 1, 0) – known as the random walk model
* ARIMA(1, 1, 0) – known as the differenced first-order autoregressive model, and so on.

Once the parameters (p, d, q) have been defined, the ARIMA model aims to estimate the coefficients α and θ, which is the result of using previous data points to forecast values.

**Limitations of the ARIMA Model**

Although ARIMA models can be highly accurate and reliable under the appropriate conditions and data availability, one of the key limitations of the model is that the parameters (p, d, q) need to be manually defined; therefore, finding the most accurate fit can be a long trial-and-error process.

Similarly, the model depends highly on the reliability of historical data and the differencing of the data. It is important to ensure that data was collected accurately and over a long period of time so that the model provides accurate results and forecasts.

**3.4 Box- Jenkins Methodology**

The Box-Jenkins approach to modelling ARIMA processes was described in a highly influential book by statisticians George Box and Gwilym Jenkins in 1970. An ARIMA process is a mathematical model used for forecasting. Box-Jenkins modelling involves identifying an appropriate ARIMA process, fitting it to the data, and then using the fitted model for forecasting. One of the attractive features of the Box-Jenkins approach to forecasting is that ARIMA processes are a very rich class of possible models and it is usually possible to find a process which provides an adequate description to the data.

The original Box-Jenkins modelling procedure involved an iterative three-stage process of model selection, parameter estimation and model checking. Recent explanations of the process (e.g., Makridakis, Wheelwright and Hyndman, 1998) often add a preliminary stage of data preparation and a final stage of model application (or forecasting).

1. Data preparation involves transformations and differencing. Transformations of the data (such as square roots or logarithms) can help stabilize the variance in a series where the variation changes with the level. This often happens with business and economic data. Then the data are differenced until there are no obvious patterns such as trend or seasonality left in the data. “Differencing” means taking the difference between consecutive observations, or between observations a year apart. The differenced data are often easier to model than the original data.
2. Model selection in the Box-Jenkins framework uses various graphs based on the transformed and differenced data to try to identify potential ARIMA processes which might provide a good fit to the data. Later developments have led to other model selection tools such as Akaike’s Information Criterion.
3. Parameter estimation means finding the values of the model coefficients which provide the best fit to the data. There are sophisticated computational algorithms designed to do this.
4. Model checking involves testing the assumptions of the model to identify any areas where the model is inadequate. If the model is found to be inadequate, it is necessary to go back to Step 2 and try to identify a better model.
5. Forecasting is what the whole procedure is designed to accomplish. Once the model has been selected, estimated and checked, it is usually a straightforward task to compute forecasts. Of course, this is done by computer.

Although originally designed for modelling time series with ARIMA processes, the underlying strategy of Box and Jenkins is applicable to a wide variety of statistical modelling situations. It provides a convenient framework which allows an analyst to think about the data, and to find an appropriate statistical model which can be used to help answer relevant questions about the data.

**Chapter Four**

**DESIGN AND METHODOLOGY**

This proponent of this capstone project used prototype method in delivering the objectives of this project.

Diagram

Description automatically generated

**Figure 4.1 Prototype Model**

Figure 4.1 shows the Prototype Model used by the proponent as a guide in developing the project entitled SAFEPASS: A QUARANTINE PASS TOOL WITH CROWD DENSITY PREDICTION USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) which is an example of the System Development Life Cycle (SDLC) model.

Using this model, it will enhance the usability, design quality of the proposed application and it will also make the development process more cost-efficient since the development cycle becomes shorter.

The phases of the prototype model involves the following steps:

**4.1 Requirements Modeling**

In Molino IV, a barangay in the city of Bacoor, Cavite, the process of implementing home quarantine passes is done manually by designated LGU or Barangay officials. Quarantine pass templates are printed on a paper with the details of their Barangay and some blank fields to be manually populated by the individual who will utilize it. The dissemination process of quarantine pass varies from one Barangay to another. Either they require their residents to claim it from their Barangay Hall or for the Barangay officials to deliver on house-to-house basis.

In the 2015 census done by Philippine Statistics Authority, the estimated population in Barangay Molino IV, Bacoor Cavite is 51, 362 and the average number of people per household is 4. Given this data, there would be around 12,841 quarantine passes to be issued to each household on a manual basis requiring a lot of time and effort.

At each designated checkpoints or basic commodity establishments, residents are required to present their quarantine pass allowing them access to the establishment. With the high volume of quarantine pass holders and limited establishments operating, there is a possibility of having crowd congestion in an establishment.

Graphical user interface, application

Description automatically generated

**Chapter Five**

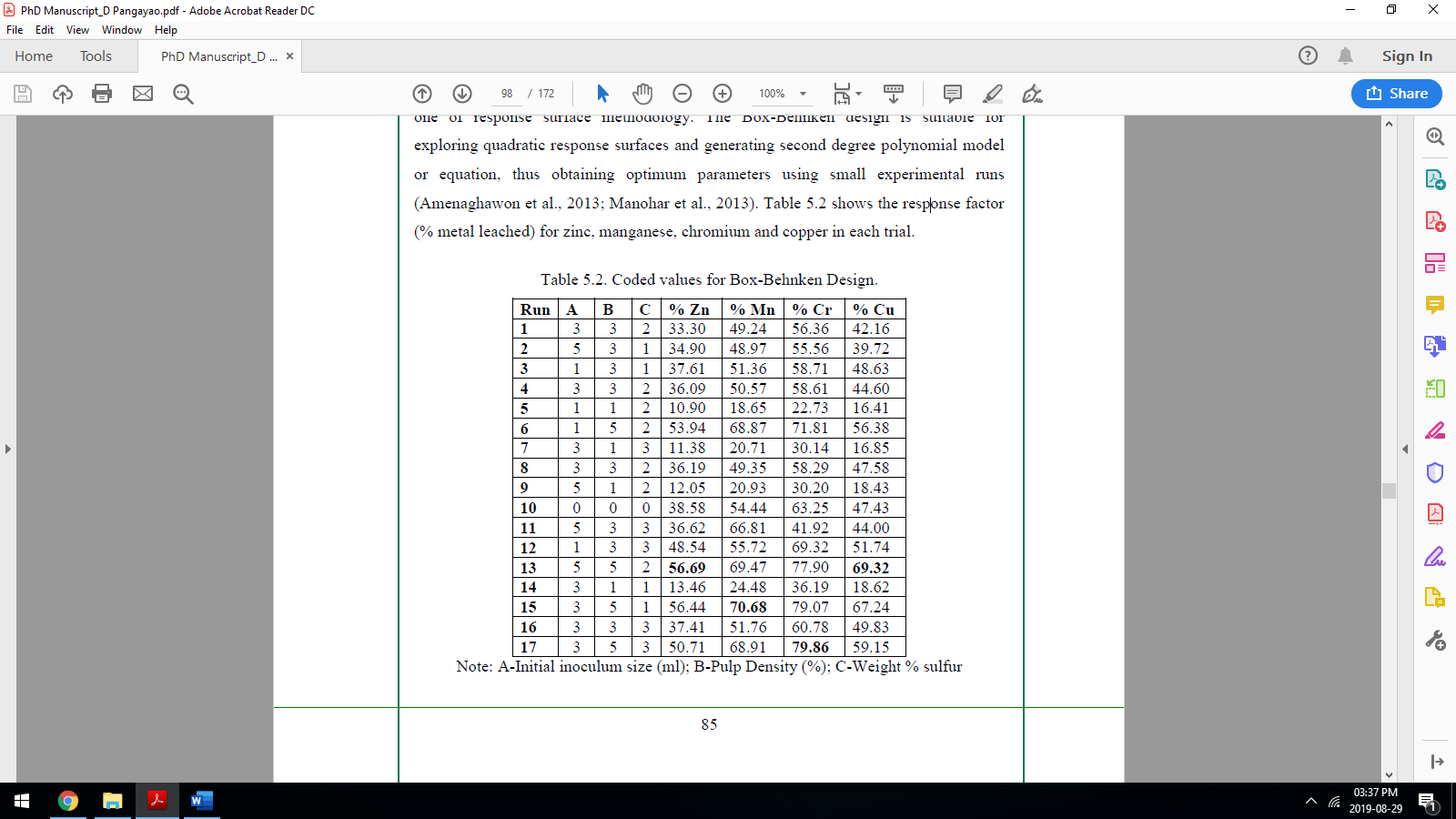
**RESULTS AND DISCUSSION**

**5.1. Tables**

All tables should be explained in the text. Figures should be inserted after they are first mentioned in the text. They should be placed as close as possible to the text while considering layout and paging. Tables should be located on a single page.

Each table should be accompanied by a table number and table heading, centered above the table. Tables are numbered consecutively within the chapter. (See Table 5.1.)

Table 5.1 Coded values for Box-Behnken Design



If a table is to be oriented in “landscape” format, then it should be oriented such that the table is facing right.

**Chapter Six**

**CONCLUSIONS AND RECOMMENDATIONS**

The thesis proposal and the thesis follow the same format, with the following exceptions:

1. The title page is changed to reflect that this is a thesis proposal.
2. A “recommendation sheet” with the adviser(s) signature is attached instead of an “approval sheet”. Signatures of the panel members, the chair and the dean are not required.
3. There is no “Acknowledgements” section in the proposal nor is there an “Abstract”.
4. There is no “Results and Discussion” section in the proposal, although a “Preliminary Results” section may be included.
5. There is no “Conclusions and Recommendations” section in the proposal.
6. The Appendix in the proposal contains additional data, information or picture relevant to the research. In the thesis, the Appendix contains supporting material for the thesis, such as raw data, calibration curves, etc.

**LIST OF REFERENCES**

*[The American Psychological Association (APA) style is used for the references. All references are listed in alphabetical order. All references are listed together. The references are not justified, not indented in the first line and with 0.5” indentation in the succeeding line. References are single space within the paragraph and are separated by 1 line.]*

Ahmaruzzaman, M. (2010). A review on the utilization of fly ash, Progress in Energy and Combustion Science, 36, 327-363.

Ajemba, R. O. and Onukwuli, O. D. (2012). Dissolution kinetics and mechanism of reaction of Udi clay in nitric acid solution. American Journal of Scientific and Industrial Research, 3, 115 – 121.

Department of Energy (DOE) (2011). www.doe.gov.ph/list-existing-power-plant

Department of Energy (DOE) (2015). www.doe.gov.ph/philippine-power-statistics

Fogler, H. S. Elements of Chemical Reaction Engineering, Fourth Edition. Pearson Education Inc. 2006

Gallardo, S., Pangayao, D., Ronquillo, R. and Salido, B. (2012). Chemical Characterization of Coal Ash from a Typical Coal-Fired CFB Boiler. 5th ASEAN Environmental Engineering Conference (AEEC), Ho Chi Minh City, Vietnam.

Levenspiel, O. Chemical Reaction Engineering, Third Edition. John Wiley and Sons. 1999

Method 3050B: Acid Digestion of Sediments, Sludges and Soils, US EPA (1996)

Pangayao, D. C. (2012). Leachability of Chromium from Coal Bottom Ash using Continuous Column Test. Directed Research, De La Salle University – Manila.

van Hullebusch, E. D., Yin, N-H, Seignez, N., Labanowski, J., Gauthier, A., Lens, P. N. L., Avril, C. and Sivry, Y. (2015). Bio-alteration of metallurgical waste by Pseudomonas aeruginosa in a semi flow-through reactor. Journal of Environmental Management, 147, 297-305.

**APPENDIX A: SAMPLING OF RAW MATERIALS / Simulation**

**APPENDIX B: ICP – OES AND AAS ANALYSIS / source code**