EARTH'S MAGNETIC FIELD AND ANOMALIES

by

NUTCHA JITCHANVICHAI

COMMITTEE MEMBERS:

Professor James Jones

Professor Krzysztof Ostaszewski

Professor Xing Wang

EARTH'S MAGNETIC FIELD AND ANOMALIES

by

NUTCHA JITCHANVICHAI

Under the Supervision of

PROFESSOR JAMES JONES

PROFESSOR KRZYSZTOF OSTASZEWSKI

PROFESSOR XING WANG

A Project Submitted in Partial
Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Mathematics

ILLINOIS STATE UNIVERSITY

2021

EARTH'S MAGNETIC FIELD AND ANOMALIES

By

NUTCHA JITCHANVICHAI

ABSTRACT

The project aims to identify and measure the economic risk focusing on the communication satellites associated with the growing of the South Atlantic Anomaly as well as to find the possible mitigation to immunize all possible economic risk in the future. The Monte Carlo simulation is an approach used to calculate the probability distributions of the economic performance indicator, which is the net present value, and measure the risk from the distribution using the value at risk as the risk indicator. The results showed that the lower variance of the loss variable can perform the better average and the standard deviation of the net present value. The change of loss variable will not affect the value at risk. In addition, adding shield of the satellites is considered as a possible mitigation to prevent the satellites from damages. The solar flares were included as the additional topic for the project. The severities from the flares, which referred to the minimum and maximum energy produced by the flares, were forecasting using the time series analysis. The final models can be used to predict pretty well, but the error of the prediction should be concerned, especially for the maximum energy. In order to improve the accuracy of the models, a larger sample size should be obtained to solve this error issue.

KEYWORDS: South Atlantic Anomaly; Monte Carlo Simulation; Solar Flares.

CONTENTS

TABLES

FIGURES

CHAPTER I: BACKGROUND OF THE STUDY

The Earth's Magnetic Field and Anomalies is a project under Innovation Consulting Community and Katie School of insurance and Risk Management at Illinois State University. They agreed to the presentation of the methodology. This paper does not present the work done in the consulting project, but rather the methodology used that is studied and worked to implement. The data used in this project are simulated based on the obtained public information. The Katie School's client for the project is Verisk. The objective is to address the following two key questions:

1. Identify and measure economic risk focusing on the communication satellites associated with the growing of the South Atlantic Anomaly (SAA)

2. Find the possible mitigation plan to immunize all possible economic risk in the future. Identifying and measuring these risks associated with the South Atlantic Anomaly (SAA) can lead to the better understanding of the growing phenomenon as well as its effects to the space technologies. Moreover, the possible mitigation plan is identified to immunize these risks in the future.

1.1 INTRODUCTION

Every life on the earth has survived from the solar winds and cosmic radiation because we have things called the Van Allen belt, which is held around the earth by the magnetic field. It is defined as

> two concentric rings of energetic particles surrounding the planet. The inner belt composed predominantly of protons from the Sun, and the outer belt contained mostly electrons from the interactions of cosmic rays in Earth's atmosphere. (Johnson-Groh, 2018, para. 4)

Therefore, the life on earth is not affected by the radiations since the high energetic particles were trapped inside the belt. However, there is an area between South America and the southern Atlantic Ocean called the South Atlantic Anomaly (SAA) where comes closer to the surface than other areas. Inside the dent has a lot of high-energy protons from solar winds and the radiation.

This phenomenon was firstly occurred 8 to 11 million years ago since the volcanic rock samples around the South Atlantic Ocean preserve an ancient record of magnetic data till these days. Furthermore, it leads to increase the flux of energetic particles using the geomagnetic intensity (tesla: T) as the index. According to the article "*South Atlantic Anomaly: 2015 through 2025*" (NASA, 2020), the observed data from 2015-2020 founds that the SAA region is expanding westward and continuing to weaken in intensity. This means the area started to split from a single valley into two cells. Splitting of the SAA means there will be more areas that can damage the objects passing through these areas. After modeling this data for the year 2025, it will show the split continuing in the future.

According to NASA (2009), "the space industry contains three main sectors: scientific, military, and commercial." Focusing on the commercial sector, this sector plays an important role in order to provide the positive economic benefit to the national economy. These technologies, including satellites, spacecrafts, international space station, etc., are the important products of the industry. The satellites these days can be used for many different purposes, such as tracking the weather or catastrophe, taking the photos of the planets or galaxies in having the better understanding about the universe, and using for communication, such as TV signals, telephones, the Internet, 5G, GPS and etc.

Although, the SAA does not affect life on the earth, either data or equipment of the satellites and spacecrafts can be damaged by the high-energy particles. Therefore, all satellites

and spacecrafts have to shut down their systems while they are passing these areas, in order to avoid either the temporary or permanent loss which may occur. This means the more lobes will exist, the more damage to the technologies' assets as well as challenge for satellites and humans in space mission will have. This is the reason why monitoring as well as modeling the phenomenon to find the possible mitigation plan for the future is essential.

1.2 LITERATURE REVIEW

1.2.1   The Occurrence of South Atlantic Anomaly

Schaefer et al. (2016) presented the new model to identify the location of the South Atlantic Anomalies (SAA). The model shows that the SAA is located in the low earth orbit (LEO). As it was mentioned previously, the SAA can damage the satellites, spacecrafts and man-made objects passing through SAA's area. In addition, the damages can be occurred, both of the big events and the smaller one. Due to the big damage from SAA, the satellites can be decommissioned. The decommissioned satellites are still in the orbit; however, no one can command or control them. While the smaller damage occurs, the astronauts can repair them in space by going on spacewalk (Wild, 2020). The objects from repairing those damages will be defined as debris.

Both of the decommissioned satellites and debris occurring due to both the SAA and not due to the SAA, can be called as space junk, which can lead to collisions of the satellites as in 2009 (Capital Technology University, 2020). The collision may identify as a secondary exposure from this phenomenon. However, O'Callaghan (2020) mentioned that the objects in the lower orbit can return to the atmosphere quicker than the higher one, which will mostly take a few years and be burnt up before reaching the ground.

1.2.2 The Satellites' Procedure

Campbell (2017) explained the communication between satellites and earth that they are using the radio wave to send signals to the antennas on the Earth. These antennas performed as both sender and receiver information normally from/to the Earth. Therefore, we might notice it almost immediately if something occurs since the sender and/or receiver is no longer working. According to the information above, it means all areas will be affected when the satellites stop working as SAA may affect the system of the satellites or spacecraft.

There is a great example case of the effect of malfunctioning satellite happened to AT&T company in 1997. Peredo and Thompson (1997) stated that AT&T was a company providing the services of the signal beam for several customers, such as ABC, Fox and PBS. These AT&T's customers were broadcasting the TV shows. Therefore, after the satellite went down on Saturday morning in 1997, AT&T had to switch all services for all customers to other satellites. This means AT&T had to cover a huge amount of money in order to remaining their services.

1.2.3 Evaluation of Economic Risk

Consideration of economic risks in a comparative analysis of nuclear technologies with different maturity levels (2019) and Economic risk analysis of decentralized renewable energy infrastructures - A Monte Carlo Simulation approach (2013) are the researches talking about the economic risk of the nuclear as well as renewable infrastructure technologies. In addition, both nuclear and renewable technologies are innovative technologies like satellite and spacecraft. Therefore, the methodology introduced on the research may be applied to the anomalies well.

The evaluation of risk indicators was identified on the research "Consideration of economic risks in a comparative analysis of nuclear technologies with different maturity levels" (2019). According to this research, there are several steps to identify economic risk. To begin with the

probability distributions of the economic performance indicator calculated by Monte Carlo method. The calculation of the distribution is used to evaluate risk indicators. There are several economic performances to get the distribution, such as net present value (NPV), present value (PV), internal rate of return (IRR), discounted payback period (DPP) and leveled cost (LC).

Then, the distribution will evaluate risk using the indicator, such as Value at Risk (VaR), Expected Shortfall (ES), Tail Value at Risk (tVaR), Expected value (ME) and Most probable value (MP). Both of the economic performance indicator and the risk indicator can be selected based on the purpose of the project. For this research, all parameters, which include external conditions and characteristics of technologies and minimum costs each year, are assumed to be distributed uniformly. The estimation of NPV as the economic performance indicator by Monte Carlo generating 10,000 times and the confidence level α for the calculation of VaR as the risk indicator equal to 95%. Moreover, the same methodology of the previous research was also performed on the research "Economic risk analysis of decentralized renewable energy infrastructures - A Monte Carlo Simulation approach" (2013).

CHAPTER II: DATA AND METHODOLOGY

2.1 DATA

The data used in the project are simulated. The variables and sources of the simulation are exhibited in Table 1.

Table 1: Explanatory variables and publicly available sources used for simulating data

Variable	Description (Unit)	Data Source
R	The annual revenue from three companies that providing the communication satellite services (USD)	- https://www.wsj.com/market-data/quotes/GSAT/financials/annual/income-statement - https://www.wsj.com/market-data/quotes/IRDM/financials/annual/income-statement
E_o	The annual other expense for communication satellite (USD)	- https://www.wsj.com/market-data/quotes/SATS,EK ,UEIC,VOXX,MCZ,NIVS,LOJN,LRAD,IGOI/ financials/annual/income-statement
L	The loss occurring when the anomalies damaged a communication satellite	- https://www.space.com/6839-space-forecast-predicts-satellite-production-boom.html
I	The initial investment sum of launching communication satellites (USD)	- https://www.space.com/6839-space-forecast-predicts-satellite-production-boom.html - https://talkingpointsmemo.com/idealab/satellites-earth-orbit

Firstly, three different companies, which are Globalstar Inc., Iridium Communications Inc. and EchoStar Corp., are selected among all companies who offer the communication satellite service. The average annual revenue (R) as well as other expense (E_0) is calculated from these mentioned companies' income statement using sale/revenue and cost of goods sold including depreciation and amortization, respectively.

The average annual revenue (R) in million USD is 859.3873 in 2019, followed by 805.3737 in 2018, 815.5687 in 2017, 780.1670 in 2016, and 1,215.2893 in 2015. In addition, the average annual other expense (E_0) in million USD is 646.3520 in 2019, followed by 586.3690 in 2018, 540.4470 in 2017, 477.8237 in 2016, and 867.8147 in 2015. Therefore, the range of annual revenue is between 780.1670 to 1,215.2893 million USD as well as the range of annual other expense is between 477.8237 to 867.8147 million USD.

Secondly, the initial investment cost (I) is the total cost of launching the communication satellite. The initial cost is calculated from two portions, which are the cost of launching and the cost of producing the satellite. According to the public sources shown in Table 1, the SpaceX's cost of launching is approximately 11.3 million USD per ton and generally the communication satellites will weigh about 1-6 tons. In addition, the cost of making a communication satellite is 51 million USD on average. Hence, the higher satellite's weight will drive the initial cost to be larger.

Lastly, the loss (L) will be considered as a parameter to identify the loss when the communication satellites are damaged by the anomalies. The occurring damage will be assumed that it occurs at year 10. The loss is the permanent loss, which will decommission the damaged satellite, since the temporary loss is on the other expense portion. This means the permanent loss

will cost the satellite's price, which approximately 99 million USD and the cost of making a communication satellite equals 51 million, according to Table 1.

2.2 METHODOLOGY

2.2.1 Monte Carlo Simulation

IBM Cloud Education (2020) explained the Monte Carlo simulation as the approach used to simulate the possible outcomes and build a model of possible outcomes by leveraging a probability distribution, such as a uniform or normal distribution, for any variable that has inherent uncertainty. In addition, it will be repeated the calculation using the different set of the random variables and the replication can be more than a thousand times to produce a large number of outcomes. They also identified three steps for applying the Monte Carlo simulation as

1) Set up the predictive model, identifying both the dependent variable to be predicted and the independent variables (also known as the input, risk or predictor variables) that will drive the prediction.

2) Specify probability distributions of the independent variables. Use historical data and/or the analyst's subjective judgment to define a range of likely values and assign probability weights for each.

3) Run simulations repeatedly, generating random values of the independent variables. Do this until enough results are gathered to make up a representative sample of the near infinite number of possible combinations (IBM Cloud Education, 2020, para. 6).

One of the simplest examples of the Monte Carlo Simulations is the calculation of the probability that two unbiased dice are getting heads. The simplest way to calculate this probability is that you can roll the unbiased dice, such as 50 or 100 times, and simply count the number of both dice are headed. After that, the probability of getting two heads is the number of

both dice are heads divided by the total occurrences and the results may not be accurate as in the theory. However, the Monte Carlo Simulation can help improve the accuracy as it will repeat the rolling dice for 1,000 times or more. This is because the more independent variables enter into the model, the more precise of predicting the dependent variables is.

Kenton (2020) stated that the Monte Carlo simulations help to explain the impact of risk and uncertainty in prediction and forecasting models for the simulation involves assigning multiple values of an uncertain variable to achieve multiple results and then to average the results to obtain an estimate. This technique can be used in many fields of studies, such as finance, engineering, science and etc. However, there is the limitation of this simulation as they assume perfectly efficient markets.

2.2.2 Economic Performance Indicator

As it mentioned in the evaluation of economic risk of literature review in chapter 1. The economic performance is calculated to identify its probability distribution. According to Monte Carlo simulation, the independent variables to calculate the economic performance will be recalculated for 10,000 times. Although there are several performance indicators as mentioned before, the performance indicator used in this project are net present value (NPV) and internal rate of return (IRR). The net present value (NPV) is the present value of the cash flows of the required rate of return of a project compared to the initial investment (Gallo, 2014). The formula of NPV is exhibited in (1) and the summary of independent variables is shown in table 2.

$$NPV = -I + \sum_{k=1}^{10} \frac{R - E_0}{(1 + i)^k} - \frac{L}{(1 + i)^{10}} \qquad (1)$$

where I = the total cost for launching the communication satellite (in million)

= cost of launching + cost of producing

Cost of launching = weight of the satellite in ton * cost of launching per ton

R = the average annual revenue (in million)

E_0 = the average annual other expense (in million)

i = the interest rate to discount the annual

k = the number of years

L = the loss from the damaged satellites (in million)

p = the varying number for variance of the loss

In addition, Fernando (2020) explained about the internal rate of return (IRR) that it is a discount rate (i) making the net present value (NPV) of all cash flows equal to zero. It used to estimate the profitability of potential investments. This means IRR and NPV depend on the same formula. The formula of IRR is exhibited in (2).

$$0 = NPV = -I + \sum_{k=1}^{10} \frac{R - E_0}{(1+i)^k} - \frac{L}{(1+i)^{10}} \tag{2}$$

Table 2: Summary of the independent variables

Variable	Distribution	Value (in million)
I	Uniform	Cost of launching = (weight ~ Uniform [1,6]) * 11.3 Cost of producing = 51
R	Uniform	[780.1670,1215.2893]
E_0	Uniform	[477.8237,867.8147]
i	Constant	3%

10

k	Constant	1, 2, 3, 4…, 10
L	Normal	$L \sim N(\mu = 105, \sigma^2 = 105*p)$ OR $L = 0$
p	Constant	0.3, 0.5 and 0.7

2.2.3 Risk Indicator

The evaluation of risk indicators was calculated using the distribution of the economic performance indicator from the previous section. Although there are several risk indicators as mentioned before, the risk indicator used in this project is Value at Risk (VaR). Kenton (2019) identified the Value at risk (VaR) as a statistic that measures and quantifies the level of financial risk or potential losses within a firm, portfolio or position over a specific time frame.

The VaR is the downside risk with a certain level of confidence (α) to specify the worst-case scenario that can be occurred at a specific time, such as monthly, quarterly, annually and etc. The level of confidence using in this project is 95% or $\alpha = 0.05$ from the left tail. The question that the VaR answers is how much the company' NPV will be in the next year, in case the SAA damaged their communication satellites with a 95% level of confidence.

2.2.4 Software

The software used in this project is RStudio version 1.3.1073. The R packages are sn, jrvFinance and cvar. The sn is used to generate the random variables for skew-normal distribution. The jrvFinance and cvar used for calculating the internal rate of return and value at risk, respectively.

CHAPTER III: RESULTS AND DISCUSSION

According to table 2, the net value of annual revenue and annual other expense is generated for ten years. In addition, there are two portions of the loss occurring when the anomalies damaged a satellite. First, all occurred loss is normally distributed with three different p, which will affect the standard deviation of the loss variables and assumed that it occurs at year 10. Secondly, the occurrence of loss is randomly positive-skew distributed. Therefore, the loss can be either 0 or $N(\mu = 105, \sigma^2 = 105 * p)$ based on the occurrence value, which is considered as 0 if the occurrence value is less than the average of the occurrence.

The net value, the loss and the initial investment sum are used to calculate Net Present Value (NPV) at the interest rate equals 3%, Internal Rate of Return (IRR) and Value at Risk (VaR) to measure the economic risk. The Value at Risk (VaR) of Net Present Value (NPV) was calculated using the historical method at the 95% level of confidence as all data are generated before calculated the Value at Risk.

Table 3: The result of the Monte Carlo simulation – three scenarios

Scenario no.	p	The average of NPV	The standard deviation of NPV	IRR	VaR at 95%
1	0.3	2,647.620	50.96746	3.004631	-2,726.106
2	0.5	2,647.618	51.01740	3.004631	-2,726.106
3	0.7	2,647.616	51.06706	3.004631	-2,726.106

The Monte Carlo Simulation will be carried out with 10,000 samples. Furthermore, the first ten simulated data are exhibited in figure 1 (Appendix A). The result of the Monte Carlo

simulation is the result of the average of NPV, IRR and VaR at the different standard variables

shown in table 3. After three scenarios' NPV are calculated, the histogram plots of all NPV are

exhibited in figure 2, figure 3 and figure 4 (Appendix A). They show that all three plots are

similar and probably normally distributed.

According to table 3, the average of NPV and IRR are positive. This means the earning

will be profitable. Furthermore, it shows that there is not much difference of the average of NPV

the standard deviation of NPV, IRR and VaR among three different p. This means that the

different standard deviation of loss does not much affect to the NPV, IRR and VaR. However,

the first scenario provided the highest average of NPV and the lowest standard deviation of NPV

while the IRR and VaR at the 95% level of confidence are equal for all scenarios.

Focusing on the first scenario's NPV and VaR, the average of Net Present Value equal to

2,647.62 million USD indicated that the average earnings will be about 2,647.62 million in the

next ten years. Then, the first scenario's VaR of NPV at 95% level of confidence equal to

-2,726.106 million means there is 95% confidence that we will not lose more than 2,726.106

million over the next ten years. Moreover, the calculation of VaR using the Gaussian method

was also performed and the results are the same as the one using the historical method. This may

assure that the average of NPV is normally distributed.

CHAPTER IV: CONCULSION

There are two key questions that this project would like to answer. First, it is to identify and measure economic risk focusing on the communication satellites associated with the growing of the South Atlantic Anomaly. The result in chapter 3 shows that the scenario no. 1, which is p = 0.3, performed the best among all three scenarios. This means the average and the standard deviation of Net Present Value will be the best if the variance of the loss variable is low. In addition, the 95% value at risk have no effect from the change of loss variable. However, the maximum loss with a 95% level of confidence that is probably occurring due to the phenomena is slightly high. Hence, the larger phenomenon will be expanding, the higher loss might be increasing.

Next, the second one is to find the possible mitigation to immunize the possible economic risk. The shield of the satellites is considered as this can prevent the satellites from damage due to the anomalies. However, adding the shield on a satellite can increase the weight of the satellite, which leads to the increase of the initial investment sum (I). Thus, adding the shield may create negative impact of the net present value. In order to answer this question, the fixed initial investment (I), which are 1, 2, 3, 4, 5 and 6 tons, is considered with the scenario no. 1 in the Monte Carlo simulation. This is because we would like to identify whether the change of initial investment sum would impact the Net Present Value. The result exhibited on the table 4 (Appendix A).

It shows that the standard deviation of net present value will not be affected by the change of initial investment sum. Less weight of satellites will provide higher average net present value; however, the increasing the weight does not create that high impact of the change of net present value. Therefore, adding the lightweight shield on the satellites can probably be

satisfied. Furthermore, the real data should be obtained and applied the methodology to ensure

whether or not there are any different. This can reflect more precise result since the Monte Carlo

takes into account all scenarios, which may not happen in the real data.

CHAPTER V: SOLAR FLARES

Space weather is a phenomenon occurred by all activities on the Sun's surface. The solar storms are one of the space weather's causes. NASA (2020) explained that there are four types of solar storms happening on the Sun, which are solar flares, coronal mass ejections (CME), high-speed solar wind and solar energetic particles. Among all kinds, the solar flares are considered the massive explosion in the solar system, which can send tons of energy through space at the speed of light. Gaughan (2017) stated that the solar flares take about 8 minutes from the Sun to Earth. Although the flares are the form of highly energetic particles, life on Earth may not be affected because of the Van Allen belt. However, there is a worst-case scenario that solar flare shut down the electricity for nine hours in Quebec, Canada, which affected 6 million people in 1989. (Gaughan, 2017)

The quantification of the flares associated with the number of sunspots, which are the dark spots on the Sun's surface compared to other areas surround them. The flares are happening followed the cycle, which is approximately 11 years. At the beginning of the cycle, the number of sunspots is minimum and will rise afterward. The peak number of spots is in the middle of the cycle and it will be a minimum number of spots again at the end of the cycle. Therefore, the more accurate prediction of the flares means the better mitigation plan that people can prepare in order to prevent all possible risks.

5.1 TIME SERIES ANALYSIS

For the solar flare forecasting, the dataset is from Kaggle (2021). The variables are the start data of the flares, the number occurrence of the flares, the minimum and maximum energy produced by the flare (kV: kilo Volt). The data set of the 4,581 data points was collected from February 2002 to December 2016. For the time series analysis of solar flares, the training set is

from February 12, 2002 to December 31, 2015. This contained 4,343 data points without the extreme data points, which are the outliers. The training set will be used for identifying two appropriated models for these time series data. The test set is from January 1, 2016 to December 27, 2016. This contained 235 data points and is used to evaluate the final model.

As the figure 5 (Appendix B), it shows that this time series of the number occurrence of the flares might be stationary since the plot does not show any explicit trends. In addition, the number occurrence of the flares might be seasonal, which is the same as the mentioned cycle about 9-11 years; however, the minimum and maximum energy does not show any cycles. Hence, it will be more beneficial to forecast the minimum and maximum energy produced by the flares.

5.1.1   Model Specification

Firstly, the Dickey-Fuller Unit-Root test on figure 6 and 7 (Appendix B) show that the minimum and maximum energy are stationary since both p-values are smaller than 0.01. Next, the ACF plot, the PACF plot, the EACF and the subset method are considered to select two possible models. As the ACF and PACF plot of the minimum and maximum energy exhibited in figure 8, 9, 12 and 13 (Appendix B), both shows that the seasonal pattern does not exist. Then, taking a look at the EACF plot in figure 10 and 14 (Appendix B), ARMA (1,1) will be one of the candidate models for both of the minimum and maximum energy. The second candidate model is suggested by a subset method in figure 11 and 15 (Appendix B). It is ARMA (2,3) for both of the minimum and maximum energy.

5.1.2   Model Fitting and Diagnostic

The minimum energy

As the model specification step, the possible models are ARMA (1,1) and ARMA (2,3). Firstly, the coefficient of ARMA (1,1) will be considered without the seasonal trend (figure 16). The result shows that all the coefficients are significant, which means they should be in the model. Considering the overfitting issue, the coefficient of ARMA (1,2) and ARMA (2,1) are significant (figure 17 and figure 18). However, the ARMA (2,1) will be considered model 1 because it provided the smallest AIC.

Then, we will consider three conditions of residuals analysis of the model: the residuals analysis, normality and Ljung-Box test. The first condition shows that the majority of data points are around zero-horizontal line (figure 19). The second condition testing with the QQ (figure 20) shows that the residuals of this model are not normal. Thus, the log-transformation was considered so that the model can be improved; however, it performed a worse result (figure 22) than the previous one. The third condition exhibited in figure 21 shows that there is no correlation since the p-value $= 0.4526 > 0.05$ (Fail to reject the null hypothesis).

Secondly, the coefficient of ARMA (2,3) will be considered without the seasonal trend (figure 23). The result shows that some coefficients are not significant. Considering the overfitting issue, the coefficient of ARMA (1,3) and ARMA (2,2) are significant (figure 24 and figure 25). However, the ARMA (2,2) will be considered model 2 because it provided the smallest AIC.

Then, we will consider three conditions of residuals analysis of the model: the residuals analysis, normality and Ljung-Box test. The first condition shows that the majority of data points are around zero-horizontal line, which is similar to model 1 (figure 26). The second condition

testing with the QQ plot (figure 27) shows that the residuals of this model are not normal. Thus, the log-transformation was considered, and the result is worse (figure 29). The third condition exhibited in figure 28 shows that there is no correlation since the p-value $= 0.3509 > 0.05$ (Fail to reject the null hypothesis).

The maximum energy

As the model specification step, the possible models are ARMA (1,1) and ARMA (2,3), which are the same to the minimum energy. As the same methodology in the minimum energy is applied, the results are similar to the minimum energy, which suggested ARMA (2,1) as the model 1 and ARMA (2,2) after the overfitting issue for ARMA (1,1) and ARMA (2,3) are considered, respectively (figure 30, 31, 32, 37, 38 and 39). Then, we will check three conditions of residuals analysis of the model: the residuals analysis, normality and Ljung-Box test. The first condition shows that the majority of data points are around zero-horizontal line (figure 33 and 40). The second condition testing with the QQ (figure 34 and 41) shows that the residuals of this model are not normal. Thus, the log-transformation was considered so that the model can be improved; however, it performed a worse result (figure 36 and 43) than the previous one. The third condition exhibited in figure 35 and 42 shows that the residuals are uncorrelated since the p-value $= 0.5853$ and $0.6177 > 0.05$ (Fail to reject the null hypothesis), respectively.

5.2 FINAL MODEL

In order to evaluate the final model, the test set will be used to calculate the criteria, which are MSE and PMAD. The most appropriate model should have a small MSE, and PMAD near to 0. The table 5 shows the comparison of MSE and PMAD between model1 and model2 of the minimum and maximum energy.

19

Table 5: The comparison of MSE and PMAD

	Minimum Energy		Maximum Energy	
	Model 1	Model 2	Model 1	Model 2
MSE	32.2679	31.8913	227.3252	229.5163
PMAD	0.3578	0.3942	0.3982	0.3754

Based on the table above, it shows that the final model of minimum energy is model 1, which is

ARMA (2,1): $Y_t = 1.1689Y_{t-1} - 0.1714\ Y_{t-2} + e_t - 0.9604e_{t-1} + 6.7416$ since it has

smaller PMAD. In addition, the final model of maximum energy is model 2, which is ARMA

(2,2): $Y_t = 1.4807Y_{t-1} - 0.4855\ Y_{t-2} + e_t - 1.31013e_{t-1} + 0.3392e_{t-2} + 13.7980$ since it

has smaller PMAD.

  Although the final models' MSE and PMAD, which determined the error part of the

model, are pretty large. This can imply that these final models can be predicted the severity of

the solar flares well, but still have to concern about the error of the prediction, especially for the

maximum energy. In addition, a larger sample size might help to solve this error issue in order to

improve the model.

REFERENCES

[1] Andrianov, K. (2017). Consideration of economic risks in a comparative analysis of nuclear technologies with different maturity levels. *Nuclear Energy and Technology*, *3*(2), 158–163. https://doi.org/10.1016/j.nucet.2017.05.012

[2] Arnold, Y. (2015). Economic risk analysis of decentralized renewable energy infrastructures – A Monte Carlo Simulation approach. *Renewable Energy*, *77*, 227–239. https://doi.org/10.1016/ j.renene.2014.11.059

[3] Campbell, A. (2017, August 7). *How do satellites communicate?* https://www.nasa.gov/ directorates/heo/scan/communications/outreach/funfacts/txt_satellite_comm.html Capital

[4] Capital Technology University. (2020, November 13). *Hazards of Space: How Satellite Missions Can Go Wrong*. https://www.captechu.edu/blog/hazards-of-space-how-satellite-missions-cango-wrong

[5] Dockrill, P. (2020, July 21). *The South Atlantic Anomaly May Go Back Far Longer Than Anybody Knew*. https://www.sciencealert.com/the-mysterious-magnetic-anomaly-in-the-south-atlanticexisted-millions-of-years-ago

[6] Fernando, J. (2020, November 13). *Internal Rate of Return (IRR)*. https://www.investopedia .com/terms/i/irr.asp

[7] Gallo, A. (2014, November 19). *A Refresher on Net Present Value*. https://hbr.org/2014/11/a-refresher-on-net-present-value

[8] Gaughan, R. (2017, April 25). *How Long for a Solar Flare to Reach Earth?* https://sciencing .com/ long-solar-flare-reach-earth-3732.html

[9] IBM Cloud Education. (2020, August 24). *Monte Carlo Simulation*. https://www.ibm.com /cloud/learn/monte-carlo-simulation

[10] Johnson-Groh, M. (2018, February 1). *Studying the Van Allen Belts 60 Years After America's First Spacecraft*. https://www.nasa.gov/feature/goddard/2018/studying-the-van-allen-belts-60- years-after-america-s-first-spacecraft

[11] Kaggle. (2021). *Solar Flares from RHESSI Mission*. Retrieved from https://www.kaggle .com/khsamaha/solar-flares-rhessi

[12] Kenton, W. (2019, April 18). *Value at Risk (VaR)*. https://www.investopedia .com/terms/v/var.asp

[13] Kenton, W. (2020, December 27). *Monte Carlo Simulation*. https://www.investopedia.com /terms/m/montecarlosimulation.asp

[14] May, S. (2014, February 12). *What Is a Satellite?* https://www.nasa.gov/audience /forstudents/5-8/ features/nasa-knows/what-is-a-satellite-58.html

[15] Merzdorf, J., & Johnson-Groh, M. (2020, August 17). *NASA Researchers Track Slowly Splitting 'Dent' in Earth's Magnetic Field*. https://www.nasa.gov/feature/nasa-researchers-track-slowlysplitting-dent-in-earth-s-magnetic-field

[16] NASA. (2020, August 17). *South Atlantic Anomaly: 2015 through 2025*. https://svs.gsfc.nasa.gov/cgi-bin/details.cgi?aid=4840&button=recent

[17] NASA. (2020, December 23). *Solar Storm and Space Weather - Frequently Asked Questions*. https:// www.nasa.gov/mission_pages/sunearth/spaceweather/index.html#q2

[18] NASA. (2021, January 10). *What are solar storms and how do they affect the Earth?* https:// image.gsfc.nasa.gov/poetry/ask/a10624.html

[19] NASA. (2021, January 10). *What Is the Solar Cycle?* https://spaceplace.nasa.gov/solar-cycles/en/

[20] O'Callaghan, J. (2020, November 13). *What is space junk and why is it a problem?* https://

www.nhm.ac.uk/discover/what-is-space-junk-and-why-is-it-a-problem.html

[21] Peredo, M., & Thompson, B. J. (1997, January 16). *AT&T Satellite Malfunction*. https://pwg.gsfc.nasa.gov/istp/cloud_jan97/att.html

[22] Schaefer, R. K., Paxton, L. J., Selby, C., Ogorzalek, B. S., Romeo, G., Wolven, B. C. and Hsieh S. Y. (2016), Observation and modeling of the South Atlantic Anomaly in low Earth orbit using photometric instrument data, *Space Weather, 14*, 330–342, doi:10.1002/2016SW001371.

[23] Smith, G. P., & Thompson, A. D. (2009, June 9). *Space Policy Development Via Macro-Economic Analysis*. https://www.nasa.gov/pdf/368983main_Applying%20a%20Macro-Economic%20 Analysis%20to%20Space%20Policy%202009_06_09.pdf

[24] Wild, F. (2020, July 28). *What Is a Spacewalk?* https://www.nasa.gov/audience/forstudents/k-4/stories/nasa-knows/what-is-a-spacewalk-k4.html

Figure 1: The first ten simulated data from r program

	I	accumPV_Net	loss_1	loss_2	loss_3	i	NPV1	NPV2	NPV3
1	78.54813	2739.158	0.00000	0.00000	0.00000	0.03	2660.609	2660.609	2660.609
2	106.83924	2806.515	103.70813	103.33221	103.02664	0.03	2622.507	2622.787	2623.014
3	85.40720	2769.273	0.00000	0.00000	0.00000	0.03	2683.866	2683.866	2683.866
4	112.19048	2796.767	0.00000	0.00000	0.00000	0.03	2684.576	2684.576	2684.576
5	115.43640	2759.074	0.00000	0.00000	0.00000	0.03	2643.637	2643.637	2643.637
6	64.87394	2731.401	114.62578	117.42683	119.70362	0.03	2581.234	2579.150	2577.456
7	92.13796	2757.964	0.00000	0.00000	0.00000	0.03	2665.826	2665.826	2665.826
8	112.72168	2814.404	0.00000	0.00000	0.00000	0.03	2701.683	2701.683	2701.683
9	93.45608	2824.842	0.00000	0.00000	0.00000	0.03	2731.386	2731.386	2731.386
10	88.09873	2690.672	0.00000	0.00000	0.00000	0.03	2602.573	2602.573	2602.573

Figure 2: The histogram plot of first scenario's NPV (p=0.3)
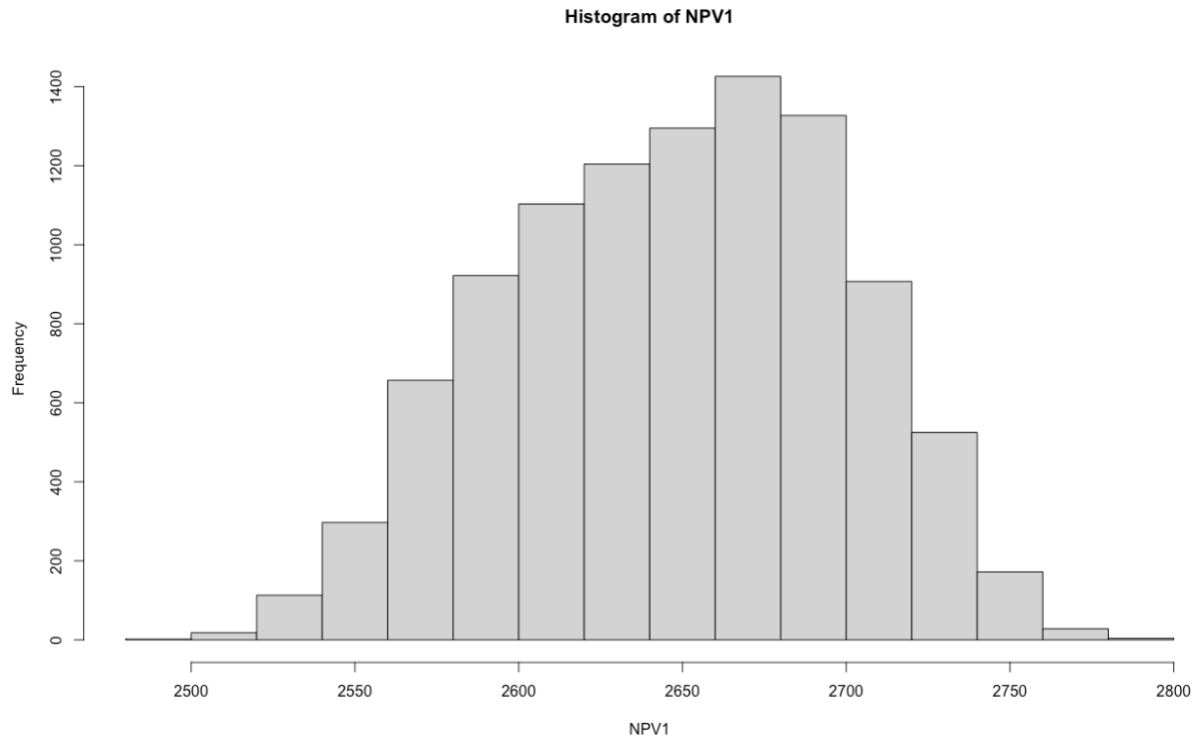
Histogram of NPV1

Figure 3: The histogram plot of second scenario's NPV (p=0.5)
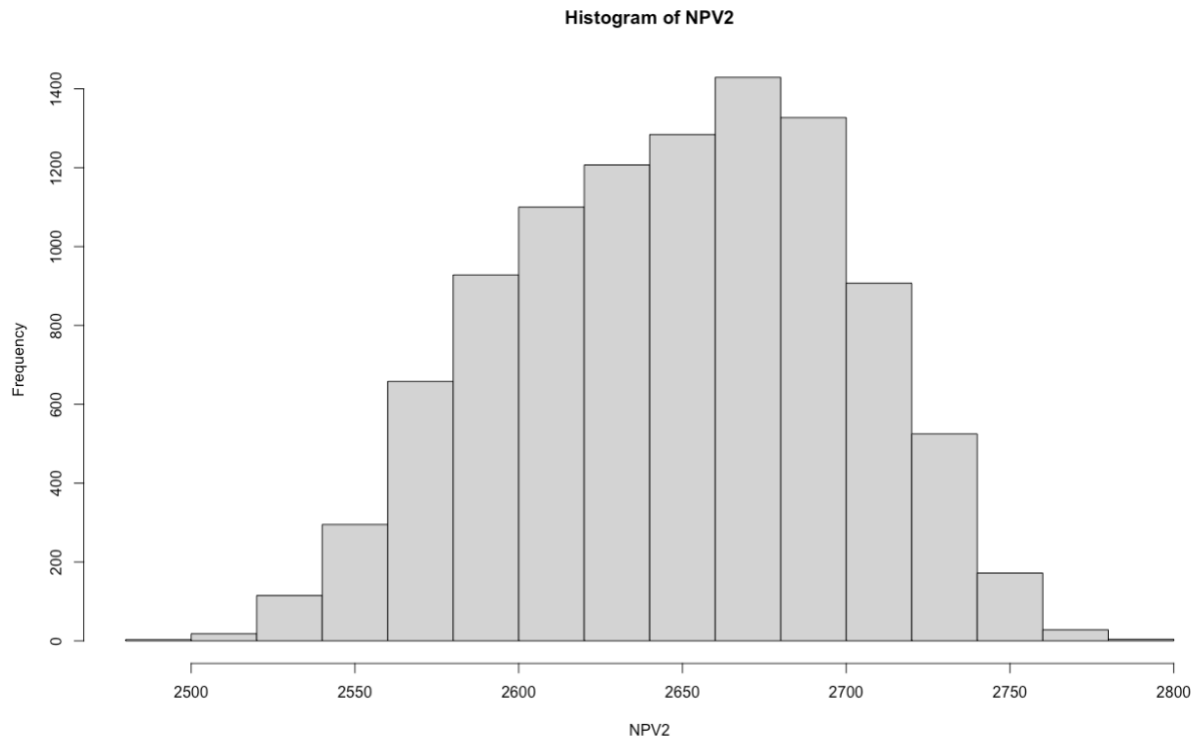

Histogram of NPV2

Figure 4: The histogram plot of third scenario's NPV (p=0.7)
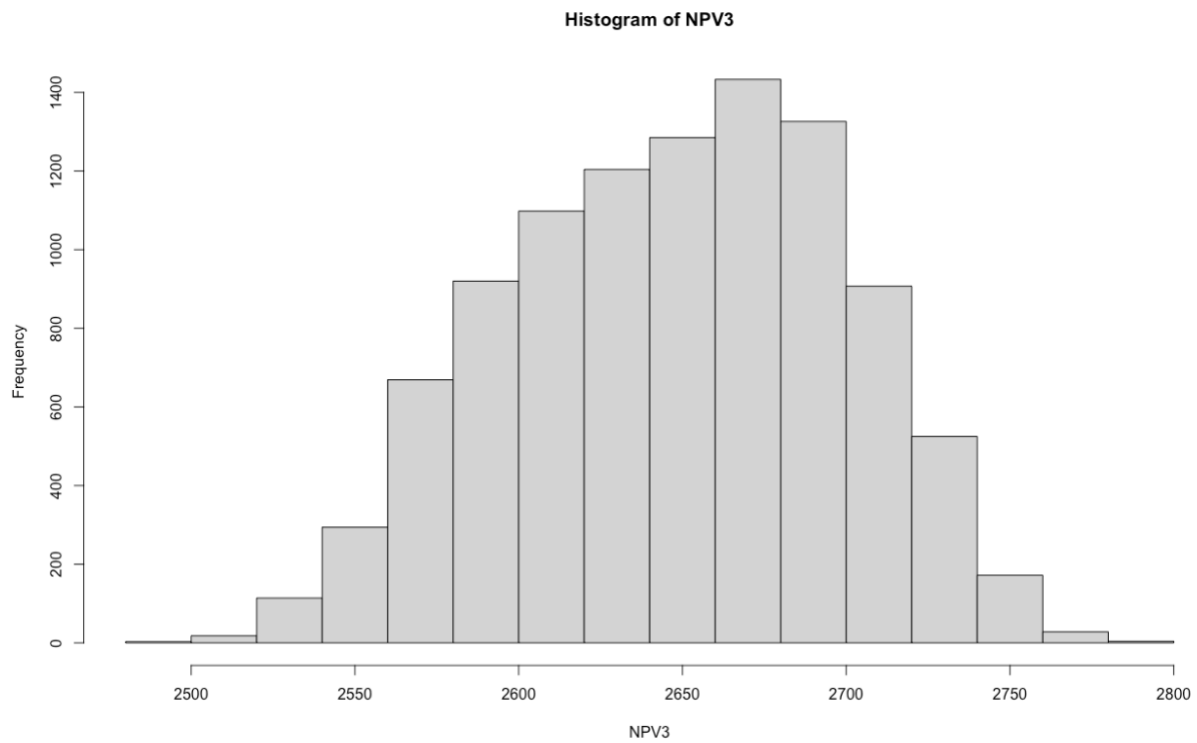

Histogram of NPV3

Table 4: The result of the Monte Carlo simulation – fixing the initial investment

Weight of the satellite (tons)	The average of NPV	The standard deviation of NPV
1	2,675.732	52.45278
2	2,664.432	52.45278
3	2,653.132	52.45278
4	2,641.832	52.45278
5	2,630.532	52.45278
6	2,619.232	52.45278

Figure 5: The plot of the number occurrence of the flares against time
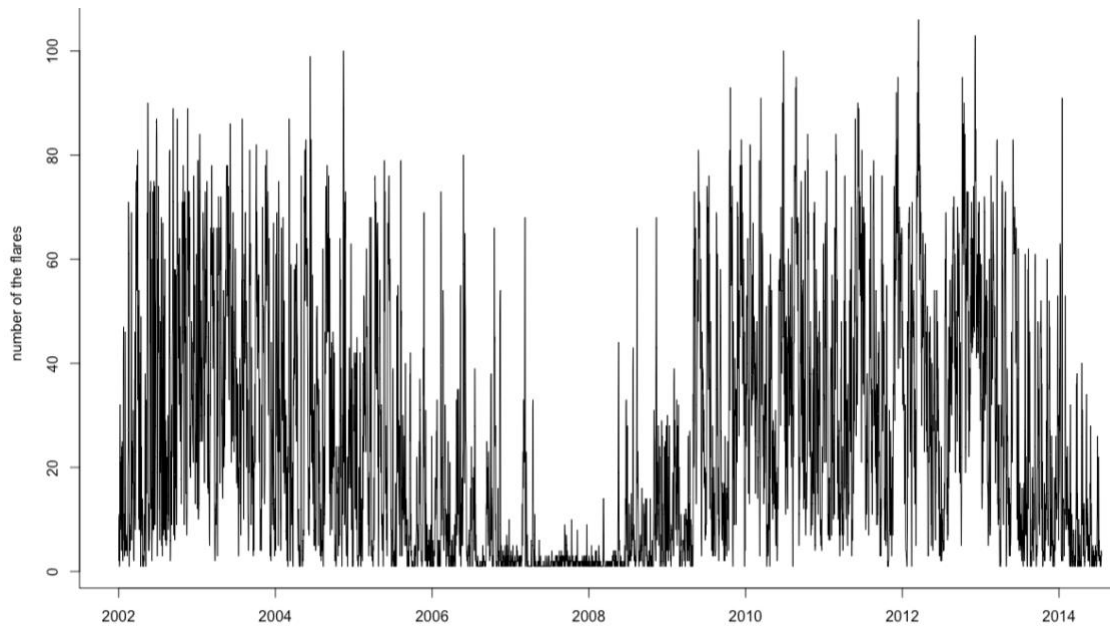
Figure 6: The Dickey-Fuller Test for minimum energy

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 1
  STATISTIC:
    Dickey-Fuller: -9.9639
  P VALUE:
    0.01
```

Figure 7: The Dickey-Fuller Test for maximum energy

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 1
  STATISTIC:
    Dickey-Fuller: -12.0147
  P VALUE:
    0.01
```

Figure 8: The acf plot of the minimum energy

Series t_min.training
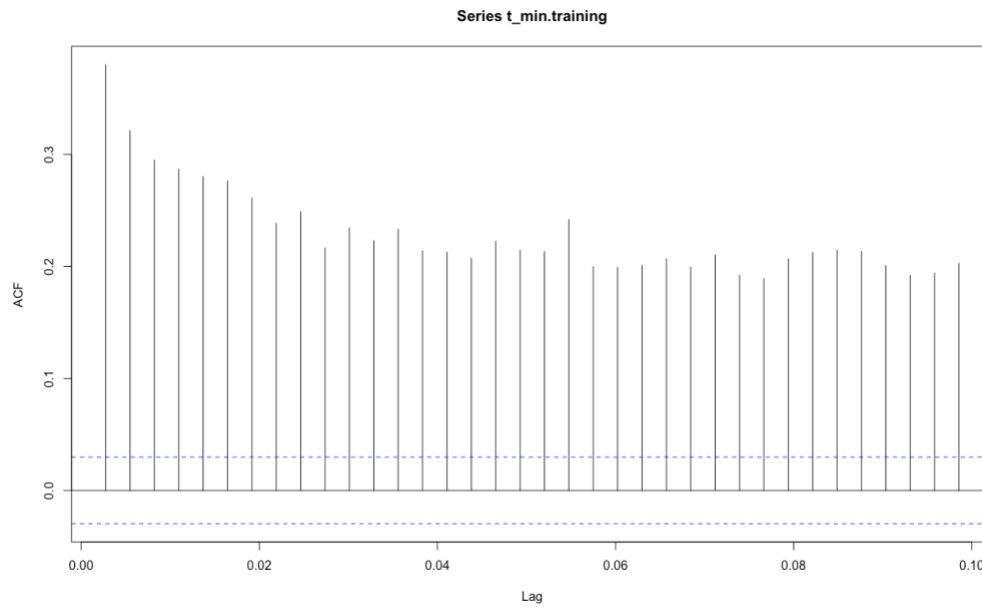
Figure 9: The pacf plot of the minimum energy

Series t_min.training
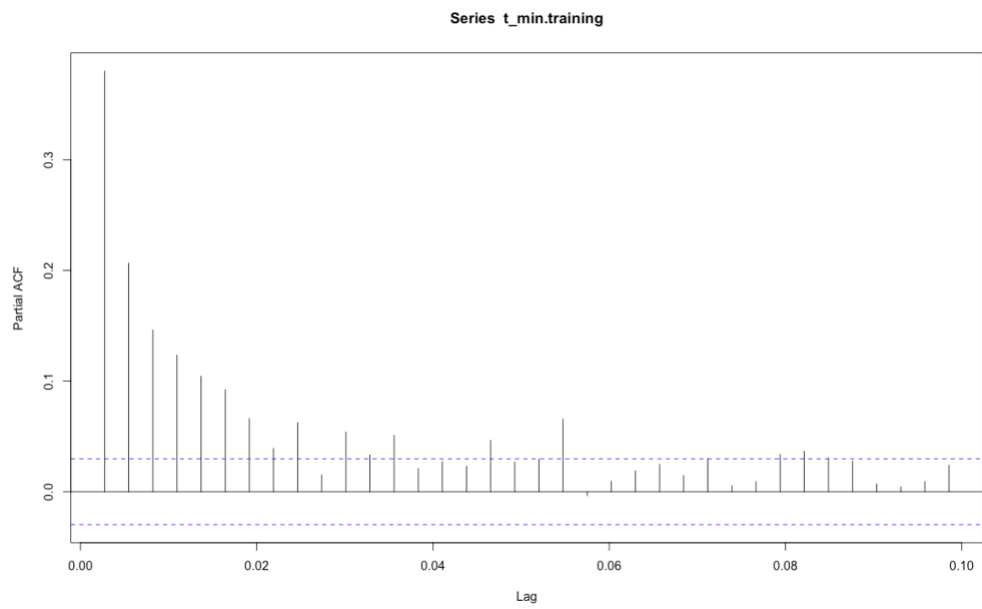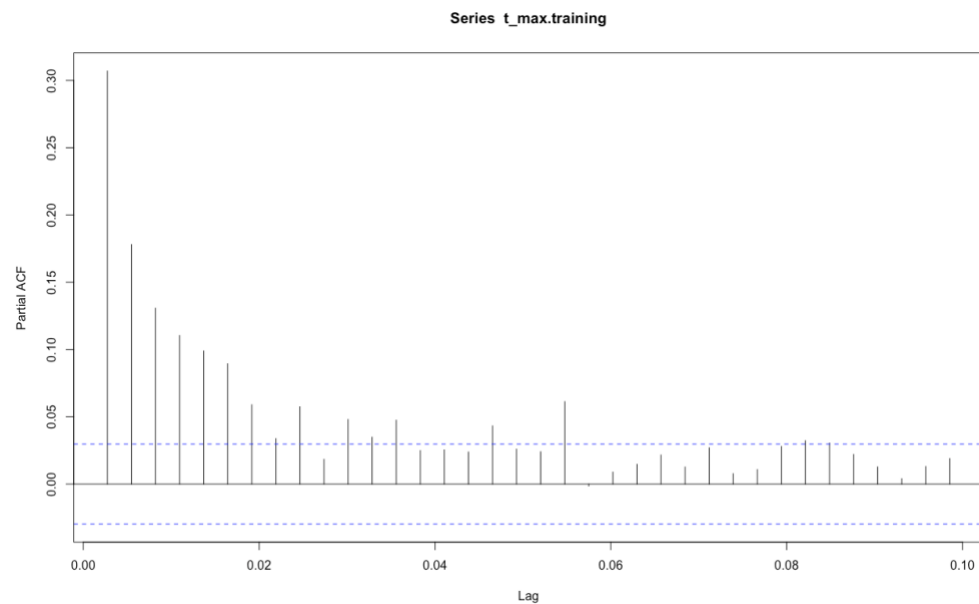
Figure 10: The eacf plot of the minimum energy

```
AR/MA
   0 1 2 3 4 5 6 7 8 9 10 11 12 13
0  x x x x x x x x x x x  x  x  x
1  x o o o o o o o x x o  o  o  o
2  x x o o o o o o o o o  o  o  o
3  x o x o o o o o o o o  o  o  o
4  x x o x o o o o o x o  o  o  o
5  x x x x x o o o o o o  o  o  o
6  x x x x x o o o o x o  o  o  o
7  x x x x x o x o o o o  o  o  o
```
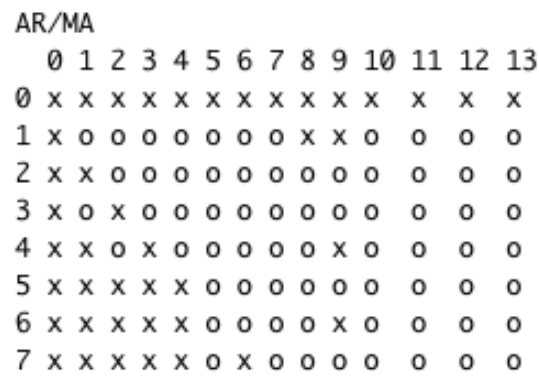
**Figure 11: The subset method of the minimum energy**
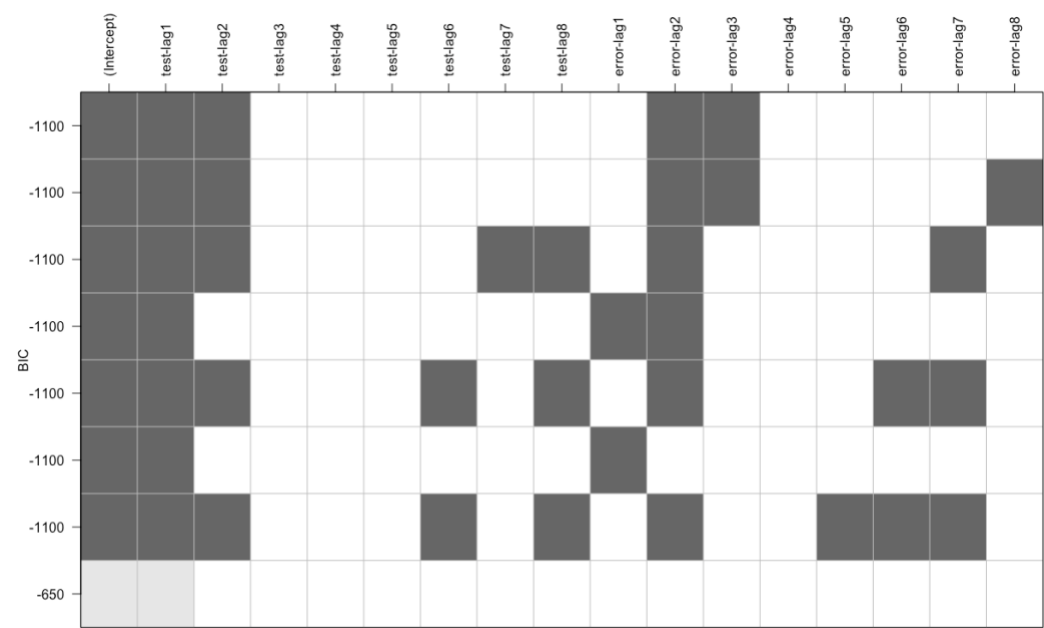
Figure 12: The acf plot of the maximum energy
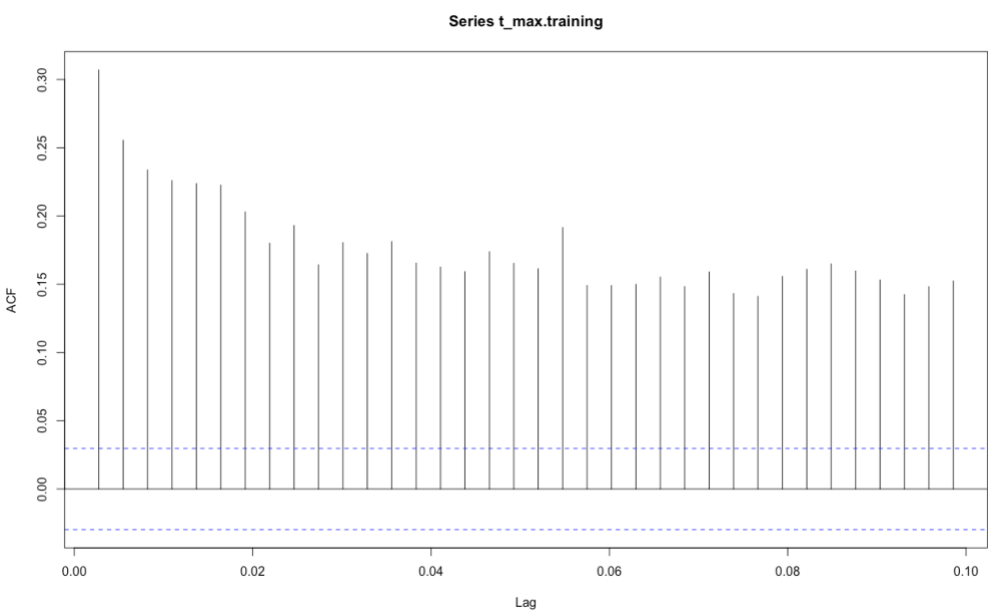
Figure 13: The pacf plot of the maximum energy

Series t_max.training

Figure 14: The eacf plot of the maximum energy

```
AR/MA
   0 1 2 3 4 5 6 7 8 9 10 11 12 13
0  x x x x x x x x x x x  x  x  x
1  x o o o o o o o x x o  o  o  o
2  x x o o o o o o o o o  o  o  o
3  x x x o o o o o o o o  o  o  o
4  x x o o o o o o o x o  o  o  o
5  x o x x x o o o o o o  o  o  o
6  x x x x x x o o o o o  o  o  o
7  x x x x x x x x o o o  o  o  o
```

Figure 15: The subset method of the maximum energy
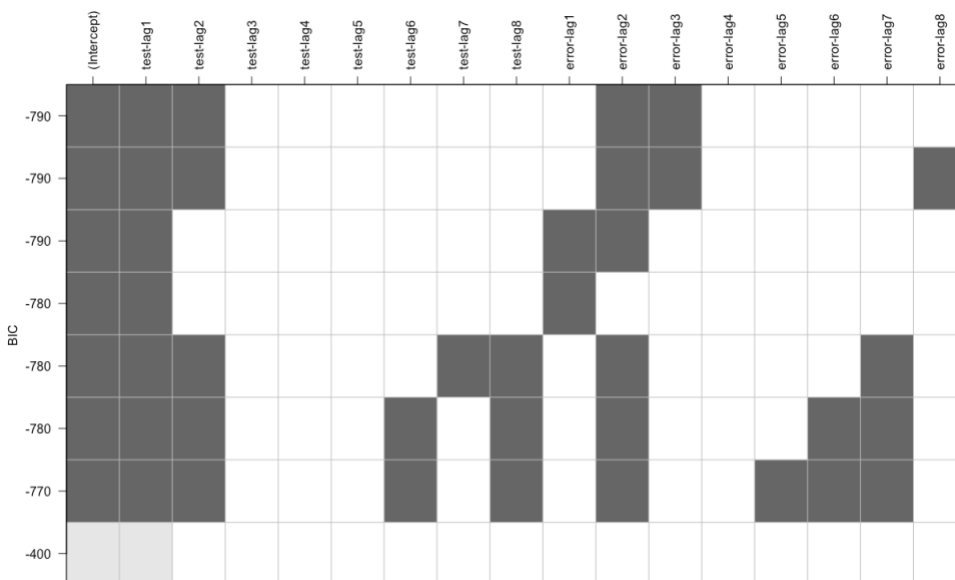
Figure 16: The coefficient of the minimum energy – ARMA (1,1)

```
Call:
arima(x = t_min.training, order = c(1, 0, 1), method = "ML")

Coefficients:
         ar1      ma1  intercept
      0.9691  -0.8317     6.7251
s.e.  0.0069   0.0184     0.2257

sigma^2 estimated as 7.529:  log likelihood = -10546.51,  aic = 21099.01
```

Figure 17: The coefficient of the minimum energy – ARMA (1,2)

```
Call:
arima(x = t_min.training, order = c(1, 0, 2), method = "ML")

Coefficients:
         ar1      ma1      ma2  intercept
      0.9927  -0.8006  -0.1250     6.7357
s.e.  0.0040   0.0157   0.0183     0.4104

sigma^2 estimated as 7.451:  log likelihood = -10524.07,  aic = 21056.14
```

31

Figure 18: The coefficient of the minimum energy – ARMA (2,1)

```
Call:
arima(x = t_min.training, order = c(2, 0, 1), method = "ML")

Coefficients:
         ar1      ar2      ma1   intercept
      1.1689  -0.1714  -0.9604     6.7416
s.e.  0.0188   0.0180   0.0100     0.6149

sigma^2 estimated as 7.427:  log likelihood = -10517.34,  aic = 21042.69
```

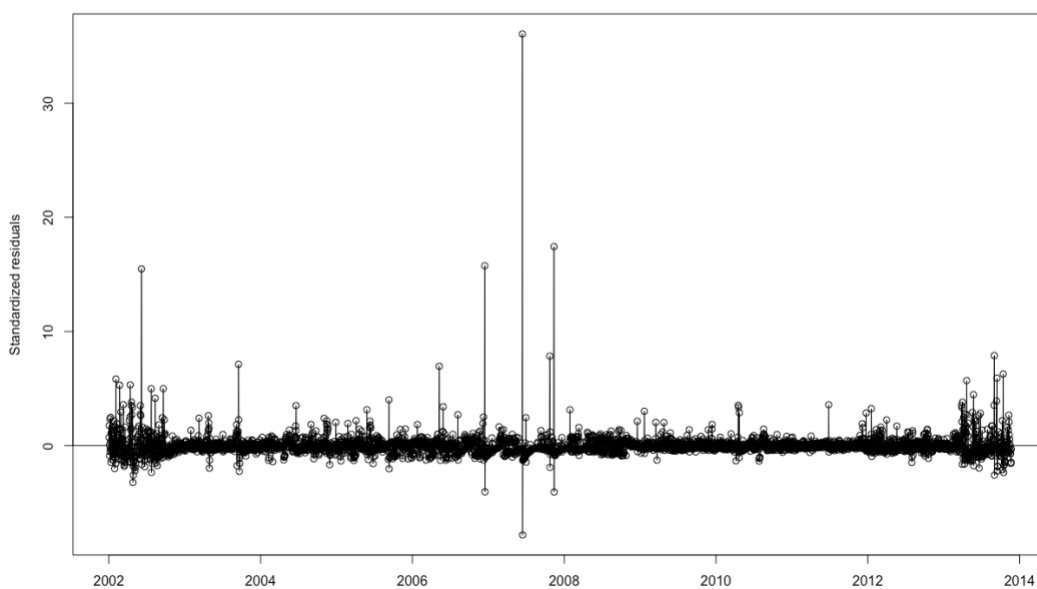**Figure 19: The residuals analysis of the minimum energy – ARMA (2,1)**

Figure 20: The normality test of the minimum energy – ARMA (2,1)
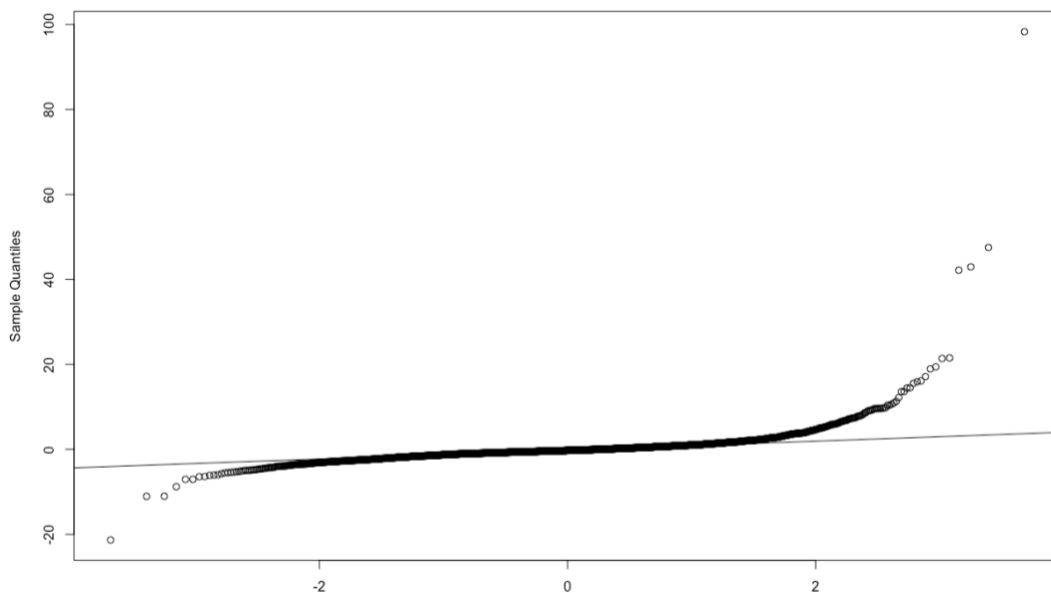
Figure 21: The Ljung-Box test of the minimum energy – ARMA (2,1)

```
          Box-Ljung test

data:  residuals(m1)
X-squared = 0.56425, df = 1, p-value = 0.4526
```

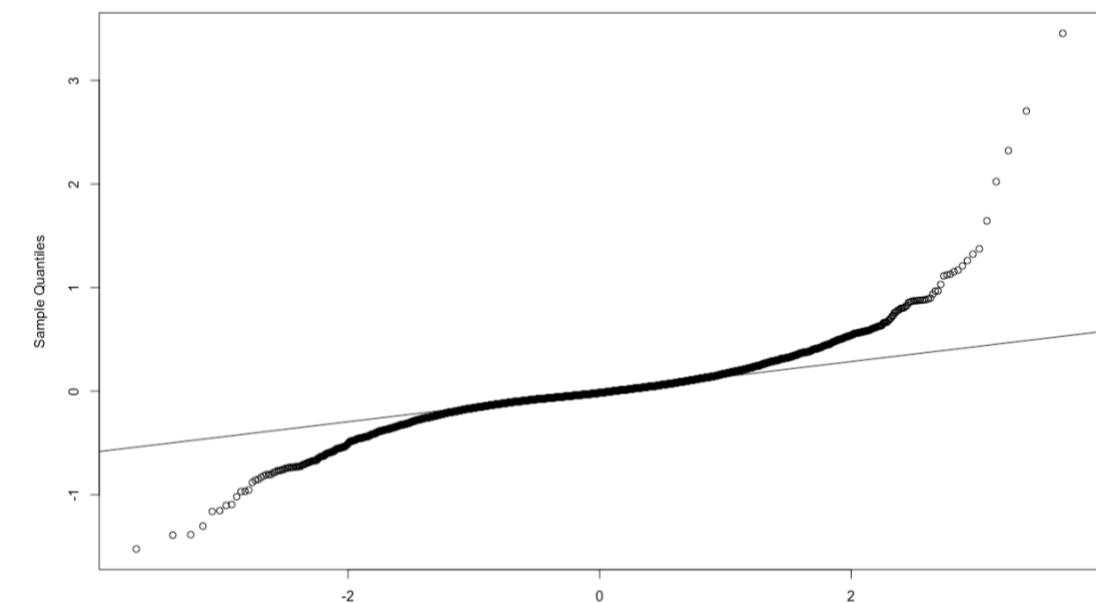**Figure 22: The normality test of the minimum energy with log-transformation – ARMA (2,1)**

Figure 23: The coefficient of the minimum energy – ARMA (2,3)

```
Call:
arima(x = t_min.training, order = c(2, 0, 3), method = "ML")

Coefficients:
         ar1     ar2     ma1      ma2      ma3  intercept
      0.0816  0.9043  0.1106  -0.8533  -0.1146     6.7288
s.e.     NaN     NaN     NaN      NaN   0.0176     0.4078

sigma^2 estimated as 7.451:  log likelihood = -10524.06,  aic = 21060.12
```

Figure 24: The coefficient of the minimum energy – ARMA (1,3)

```
Call:
arima(x = t_min.training, order = c(1, 0, 3), method = "ML")

Coefficients:
         ar1      ma1      ma2      ma3   intercept
      0.9978  -0.8055  -0.0752  -0.0773     6.7661
s.e.  0.0016   0.0153   0.0201   0.0152     0.7241

sigma^2 estimated as 7.409:  log likelihood = -10512.15,  aic = 21034.3
```

Figure 25: The coefficient of the minimum energy – ARMA (2,2)

```
Call:
arima(x = t_min.training, order = c(2, 0, 2), method = "ML")

Coefficients:
         ar1      ar2      ma1     ma2  intercept
      1.7495  -0.7495  -1.5753  0.5817     6.7743
s.e.  0.0410   0.0410   0.0493  0.0483        NaN

sigma^2 estimated as 7.348:  log likelihood = -10494.65,  aic = 20999.3
```

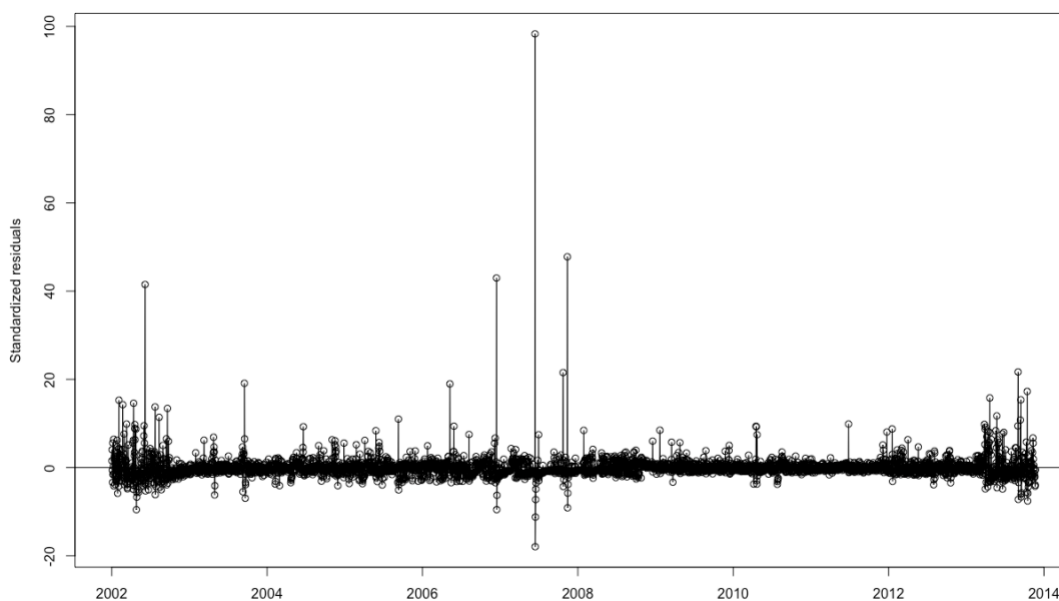**Figure 26: The residuals analysis of the minimum energy – ARMA (2,2)**

34

Figure 27: The normality test of the minimum energy – ARMA (2,2)
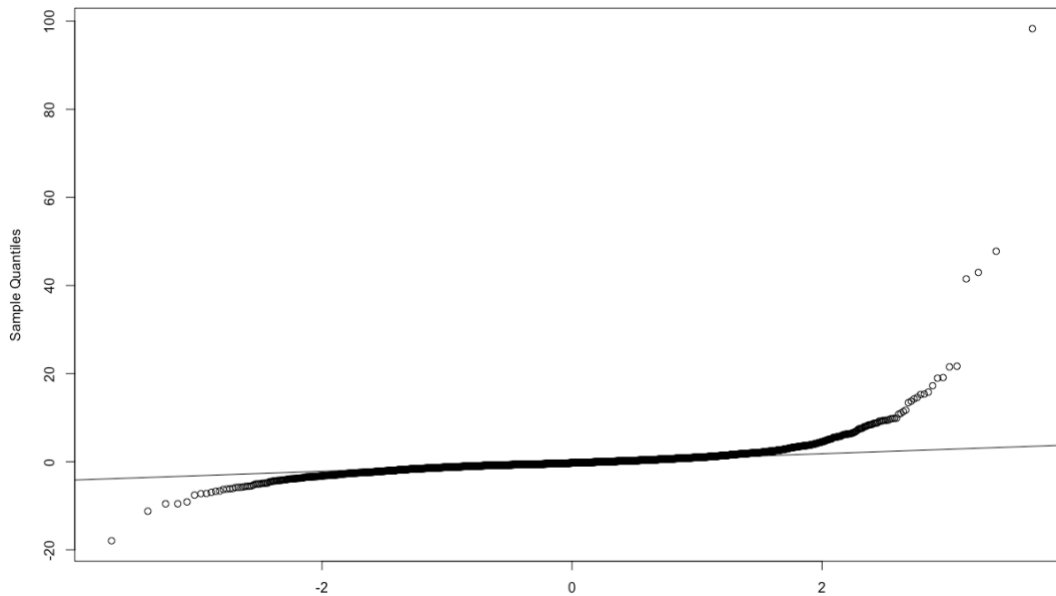

Figure 28: The Ljung-Box test of the minimum energy – ARMA (2,2)

```
        Box-Ljung test

data:   residuals(m2)
X-squared = 0.87017, df = 1, p-value = 0.3509
```

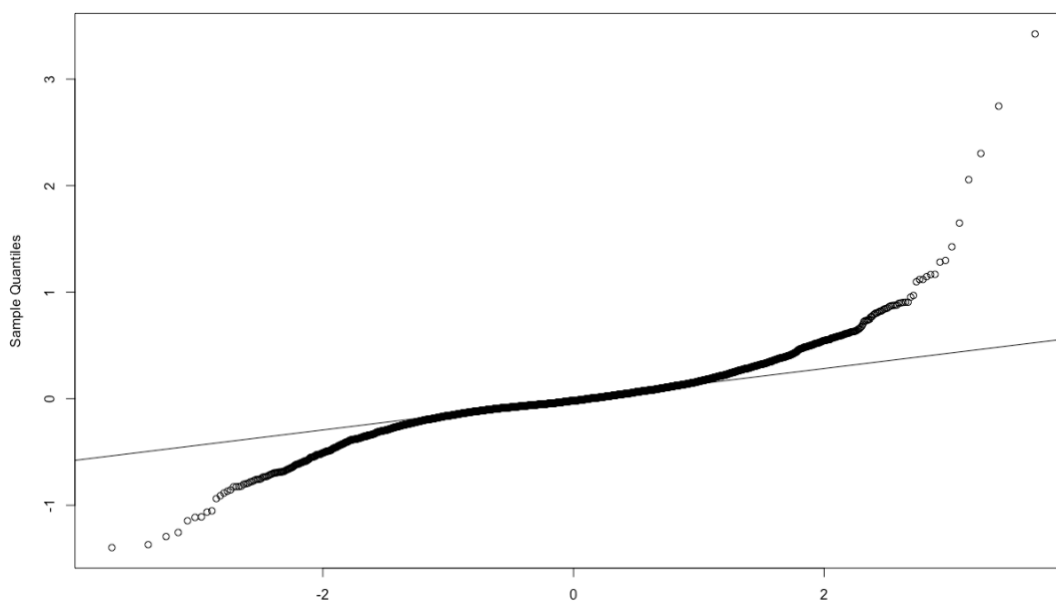**Figure 29: The normality test of the minimum energy with log-transformation – ARMA (2,2)**

Figure 30: The coefficient of the maximum energy – ARMA (1,1)

```
Call:
arima(x = t_max.training, order = c(1, 0, 1), method = "ML")

Coefficients:
         ar1      ma1  intercept
      0.9726  -0.8653    13.7211
s.e.  0.0074   0.0191     0.5066

sigma^2 estimated as 46.65:  log likelihood = -14506.94,  aic = 29019.88
```

Figure 31: The coefficient of the maximum energy – ARMA (1,2)

```
Call:
arima(x = t_max.training, order = c(1, 0, 2), method = "ML")

Coefficients:
         ar1      ma1      ma2  intercept
      0.9954  -0.8371  -0.1135    13.7667
s.e.  0.0028   0.0150   0.0161     1.0688

sigma^2 estimated as 46.22:  log likelihood = -14487.45,  aic = 28982.9
```

Figure 32: The coefficient of the maximum energy – ARMA (2,1)

```
Call:
arima(x = t_max.training, order = c(2, 0, 1), method = "ML")

Coefficients:
         ar1      ar2      ma1  intercept
      1.1378  -0.1403  -0.9660    13.7655
s.e.  0.0176   0.0170   0.0084     1.3156

sigma^2 estimated as 46.13:  log likelihood = -14483.03,  aic = 28974.06
```

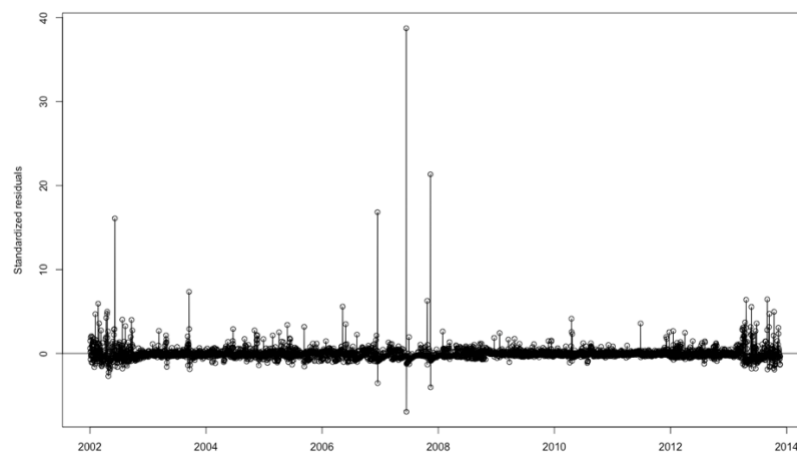**Figure 33: The residuals analysis of the maximum energy – ARMA (2,1)**

Figure 34: The normality test of the maximum energy – ARMA (2,1)

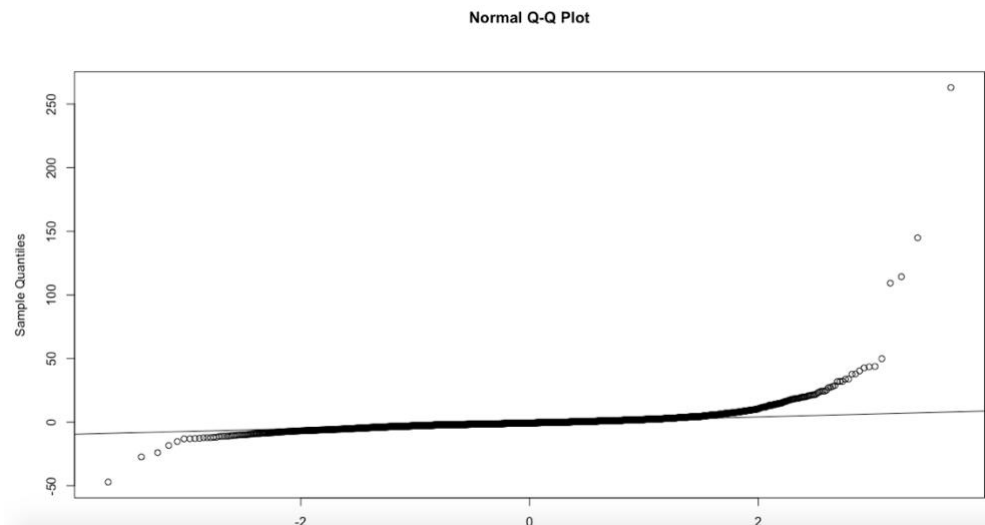
Normal Q-Q Plot

Figure 35: The Ljung-Box test of the maximum energy – ARMA (2,1)

```
       Box-Ljung test

data:   residuals(m1)
X-squared = 0.2977, df = 1, p-value = 0.5853
```

Figure 36: The normality test of the maximum energy with log-transformation – ARMA (2,1)
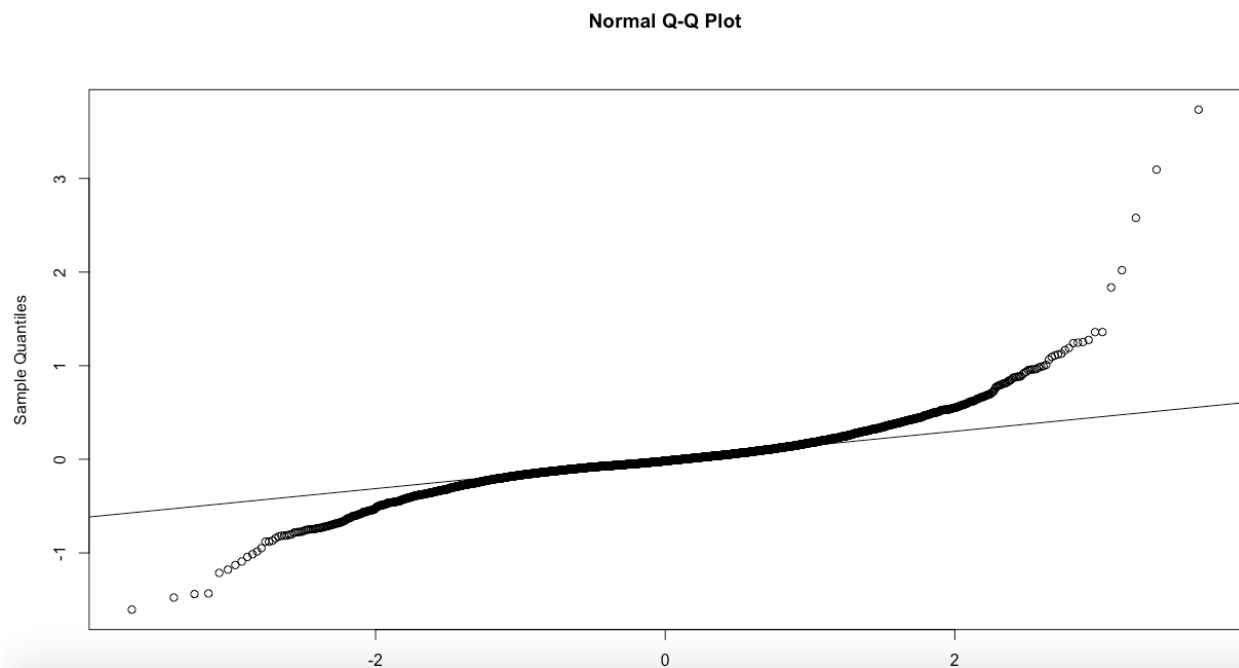

Normal Q-Q Plot

Figure 37: The coefficient of the maximum energy – ARMA (2,3)

```
Call:
arima(x = t_max.training, order = c(2, 0, 3), method = "ML")

Coefficients:
         ar1      ar2      ma1      ma2      ma3   intercept
      0.0774   0.9137   0.0808  -0.8803  -0.1058     13.7306
s.e.     NaN      NaN      NaN      NaN   0.0150      1.0658

sigma^2 estimated as 46.22:  log likelihood = -14487.42,  aic = 28986.84
```

Figure 38: The coefficient of the maximum energy – ARMA (1,3)

```
Call:
arima(x = t_max.training, order = c(1, 0, 3), method = "ML")

Coefficients:
         ar1      ma1      ma2      ma3   intercept
      0.9978  -0.8386  -0.0606  -0.0657     13.7780
s.e.  0.0016   0.0153   0.0203   0.0151      1.5275

sigma^2 estimated as 46.03:  log likelihood = -14478.27,  aic = 28966.55
```

Figure 39: The coefficient of the maximum energy – ARMA (2,2)

```
Call:
arima(x = t_max.training, order = c(2, 0, 2), method = "ML")

Coefficients:
         ar1      ar2      ma1      ma2   intercept
      1.4807  -0.4855  -1.3101   0.3392     13.7980
s.e.  0.2192   0.2157   0.2279   0.2090      0.6248

sigma^2 estimated as 46.02:  log likelihood = -14477.39,  aic = 28964.77
```

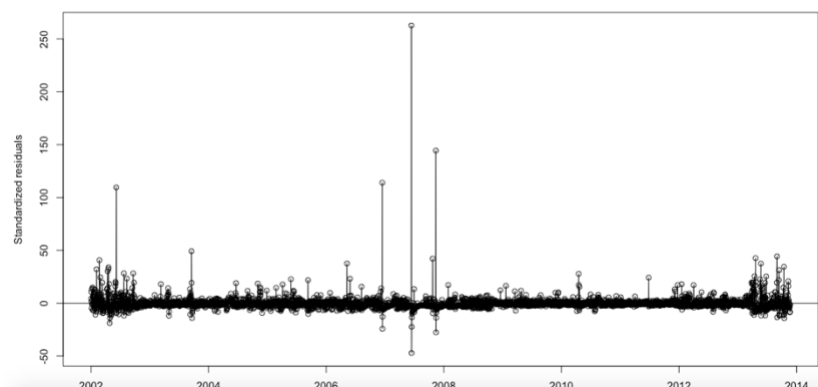**Figure 40: The residuals analysis of the maximum energy – ARMA (2,2)**

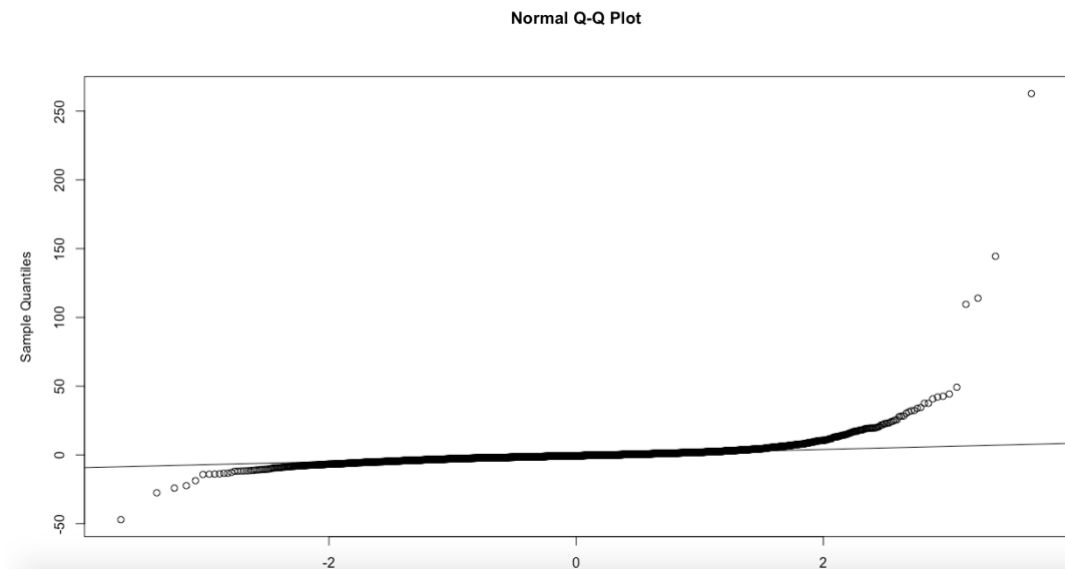**Figure 41: The normality test of the maximum energy – ARMA (2,2)**

Normal Q-Q Plot

Figure 42: The Ljung-Box test of the maximum energy – ARMA (2,2)

```
        Box-Ljung test

data:   residuals(m2)
X-squared = 0.24908, df = 1, p-value = 0.6177
```

Figure 43: The normality test of the maximum energy with log-transformation – ARMA (2,2)

Normal Q-Q Plot