MODELLING OF EXTREME TEMPERATURE AND ITS CORRELATION WITH ENSO: A CASE STUDY IN PERLIS

by

MOHD SHAZLI BIN ISMAIL
(1632121989)

A dissertation submitted in partial fulfillment of the requirements for the degree of
Master of Science (Engineering Mathematics)

Institute of Engineering Mathematics
UNIVERSITI MALAYSIA PERLIS

2017

ACKNOWLEDGMENT

First, I would like to thank Allah the almighty whom give me strength and wisdom to endure this journey and completed this dissertation. To my lovely wife and kids whom fully supported me without any hesitation and continuous love and cared for me so much all the way. Not to forget, my parents who always wanted me to further my studies, this is a small gift for them.

I would first like to thank my dissertation advisor Dr. Norazrita Mohd Amin of the Institute of Engineering Mathematics at Universiti Malaysia Perlis. The door to Dr. Norazrita office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this paper to be my own work, but steered me in the right the direction whenever she thought I needed it.

Finally, I must express my very profound gratitude to everybody who involves directly or indirectly for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

COADS	Comprehensive
ENSO	El Nino southern oscillation
ESRL	Earth system research laboratory
EV	Extreme value
EVT	Extreme value theory
GEV	Generalized extreme value
GPD	Generalized pareto distribution
IID	Independent identically distributed
KS	Kolmogorov Smirnov
LM	L moments
MEI	Multivariate ENSO index
MLE	Maximum likelihood estimator
MMD	Malaysian Meteorological Department
NOAA	National Oceanic and Atmospheric Administration
ONI	Oceanic nino index
PC	Principal component
PCA	Principal component analysis
POT	Peak over threshold
QQ	Quantile-quantile
RMSE	Root mean squared error
SOI	Southern oscillation index
SST	Sea surface temperature

LIST OF SYMBOLS

C	Celsius
μ	Location parameter
σ	Scale parameter
ξ	Shape parameter

PEMODELAN SUHU MELAMPAU DAN KORELASINYA DENGAN ENSO. SATU KAJIAN KES DI PERLIS

ABSTRAK

Teori nilai ekstrim (EV) telah meningkatkan perhatian penyelidik untuk memodelkan dan meramalkan peristiwa bencana atau risiko tinggi. Memandangkan peristiwa alam sekitar yang melampau boleh menyebabkan kehilangan harta benda yang besar dan menjejaskan kehidupan manusia, oleh sebab itu adalah penting untuk memahami tingkah laku kejadian yang tidak biasa ini dan meramalkan kejadian yang akan datang. Teori EV memberikan beberapa pemahaman kepada taburan ekor di mana model standard telah terbukti tidak boleh dipercayai. Pengagihan nilai ekstrim yang umum (GEV) digunakan untuk memodelkan ekstrim berdasarkan kaedah maxima blok. Kesimpulan peristiwa ekstrem alam sekitar adalah penting sebagai garis panduan dalam merancang struktur untuk bertahan di bawah keadaan yang sangat melampau. Suhu melampau menyebabkan pelbagai kesan dan boleh dikaitkan dengan kesihatan manusia dan kerosakan bahan. Sementara itu, El Nino Southern Oscillation (ENSO), adalah keadaan iklim yang berlaku sepanjang tahun di seluruh dunia. Teori EV diterapkan melalui GEV dengan dua pendekatan, Penganggar kebolehjadian maksimum (MLE) dan L-Moments (LM) untuk memodelkan suhu yang melampau di Chuping, Perlis dan dengan itu menganggarkan tahap ekstrim masa depan. Masalah saiz blok dibincangkan. Ujian kebagusan suaian (GOF) dilakukan dengan menggunakan punca ralat kuasa dua min (RMSE) dan ujian Kolmogorov Smirnov (KS) untuk menentukan pendekatan terbaik. Dapatan keputusan mendapati MLE lebih baik daripada LM untuk blok tahunan tetapi LM menyamakan kedudukan untuk blok bulanan. Tahap kembali dijangka melebihi nilai maksimum $40.1^{o}C$ pada tempoh pulangan tertentu. Koefisien korelasi sederhana, r diperolehi dalam mencari hubungan antara suhu melampau dan ENSO. Oleh itu, dibuktikan bahawa peristiwa ENSO mempengaruhi suhu melampau di Chuping, Perlis.

MODELLING OF EXTREME TEMPERATURE AND ITS CORRELATION WITH ENSO. A CASE STUDY IN PERLIS

ABSTRACT

Extreme value (EV) theory has raised researcher's attention for modelling and forecasting the catastrophic or high-risk events. Since extreme environmental events may cause huge loss of properties and affect human life, therefore it is significant to understand the behavior of such uncommon events and predict the upcoming. EV theory affords some understanding to the tail of a distribution where standard models have proved unreliable. The generalized extreme value (GEV) distribution is used to model the extreme based on block maxima method. Inference of the extremes of environmental events is essential as a guideline in designing structures to survive under the utmost extreme condition. Extreme temperature caused various effect and can be associated with human health and material damage. Meanwhile, El Nino Southern Oscillation (ENSO), is a climate condition that happened all year round globally. The EV theory is applied through GEV with two approaches, Maximum Likelihood Estimate (MLE) and L-Moments (LM) to model the extreme temperature in Chuping, Perlis and hence estimates the future extreme levels. The issue of block size is discussed. The goodness of fit (GOF) test is done using root mean squared error (RMSE) and Kolmogorov Smirnov (KS) test to determine the best approach. From the results, it is found that MLE is better than LM for the yearly block but LM squared things up for the monthly block. The return level is expected to exceed the maximum value $40.1^{o}C$ at certain return periods. The moderate correlation coefficient, r is obtained in finding the relation between extreme temperature and ENSO. So, it is proofed that ENSO events influences extreme temperature in Chuping, Perlis.

CHAPTER 1 : INTRODUCTION

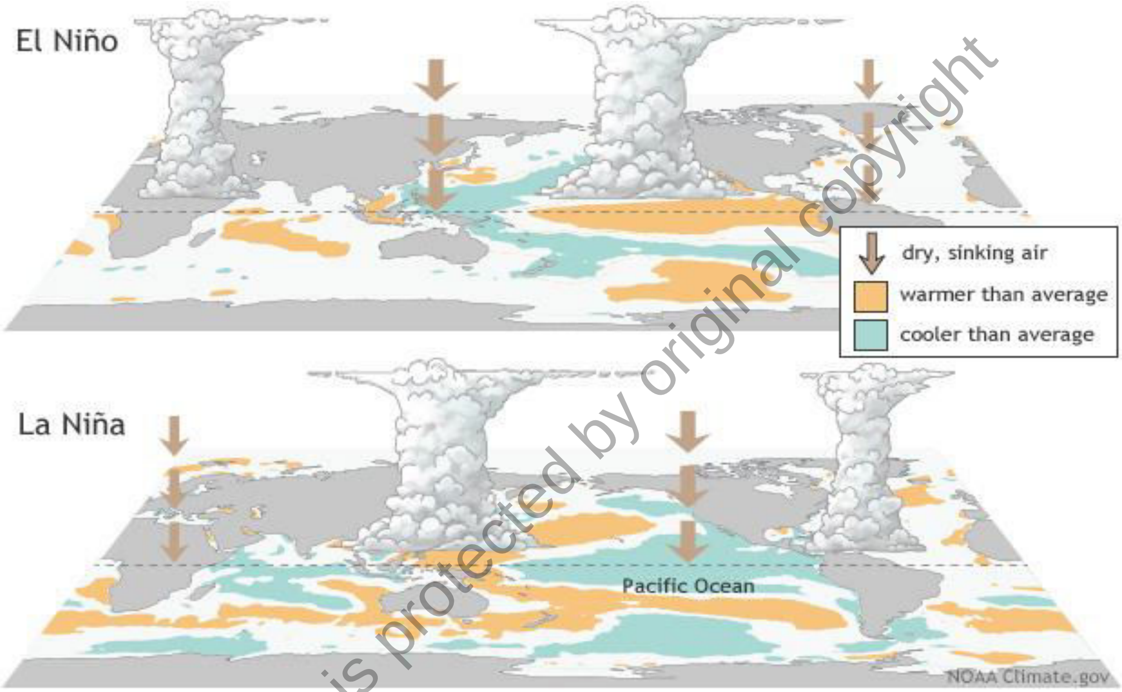
1.1 Background Study

El- Niño is a climate phenomenon which happens once every two to seven years. This phenomenon occurred when the hot sea current was replaced with the cold one outside of west coast Peru, South America. The El- Niño name itself was said to be given by Peruvian fisherman when he noticed the change of sea water temperature while fishing in early 20^{th} century.

La Niña on the other hand, is the opposite characteristic of El- Nino. El- Nino represent the warm phase while La Niña signify cold phase. In the context of interannually variability, El Niño and La Niña represent, respectively, warm and cold anomalies of sea surface temperature (Bobrowsky, 2013).

Southern Oscillation on the other hand is the term for atmospheric pressure changes between the east and west tropical Pacific that accompany both El- Niño and La Niña episodes in the ocean. El- Niño, La Niña and the Southern Oscillation together comprise a complex system of climate fluctuations termed the El- Niño Southern Oscillation (ENSO) phenomenon (Daud, Zakaria, Sahat, Moin, & Ismail, 2013). This climate condition lasts between 6-18 months during its occurrence. And given the fact that 70 percent of the world is covered by water (sea), that is why this phenomena affects the whole world.

Figure 1.1 shows El- Niño and La Niña phases which is a cyclic climate trend across tropical pacific called ENSO. This kind of condition change back and forth unevenly every two to seven years, and each phase triggers predictable interruptions of temperature, precipitation, and winds. These changes disturb the large-scale air movements in the tropical region, thus triggering a cascade of global side effects



(NOAA Climate.gov, 2016)

Figure 1.1: Example of El-Nino and La-Nina graphical description. (NOAA Climate.gov, 2016).

1.2 Problem Statement

In South East Asia or Malaysia specifically, precipitation anomalies vary spatially and seasonally with ENSO evolution (Juneng & Tangang, 2005; Tangang, Juneng, & Ahmad, 2007). A lot of ENSO effect is done on precipitation. Winter, Tawn, & Brown (2016) claimed that the effect of ENSO on mean global temperatures has been well studied but the impact on extreme temperature is less stated. To date, only a small number of research has been done on ENSO effect and temperature (Winter et al., 2016).

In Malaysia, ENSO effect on climate has been studied only on rainfall (Daud et al., 2013; Juneng & Tangang, 2005). Whereas globally, other than temperature, the ENSO effect is studied on precipitation (Davey et.al, 2014), westerly and easterly wind burst (Hu & Fedorov, 2017). From this, we knew that there is still no research being done on correlation of ENSO on maximum temperature in Malaysia.

Chuping, Perlis is considered as one of the hottest place in Malaysia. It rose to fame when it recorded the highest temperature in Malaysia at 40.1°C, on April 9th, 1998 (General Climate Information, 2011). This record is still unbeatable till now in Malaysia. So, did ENSO caused the extreme temperature in Chuping? Is there a correlation between ENSO and such extreme temperature?

There are a lot of negative impact of high temperature, in this case, the extreme temperature to environment and people. The dry spell, forest fires, human health are some main causes. The effect of extreme temperature on the environment can cause the chain reaction. For example, high or extreme temperature may lead to the dry spell. This may lead to less rain. Then it may cause forest fire started and it may lead to such catastrophic impact, loss of habitat of flora and fauna, food supply and water supply to human and animals. This kind of effect also being stated by Hasan, Salam, & Kassim, (2013)

This hazardous event is rarely to occur but it might happen rather sooner or later. The effect of high or extreme temperature on human also being considered. As stated before, human being also being affected by forest fires (e.g. loss of food supply and water supply). But the most affected one is human health. Heat stroke and skin cancer may endanger human's life. In Malaysia heat stroke cases have come to headline because it leads to death. In Europe, heatwaves event in 2003 causes 40000 casualties of heat-related (Winter & Tawn, 2016). Since the negative effect of extreme temperature seems worst, it is important to prevent and overcome the causes or taking cautionary steps in facing this situation.

1.3 Objectives

The objectives of this research are:

i. To model the extreme temperature using block maxima method

ii. To predict the return level of extreme temperature in certain period

iii. To examine whether ENSO correlate with extreme temperature in Chuping

1.4 Research Scope

The scope of this study is at the maximum temperature in Chuping, Perlis. The period of maximum temperature that has been considered for duration between January 1997 until December 2016. This period was chosen because two mass ENSO events are believed to happen in 1998 and 2016. The Multivariate ENSO Index (MEI) average bimonthly mean index is also being selected between those periods.

1.5 Expected Outcome

By using Generalized Extreme Value (GEV) distribution, extreme temperature in Chuping will be model and the return level that is expected to exceed a certain level after several return periods is computed. It is also expected that there is a correlation between ENSO and extreme temperature in Chuping.

From the outcome of the research, it is hope that local authorities such as town council and health department can act wisely during ENSO period by taking extra cautionary steps in spreading information to public on the severity of health if affected by extreme temperature condition.

CHAPTER 2 : LITERATURE REVIEW

2.1 Application of Extreme Value Theory

Extreme Value (EV) hypothesis was first created by Fisher and Tippett and formalized by Gnedenko. It demonstrates the appropriation of the block maxima of an example of independent identically distributed (iid) factors joins to an individual from the purported EV distribution (Faranda, Lucarini, Turchetti, & Vaienti, 2011). It all starts with the study of stochastically series. Then it becomes interesting in many areas. It has been practiced in many disciplines such as extreme precipitation and climate events. Nathan et al., (2016) did the research in extreme precipitation. While Abaurrea, Asin, & Cebrian (2015); García-Cueto, Cavazos, de Grau, & Santillán-Soto (2014); Tangang et al., (2012) used extreme value theory (EVT) in various climate events. EVT models the behaviour of extreme observations, either maxima or minima (García-Cueto et al. 2014). Studies related to the analysis and modelling of extreme climate events use general circulation models and the EVT as essential tools (García-Cueto, et.al, 2013).

In a refinement to the central limit theorem, where every one of the qualities in a period arrangement is viewed as, EVT examines the union of the qualities in the tail of a dispersion towards Fréchet, Gumbel or Weibull probability distributions (Ghil et al., 2011). The distribution also known as type 1, type 2 and type 3 which referred to Fréchet, Gumbel and Weibull distribution respectively.

A lot of statistical methods did not pay too much attention to the high or low end of the tail distribution. It utmost concerned is on the center of the distribution (Amin, Adam, & Aris, 2015).

Generally, there are two approaches stated by Fanone, Gamba, & Prokopczuk (2013) to recognize extremes in real data. The first one is considering the maximum data in periods, for example, months or years. It is also called block maxima. It corresponds to Generalized Extreme Value (GEV) distribution. The second approach is Generalized Pareto Distribution (GPD) which focuses on the exceedances of a certain high threshold.

2.1.1 Application of Block Maxima Method

The block maxima method is inspired by the thesis of Gumbel which usually applied in environmental sciences, especially to analyze extremes of a given series of observations. It forms a sample of block maxima which is constructed by dividing the early high frequency of observation into blocks of equal size (Kojadinovic & Naveau, 2016). The block maxima approach in EV theory consists of dividing the observation period into an overlapping periods of equal size and restricts attention to the maximum observation in each period (Ferreira & De Haan, 2015).

The block maxima method for EV analysis basically consists of the following. Firstly, long series of data is being partitioned into blocks where for each block, the maximum is computed. Then EV is fitted to the sample of block maxima. Typically, the blocks represent the months and years of data. So, the name monthly and annual maxima series are used and the tail quantiles are figured by using the fitted distribution (Bücher & Segers, 2014). Hasan et al., (2013) formed series of block maxima $P_{q,1}, P_{q,2}, ..., P_{q,p}$ from a group of extreme values of equal block length *n*. It was then fitted in GEV. They also added that an analysis of annual maximum data is likely to be more robust compared to shorter blocks. Block Maxima seems as an essential method in implementing EV theory.

2.1.2 Application of Extreme Value Distribution

GEV distribution is a broadly utilized model for extraordinary occasions tails one of three sorts of distribution, including Gumbel, Fréchet, and Weibull (Jahanbaksh Asl, Khorshiddoust, Dinpashoh, & Sarafrouzeh, 2013). This distribution is commonly used one compared to other outrageous EV family. A solitary GEV dissemination is because of a blend of a group of consistent likelihood conveyance which was proposed for factual strength (Ahmat, Yahaya, & Ramli, 2015). There are 3 types of parameters used in this distribution. Location parameter, μ, to determine the shifting of a distribution in specified direction on the horizontal axis. Next is scale parameter, σ, to measure the dispersion of a distribution also its concentration. Lastly the shape parameter, λ or ξ, mention the shape of a distribution and its tail (Ahmat et al., 2015; Hasan, Radi, & Kassim, 2012; Hasan et al., 2013; Amin et al., 2015).

García-Cueto et al., (2013) used GEV for modelling annual block maxima in Mexicali, Mexico. GEV also is commonly used model for extreme events where the issue of non-stationarity of the extreme sequences is handled by regression modelling of the GEV parameters (Bocci, Caporali, & Petrucci, 2013). Meanwhile, Tye & Cooley (2015) used

GEV in their studies to determine how large rainfall that lead to 2013 Colorado flood affect the GEV parameter estimates applied by designers and planners.

On the other hand, Huang, Xia, Guo, & Yang, (2013) fitted the GEV distribution to the yearly outrageous temperature and the parameters were assessed by the ML technique. The arrival levels for 10,50 and 100 years return periods assessed by the profile likelihood strategy were gotten for the yearly extraordinary temperature. Fuentes, Henry, & Reich (2013) used a spatial structure for GEV parameters in their application of extreme temperature data.

In Malaysia specifically, Zin & Jemain (2010) applied GEV in statistical distributions of yearly intense sequence and part moisture less series of drought occasion between 1975-2004. Hasan et al., (2012) used GEV in modelling extreme temperature in Bayan Lepas, Penang between year 2000-2009, and modelling the distribution of extreme share return. While Hasan et al., (2013) used GEV in modelling annual extreme temperature for several places across Malaysia. This distribution also had been used by Ahmat et al., (2015) in predicting PM_{10} extreme concentrations in urban monitoring stations in Selangor and Mohd Amin et al., (2015) used GEV in EV analysis for modeling high PM_{10} level in the southern state of Peninsular Malaysia, Johor Bahru.

Sun et al., (2017) used nonstationary GEV distributions to model extreme precipitation over global land for the period 1979–2015.Wehner (2010) applied GEV in investigating three sources of uncertainty in the calculation of extreme value statistics for observed and modeled climate data which is inter-model differences in formulation, unforced internal variability, and choice of statistical model. In Europe, Siliverstovs et al., (2010) practiced GEV in modeling maximum temperatures in Switzerland monitored in twelve locations. Bayesian framework is used in determining the GEV parameters.

2.2 Approach for Parameter Estimate

There are some methods of parameter estimate such as least-squares estimation, maximum likelihood estimation (MLE), L moments and probability weighted moment (PWM) (Myung, 2003). Benstock & Cegla (2017) used MLE as a method of estimating the parameters of a distribution from a given set of data. Mazas, Garat, & Hamm (2014) in their paper used MLE with 2 parameters by Peak Over Threshold (POT) method in determining extreme environmental variables which eventually they found that L-moments with 3 parameters is better alternative than MLE. In climate extreme events analysis, Jahanbaksh Asl et al.(2013) suggested in using MLE method for estimating parameters of maximum and minimum temperature, maximum wind speed, and maximum precipitation series of Zanjan, Iran.

García-Cueto et al. (2013) stated three reasons on why he applied MLE to estimate parameters. The first one is the data sample is large or bigger than fifty (> 50), so it can be compared in performance with other methods. Secondly, covariate information is easily put into parameter estimations. Finally, error limits for the parameter estimator is easily obtained compared with other alternative methods. MLE also being used to estimate parameters for non-homogenous common Poisson shock process in determining extreme heat events in Barcelona, Spain (Abaurrea, Asin, & Cebrian, 2015).

Hosking introduced the LM method in 1990. Since, it is used in estimating parameters of a distribution, regularly in the field of connected research, for example, meteorology, hydrology and structural building (Wan Zin, Jemain, & Ibrahim, 2009). LM depends on arrange insights where the first until the point when fourth request relate to measures of area, scale, skewness, and kurtosis separately.

2.3 Correlation

Correlation is an important process in finding relation between two variables. In statistic, the common coefficient used in finding correlation is Pearson correlation coefficient. It is otherwise called the sample correlation coefficient (r), product moment correlation coefficient, or coefficient of correlation (Kutner, Nachtsheim, Neter, & Li, 1996). It quantifies the straight connection between two arbitrary factors. For example, when the estimation of the indicator is expanded or diminished by some sum, the other variable changes relatively with it as such straightly (Zou, Tuncali, & Silverman, 2003).

2.4 ENSO Indexes

Methods for determining the occurrence of ENSO is called ENSO indexes. There are several indexes that are being used by researches nowadays in determining ENSO. Sea Surface Temperature (SST), Southern Oscillation Index (SOI), Oceanic Nino Index (ONI) and Multivariate ENSO index (MEI) are some of popular ENSO indexes.

SST is the most widely used index in determining ENSO. The process of SST is described next. SST in the equatorial Pacific is accompanied and reinforced by changes in atmospheric circulation. Through atmospheric dynamical processes, such changes extend well beyond the tropical Pacific, and influence, for example, the patterns of seasonal winds, rainfall and temperature worldwide (Davey et al., 2014). As being described before, the occurrence of SST anomalies made such an impact to the temperature globally. According to Bobrowsky (2013), anomalies here means climatological or normal values associated with the mean seasonal cycle.

Malaysia is also being affected by the anomalies that happened in the ocean. Since Malaysia is situated in the middle of two huge seas, the Pacific Ocean toward the east and the Indian Ocean toward the west its atmosphere is additionally emphatically affected by regular atmosphere inconstancy related with these seas (Tangang et al., 2012). Further understanding of SST especially on its region, is depicted in Figure 2.1.

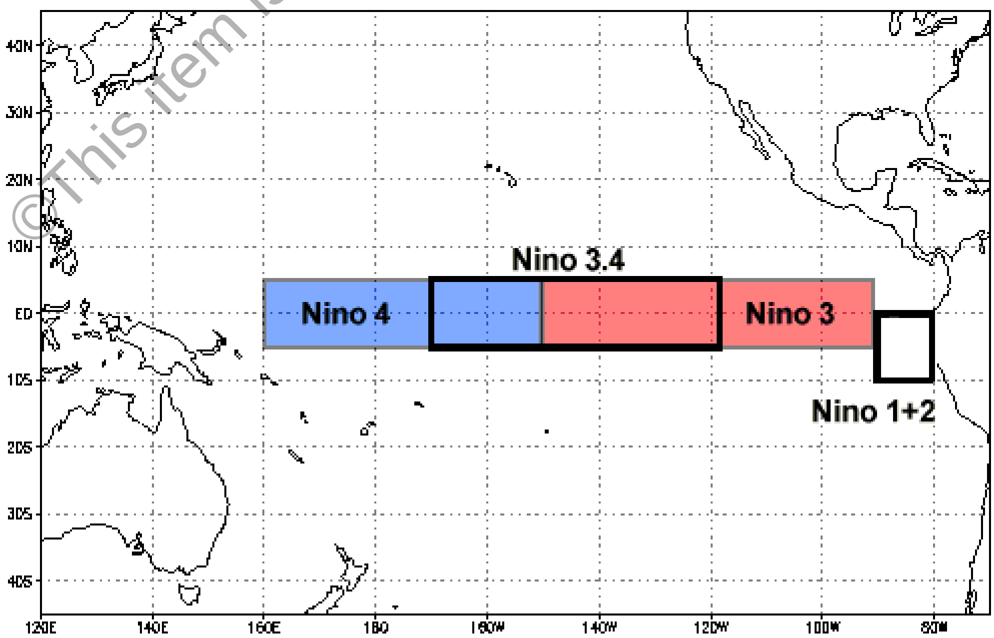


Figure 2.1: SST Nino region in equatorial pacific. (NOAA Climate.gov, 2016).