

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Point process models for diurnal variation rainfall data

A thesis presented in partial fulfilment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
IN STATISTICS

at Massey University
Albany (Auckland), New Zealand.

Norazlina binti Ismail

September 2014

Abstract

The theoretical basis of the point process rainfall models were developed for midlatitude rainfall that have different temporal characteristics from the tropical rainfall. The diurnal cycle, a prominent feature in the tropical rainfall, is not represented in the point process models. An extension of the point process models were developed to address the diurnal variation in rainfall. An observed indicator of the rainfall, X is added to the point process models. Two point process models, Poisson white noise (PWN) and Neyman-Scott white noise (NSWN) model were used as the main rainfall event, Y . The rainfall is modelled assuming two cases for the variable X , independent and dependent. Bernoulli trials with Markov dependence are used for the dependent assumption. To allow the model to display the diurnal variation and correlation between hours, the model was fitted to monthly rainfall data by using the properties of two hour blocks for each month of the year. However, the main point process models were assumed the same for each of the 12 blocks, thus having only one set of point process parameters for the models for each month. There are 12 rainfall occurrence parameters and 12 Markov dependence parameters, one for each block. A total of six models were fitted to the hourly rainfall data from 1974 to 2008 taken from a rain site in Empangan Genting Klang, Malaysia.

The PWN and NSWN models with X were first fitted with the assumption that the rainfall indicators are independent between the hours within the two hour block. Simulation studies showed the model does not fit the moments properties adequately. The models were then modified based on a dependence assumption between the hours within the two hour block. These models are known as the Markov X-PWN and Markov X-NSWN models. Both models improve the fit of the moment properties. However, having only one point process model to represent the rainfall events for Malaysia rainfall data was not sufficient. Since tropical rainfall consists of two types of rain, convective and stratiform, the PWN and Markov X-NSWN model were superposed to represent the two types of rainfall. A simple method by assuming non-homogenous PWN process for every two hour block did not fit well the daily diurnal variation. A comparison between the six models show that the superposed PWN and Markov X-NSWN model improved the fitting of mean, variance and autocorrelation. The superposed model was then simplified to an 8-block model to reduce the number of parameters. This modification to the point process models succeeded in describing the diurnal variation in the rainfall, but some of

the models were not able to fit other properties that were not included in the parameter estimation process such as the extreme values.

Acknowledgement

Alhamdulillah. The perfect, most beautiful praise is only for ALLAH. Thank you ALLAH.

2:286 On no soul doth ALLAH place a burden greater than it can bear. It gets every good that it earns, and it suffers every ill that it earns.

2:216 ... But perhaps you hate a thing and it is good for you; and perhaps you love a thing and it is bad for you. And ALLAH knows, while you know not.

I would like to express my sincere appreciation to my supervisors, Dr Barry McDonald and Associate Professor Dr Paul Cowpewart, for their expertise, assistance, guidance, and patience throughout my PhD studies. It was a real privilege for me to have them both as my supervisors, Dr Paul, a well known researcher in his field and Dr Barry, who is willing to learn with me. Thank you so much Dr Barry for your constant support, availability, constructive suggestions, and careful editing which were determinant for the accomplishment of the work presented in this thesis.

I would like to extend my sincere gratitude to Freda, Annette, Yan Ou, Mike Yap and Lyn for their help and support with the basic needs to survive as a student at Massey University. To Dr Peter Kay, thank you for the guidance in my C programming.

To my dear husband, Mohd Nihra Haruzuan, thank you for supporting my dream and thank you for all the sacrifices you have made to see me reach my goals. To my sons, Aqil and Wafi, you guys are a precious gifts from ALLAH, the smiles and laughters brings mama joy everyday! To my parents Ismail and Zainab, my brothers and sisters, Nizam, Ani (Hemly too), Yan (& Adi) and Iwan, thank you for the support and encouragement. Not forgetting my parent in-law and my sister's in-law. I am blessed with a happy and loving family.

Completing this work would have been all the more difficult were it not for the support and friendship provided by the Malaysian community in Albany, the other fellow post-graduates and also the members of the Statistics Department. Special dedication to John Xie, for his help and advice.

Sincere appreciation is also due to the staff of Malaysia Meteorological Department and Drainage and Irrigation Department for providing the daily rainfall data. Finally, I would like to thank UTM and Ministry of Higher Education of Malaysia under the SLAI program for giving me the opportunity and funding which allowed me to further my study in New Zealand.

65:2-3 And whoever fears ALLAH - He will make for him a way out. And will provide for him from where he does not expect. And whoever relies upon ALLAH - then He is sufficient for him. Indeed, ALLAH will accomplish His purpose. ALLAH has already set for everything a [decreed] extent.

Contents

Abstract	i
Acknowledgement	iii
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 Introduction	1
1.2 Rationale of the research	3
1.3 Research objectives	5
1.4 Research approach	5
1.5 Scope of study	6
1.6 Thesis outline	6
2 Literature review	8
2.1 Introduction	8
2.2 Stochastic rainfall models	10
2.2.1 Two-part models	10
2.2.2 White noise model	11
2.2.3 Rectangular pulse model	12
2.2.4 Pulse model	16
2.3 Malaysia point process model	17
2.4 Diurnal variation	18
2.5 Spatial rainfall	19
2.6 Concluding remarks	21
3 Methodology, data description & exploratory analysis	22
3.1 Point process model	22
3.1.1 Poisson white noise model	23
3.1.2 Neyman-Scott white noise model	24
3.1.3 Bernoulli trials with Markov dependence	26

3.1.4	The X latent point process model	27
3.1.5	Moments for two hour blocks	28
3.1.6	Superposed model	30
3.1.7	Model parametrisation and fitting	30
3.1.8	Simulation	31
3.1.9	Model assessment	32
3.2	Data description	33
3.3	Exploratory analysis	34
3.4	Summary	39
4	Non-homogenous PWN	40
4.1	Introduction	40
4.2	PWN model specification & properties	40
4.3	Fitting procedure	41
4.4	Analysis	42
4.4.1	Parameter estimates	42
4.4.2	Moments	42
4.4.3	Monthly moments	48
4.5	Summary and conclusion	49
5	X-PWN model & Markov X-PWN model	50
5.1	Introduction	50
5.2	X-PWN model specification & properties	51
5.3	Fitting procedure & simulation	51
5.4	Analysis	51
5.4.1	Parameter estimates	51
5.4.2	Moments	54
5.4.3	Extreme values	54
5.4.4	Monthly moments	58
5.5	Markov X-PWN model specification & properties	58
5.6	Fitting procedure & simulation	59
5.7	Analysis	59
5.7.1	Parameter estimates	59
5.7.2	Moments	60
5.7.3	Extreme values	63
5.7.4	Monthly moments	63
5.8	Summary and conclusion	69
6	X-NSWN model & Markov X-NSWN model	70
6.1	Introduction	70
6.2	X-NSWN model specification & properties	70

6.3	Fitting procedure	71
6.4	Analysis	72
6.4.1	Parameter estimates	72
6.4.2	Moments	72
6.4.3	Extreme values	78
6.4.4	Monthly moments	78
6.5	Markov X-NSWN model specification & properties	78
6.6	Fitting procedure	81
6.7	Analysis	81
6.7.1	Parameter estimates	81
6.7.2	Moments	82
6.7.3	Extreme values	90
6.7.4	Monthly moments	90
6.8	Summary and conclusion	92
7	Superposed model and 8-block model	93
7.1	Introduction	93
7.2	Superposed Markov X-NSWN and PWN model specification & properties	94
7.3	Fitting procedure & simulation	95
7.4	Analysis	97
7.4.1	Parameter estimates	97
7.4.2	Moments	97
7.4.3	Extreme values	104
7.4.4	Monthly moments	104
7.5	8-block model specification & properties	107
7.6	Fitting procedure	109
7.7	Analysis	109
7.7.1	Parameter estimates	109
7.7.2	Moments	109
7.7.3	Extreme values	117
7.7.4	Monthly moments	117
7.8	Summary and conclusion	119
8	Conclusions and recommendations	121
8.1	Conclusions	121
8.2	Recommendations for future work	123
	Bibliography	125

List of Figures

3.1	PWN plot	24
3.2	NSWN plot	24
3.3	Peninsular Malaysia	33
3.4	Kuala Lumpur rainfall stations	34
3.5	Annual rainfall totals: the Genting Klang data (1973-2008)	35
3.6	The monthly mean rainfall. The Southwest monsoon from May to August and Northeast monsoon from November to February.	36
3.7	Daily mean rainfall	36
3.8	Hourly mean rainfall	37
3.9	Diurnal rainfall pattern for every month	38
4.1	The mean and variance of January, February, March and April	45
4.2	The mean and variance of May, June, July and August	46
4.3	The mean and variance of September, October, November and December	47
4.4	The hourly mean and variance for every month	48
5.1	The mean of 2 hour blocks for every month	55
5.2	The variance of 2 hour blocks for every month	56
5.3	Annual extreme values for X-PWN	57
5.4	The hourly mean and variance for every month	58
5.5	The mean of 2 hour blocks for every month	64
5.6	The variance of 2 hour blocks for every month	65
5.7	The autocorrelation of 2 hour blocks for every month	66
5.8	Annual extreme values for Markov X-PWN	67
5.9	The hourly mean, variance and autocorrelation for every month	68
6.1	The mean of 2 hour blocks for every month	73
6.2	The variance of 2 hour blocks for every month	76
6.3	The autocorrelation of 2 hour blocks for every month	77
6.4	Annual extreme values for Markov X-PWN	79
6.5	The hourly mean, variance and autocorrelation for every month	80
6.6	The mean of 2 hour blocks for every month	86

6.7	The variance of 2 hour blocks for every month	87
6.8	The autocorrelation of 2 hour blocks for every month	88
6.9	Annual extreme values for Markov X-PWN	89
6.10	The hourly mean, variance and autocorrelation for every month	91
7.1	The hourly rainfall from the (a) Markov X-NSWN process (b) PWN process (c) superposed process	96
7.2	The mean of 2 hour blocks for every month	101
7.3	The variance of 2 hour blocks for every month	102
7.4	The autocorrelation of 2 hour blocks for every month	103
7.5	Annual extreme values for Markov X-PWN	105
7.6	The hourly mean, variance and autocorrelation for every month	106
7.7	The hourly mean of 3 hour blocks for every month	108
7.8	The mean of 3 hour blocks for every month	113
7.9	The variance of 3 hour blocks for every month	114
7.10	The autocorrelation of 3 hour blocks for every month	115
7.11	Annual extreme values for 8-blocks model	116
7.12	The hourly mean, variance and autocorrelation for every month	118

List of Tables

3.1	Parameters Description	29
3.2	Sample statistic of lag 1 autocorrelation for pooled subsamples by month. .	37
4.1	Parameter estimates for the pulse depth. The units are mm.	43
4.2	Parameter estimates for the rate of rainfall occurrences. The units are hour ⁻¹	44
5.1	Parameter estimates for the pulse rate λ and depth η . The units hour ⁻¹ for λ and mm for η	52
5.2	Parameter estimates of alpha. The units are hour ⁻¹	53
5.3	Parameter estimates for the pulse rate λ and depth η . The units hour ⁻¹ for λ and mm for η	60
5.4	Parameter estimates of alpha. The units are hour ⁻¹	61
5.5	Parameter estimates of gamma. The units are hour ⁻¹	62
5.6	Model comparison in terms of MSE	63
6.1	Parameter estimates for the pulse depth. The units are hour ⁻¹ for λ and β with η the unit is mm and ν for number of pulses.	74
6.2	Parameter estimates of α . The units are hour ⁻¹	75
6.3	Parameter estimates for the pulse depth. The units are hour ⁻¹ for λ and β with η the unit is mm and ν for number of pulses.	82
6.4	Parameter estimates of alpha. The units are hour ⁻¹	83
6.5	Parameter estimates of gamma. The units are hour ⁻¹	84
6.6	Model comparison in terms of MSE	85
7.1	Parameter estimates for the pulse depth. The units are hour ⁻¹ for all estimates, except η the unit is mm and ν for number of pulses.	97
7.2	Parameter estimates of alpha. The units are hour ⁻¹	98
7.3	Parameter estimates of gamma. The units are hour ⁻¹	99
7.4	Model comparison in terms of MSE	100
7.5	Parameter estimates for the NSWN and PWN process. The units are hour ⁻¹ for all estimates, except η the unit is mm and ν for number of pulses.	110

7.6	Parameter estimates of alpha. The units are hour ⁻¹	111
7.7	Parameter estimates of gamma. The units are hour ⁻¹	112
7.8	Model comparison in terms of MSE	117