

STATISTICAL ANALYSIS ON RAINFALL AND DISCHARGE DATA: IN  
STUDY AREA IN JOHOR RIVER BASIN

AZWANIZA BINTI POHARAN @ BUNARI

A dissertation submitted in fulfillment of the  
requirements for the award of the degree of  
Master of Mathematics (Statistics)

Faculty of Sciences  
Universiti Teknologi Malaysia

JANUARY 2014

## **ACKNOWLEDGEMENT**

In the name of ALLAH S.W.T, the Most Gracious and the Most Merciful Lord, I want to express my gratitude of thanks to ALLAH S.W.T. for giving me a good health to enable me to complete this master thesis.

First and foremost, I would like to express my deepest thank to my helpful supervisor Dr. Shariffah Suhaila Binti Syed Jamaludin for their helps, supports and guidance that really benefited me in completing this study. All the kindness and interesting discussion that were given to helped me a lot. I really appreciated the opportunity of completing my master thesis under their supervision.

Lastly, I would like to give my hearty thank to my beloved parents for their unconditional love, inspiring support, encouragement and positive spirit. All of your supports and contributions were really meaningful to me and only ALLAH S.W.T can recall your kindness. Thank you so much.

*To my beloved family and friends!*

## **ABSTRACT**

Johor river is one of the several river that drain into Tebrau Strait ( Selat Tebrau). The Johor river catchment is centered about 35E, 0.5N. Rainfall and flow data for the period 1980 – 2011 were used in this study. The time series of monthly values of rainfall and discharge were analyzed using statistical methods. Trend analysis was performed using Mann-Kendall trend test. This was done in an attempt to determine whether or not there have been any significant change in rainfall and discharge over this catchment. Forecasting analysis which is ARIMA model and Smoothing Technique were used to select best model to predict for 5 years ahead. Best fitted model were analyze using statistical criteria of Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) The best model used for forecasting the rainfall and discharge data is ARIMA model. While rainfall and discharge data analysis for the station slightly show the same trend before and after forecast.

## TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
<b>1</b>	<b>INTRODUCTION</b>	
1.1	Introduction	1
1.2	Background of Study	4
1.3	Statement of Problem	6
1.4	Objectives of Study	7
1.5	Scope of Study	7
1.6	Significance of Study	8
<b>2</b>	<b>LITERATURE REVIEW</b>	
2.1	Introduction	10
<b>3</b>	<b>METHODOLOGY</b>	
3.1	Introduction	17
3.2	Descriptive Statistics	18
3.3	Summary Statistics	18
3.4	Non-Parametric Analysis	19
3.5	Forecasting Technique Analysis	22

3.6	Error Measure	30
3.7	Research Framework	31
3.8	Gantt Chart	32
<b>4</b>	<b>RESULT</b>	
4.1	Introduction	33
4.2	Data	33
4.3	Summary Statistic by Station	34
4.4	Stations basis trend analysis	40
4.5	Stations basis forecasting analysis	47
<b>5</b>	<b>CONCLUSION AND RECOMMENDATION</b>	
5.1	Introduction	62
5.2	Conclusions	63
5.3	Recommendations	64

## LIST OF TABLES

TABLE NO	TITLE	PAGE
3.1	Gantt Chart	32
4.1	The names of the stations with latitude and longitude of rainfall data	36
4.2	The names of the stations and descriptive statistics of rainfall daily data	37
4.3	The descriptive statistics of discharge daily data	39
4.4	The homogeneous test of rainfall and discharge data	41
4.5	Mann-Kendall trend test	43
4.6	Beginning of detected trend	45
4.7	ADF test on rainfall data for S7	50

<b>4.8</b>	ADF test on rainfall data for S7 after seasonal difference	51
<b>4.9</b>	Comparison between Seasonal ARIMA Model	53
<b>4.10</b>	The Best of Seasonal ARIMA Model for all Stations Rainfall and Discharge	54
<b>4.11</b>	Comparison between Single Exponential Smoothing, Double Exponential Smoothing and Holt Winters trend and seasonality Model	57
<b>4.12</b>	Comparison between Single Exponential Smoothing, Double Exponential Smoothing, Holt Winters trend and seasonality model and seasonal ARIMA model for S7	58



## LIST OF FIGURES

<b>FIGURE NO</b>	<b>TITLE</b>	<b>PAGE</b>
1	Location of Johor state in Peninsular Malaysia	35
2	Sequential values of statistics from the Mann-Kendall test for all stations	46
3	Monthly rainfall data for S7	48
4	ACF and PACF of S7 after seasonal difference	51
5	Forecasting result on all stations	60
6	The trend of the S7 before and after forecast	61

## LIST OF ABBREVIATIONS

MK	-	Mann Kendall
CV	-	Coefficient of Variation
JB	-	Johor Bahru
PWMK	-	Pre-Whitening Mann Kandell

## **LIST OF APPENDICES**

<b>APPENDIX</b>	<b>TITLE</b>	<b>PAGE</b>
A	Forecasting analysis using ARIMA model	i - xiv

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 Introduction**

Malaysia is a country located in Southeast Asia and straddling the South China Sea. There are two distinct parts in Malaysia being Peninsular Malaysia to the west and East Malaysia to the east. Peninsular Malaysia is located south of Thailand, north of Singapore but east of the Indonesian island of Sumatra, Indonesia. East Malaysia is located on the island of Borneo and shares borders with Brunei and Indonesia. Malaysia consist of high land area (above 180 m above sea level), lowlands (less than 180 m above sea level), plains, basins and coastline.

Johor, "the southern gateway", is the third largest state in Peninsular Malaysia. Johor covering a total area of 19 984 square kilometers of land in the southern part of the peninsula. The state is located between latitudes 1° 20" N and 2° 35" N and longitude 102° turn is 40" T to 100° 16" T. It is bordered to the north of Pahang, Malacca and Negeri Sembilan on the west, Straits of Johor and Malacca Straits to the south and the South China Sea in the East. The capital city of Johor is Johor Bahru, located in the southern part of Johor.

In addition, the forecasting rainfall method can be used in hydroelectric generator industry. Field significance allows the determination of the percentage of tests that are expected to show a trend, at a given local (nominal) significance level, purely by chance, reflecting the cross-correlation in the data. A resampling, approach was used to determine the critical value for the percentage of stations expected to show a trend by chance. Therefore, ability to predict future rainfall and discharge is very beneficial. This model can be applied to predict rainfall and discharge for the next few years. After that, the exact rainfall can be estimated.

## REFERENCES AND BIBLIOGRAPHY

- Nasir Nayan, Jamaluddin Md Jahi, Abdul Latif Mohamed (2010) Perubahan gunatanah dan penduduk zon pinggir pantai Negeri Perak 1984-2004. In: Sakinah Harith, Hasmah Abdullah, Rapeah Suppian, Haliza Abdul Rahman (eds) *Proceeding 2nd National Conference on Environment & Health 2010*, pp. 85-92. Anjuran School of Health Sciences & PERSALA Universiti Sains Malaysia.
- Shaharuddin Ahmad, Noorazuan Md. Hashim (2006) Menganalisis pola dan arah aliran hujan di Negeri Sembilan menggunakan kaedah GIS Poligon Thiessen dan Kontur Isoyet. *Geografia* 3 (2), 1-12.
- Suppiah R, Hennessy KJ (1998) Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990. *International Journal of Climatology* 10, 1141–1164.
- Yue S, Pilon P, Phinney B (2003) Canadian streamflow trend detection: impacts of serial and crosscorrelation. *Hydrological Science Journal* 48, 51-63.
- Jamaludin, S. and Jemain, A. A. (2007). Fitting Daily Rainfall Amount in Peninsular Malaysia Using Several Types of Exponential Distributions. *Journal of Applied Sciences Research.*, 3(10), 1027-1036.
- Nicholas, A. O. (1987). Forecasting seasonal rainfall for agricultural decision-making in northern Nigeria. *Agricultural and Forest Meteorology* 107, 193-205.
- Amendola, A. (2003): Forecasting performance of regime-switching models in hydrological time series. *Giornata di Studio: Metodi Statistici e Matematici per le analisi Idrologiche-Roma*, Am. Soc. Eng. 116, pp. 770–779.
- Box, GEP., Jenkins, GM. (1970): *Time Series Analysis: Forecasting and Control*. Holden-Day, Boca Raton, Fla., USA.
- Amendola, A., Storti G. (1999): A threshold model for rainfall-flow non-linearity. *Book of short papers*, S.Co.
- Morgan RPC (1974) Estimating regional variations in soil erosion hazard in Peninsular Malaysia. *Malaysian Nature Journal* 28, 94-106.

- Shahid S (2009) Rainfall variability and the trends of wet and dry periods in Bangladesh. *International Journal of Climatology*, doi:10.1002/joc.2053.
- Box, G.E.P. and G.M. Jenkins, 1976. *Time Series Analysis: Forecasting and Control*. Revised Edn., Hodey-Day, San Francisco.
- Walter Vandaele, 1983. *Applied time series and Box-Jenkins Models*. Academic Press Inc., Orlando, Florida, ISBN: 10: 0127126503, pp: 417.
- Montgomery, D.C. and L.A. Johnson, 1967. *Forecasting and Time Series Analysis*. McGraw-Hill Book Company, <http://www.abebooks.com/Forecasting-Time-Series-Analysis-Montgomery-Douglas/1323032148/bd>
- Chiew, F.H.S., M.J. Stewardson and T.A. McMahon, 1993. Comparison of six rainfall-runoff modeling approaches. *J. Hydrol.*, 147: 1-36. <http://cat.inist.fr/?aModele=afficheN&cpsidt=4788795>
- ‘Azumi SD, Supiah S, Azmi A (2010) Modelling the distribution of rainfall intensity using hourly data. *American Journal of Environmental Science* 6 (3), 238- 243.
- Wan Ruslan Ismail (1994) *Pengantar hidrologi*. Dewan Bahasa dan Pustaka, Kuala Lumpur. Ward RC (1975) *Principles of hydrology*. McGraw Hill, London. Edisi Kedua.
- Strangesway I (2007) *Precipitation: Theory, measurement, and distribution*. Cambridge University Press, United Kingdom.
- Blaikie P, Cannon T, Davis I, Wisner B (1994) *At Risk: Natural hazards, people vulnerability and disasters*. Routledge, London.
- Fauchereau NS, Trzaska MR, Richard Y (2003) Rainfall variability and changes in Southern Africa during the 20th century in the global warming context. *Natural Hazards* 29, 139- 154.
- Ruksana HR, Syed Hafizur Rahman, Samarendra Karmakar, Ghulam Hussain (2006) Trend analysis of climate change and investigation on its probable impacts on rice production at Satkhira, Bangladesh. *Pakistan Journal of Meteorology* 6 (2), 37- 50.
- Wilby RL, Perry GLW (2006) Climate change biodiversity and the urban environment: A critical review based on London, UK. *Progress in Physical Geography* 30 (1), 73- 98.

- Bravo MT (2009) Voices from the sea ice: The reception of climate impact narratives. *Journal of Historical Geography* 35, 256- 278.
- Enzel Y, Wells SG (1997) Extracting Holocene paleohydrology and paleoclimatology information from modern extreme flood events: An example from Southern California. *Geomorphology* 19, 203- 226.
- Lee JG, Heaney JP, Asce M (2003) Estimation of urban imperviousness and its impacts on storm water systems. *Journal of Water Resources Planning and Management* 129 (5), 419- 426.
- Bravo MT (2009) Voices from the sea ice: The reception of climate impact narratives. *Journal of Historical Geography* 35, 256- 278.
- Hansen, J. W. and Ines, A. V. M. (2005). Stochastic disaggregation of monthly rainfall data for crop simulation studies. *Agricultural and Forest Meteorology* 131, 233–246.
- Nnaji, A. O. (2001). Forecasting seasonal rainfall for agricultural decision-making in northern Nigeria. *Agricultural and Forest Meteorology* 107, 193-205.
- Delleur JW, Tao PC, Kavass ML. 1976. An evaluation of practicality and complexity of some rainfall and runoff time series models. *Water Resources Research* 12: 953–970.
- Salas JD, Obeysekera JTB. 1982a. ARMA model identification of hydrologic time series. *Water Resources Research* 18: 1011–1021.
- Salas JD, Delleur JW, Yevjevich VM, Lane WL. 1980. *Applied Modeling of Hydrologic Time Series*. Water Resources Publications: Littleton, CO.
- Hipel KW, McLeod AE. 1994. *Time Series Modeling of Water Resources and Environmental Systems*. Elsevier: Amsterdam.
- Box GEP, Jenkins GM. 1976. *Time Series Analysis, Forecasting and Control*, Revised edn. Holden – Day: San Francisco, CA.
- Chbab, E.H., Buiteveld, H. and Diermanse, F., 2006. Estimating exceedance frequencies of extreme river discharges using statistical methods and physically based approach. *Wasser- und Abfallwirtschaft* 58: 35-43.



- Eberle, M., Sprokkereef, E., Wilke, K. and Krahe, P., 2001. Hydrological modelling in the river Rhine basin, Part II: Report on hourly modeling Report No. 1338, Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany.
- A.R. Schmidt. Analysis of stage-discharge relations for open-channel flows and their associated uncertainties. PhD thesis, University of Illinois, Urbana-Champaign, 2002. 3, 5,9, 17.
- A. Petersen-Overleir and T. Reitan. Bayesian analysis of stage-fall-discharge models for gauging stations affected by variable backwater. *Hydrological Processes*, 23(21):3057–3074, 2009a. 13.
- Petersen-Overleir. Fitting depth-discharge relationships in rivers with floodplains. *Hydrology Research*, 39(5-6):369–384, 2008. 11, 12.
- Petersen-Overleir. Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 51(3):365–388, 2006. 11, 12.
- Olivier, G. Pierrefeu, M. Scotti, and Blanquart B. (uncertainty of discharge measured with relating curve between level and discharge). In SHF conference, *Hydrological measurements and uncertainties*, 2008.11, 15.