

Temporal Trends of Discrete Extreme Events – A Case Study

S N Rahmat¹, N Jayasuriya², M Bhuiyan² and M S Adnan¹

¹ Faculty of Civil and Environmental Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat Johor, MALAYSIA

² Senior Lecturer, School of Civil, Environmental and Chemical Engineering, RMIT University, GPO Box 2476, Melbourne, Victoria 3001, AUSTRALIA

E-mail: nazahiya@uthm.edu.my

Abstract: Investigating trends in discrete events is essential for the study of changing patterns of extreme events. Temporal trends in the inter-arrival times of occurrence of drought events were examined for 21 selected stations across Victoria, Australia. In the present study, the Standardize Precipitation Index (SPI) was applied for 12-month time scale to identify drought. A drought event here is defined as a period in which the SPI is continuously negative and reaching a value of -1.0 or less. Often, nonparametric tests are commonly used to test for trends including in discrete events. However, discrete events are not constant because of the presence of zero values or non-normality of data. The methodology applies to long-term records of event counts and is based on the stochastic concepts of Poisson process and standard linear regression. Overall, of the 21 stations, 15 showed statistically significant increasing frequency indicates those events are becoming more frequent. Only one station gave insignificant result. The remaining 5 stations showed the time between events was significantly increasing designates droughts are becoming less frequent.

Keywords: Parametric trend test poisson process standardize precipitation index (SPI).

1. Introduction

Measuring trends in extreme events has received much attention in recent years. The International Panel on Climate Change (IPCC), in its assessment of the vulnerability of the regional impacts of climate change, reports the potential risk of variations in extreme events [1]. Parametric and non-parametric methods are the most commonly used techniques to detect significant trends in hydro-climatologic time series [2]. While nonparametric techniques, such as a Mann-Kendall test for trend or a Sequential Mann-Kendall test for a change point have been widely used and applied in the case of extremes, they were not developed for dealing with discrete events as such [3].

When common statistical tests are used to test for trends in discrete events, fundamental assumptions are often violated because of the use of counted data, the presence of zero values, or non-normality of data, where little is known about the distribution of the underlying dataset [4]. An alternative approach common in hydrology has been to apply regression analysis with a trend variable to hydrologic time series transformed to an approximate normal distribution. For instance, Ross [4] applied the logarithmic transformation to stream flow, as consistent with a lognormal distribution.

Time series of climatological extreme events, however, become counting processes once a threshold level is identified. A counting process is simply a stochastic process by which a running count of the number of events occurring in a prescribed time increment is observed [5,6]. Keim and Cruise [6] applied Poisson process and linear trend analysis to detect changes in precipitation. They concluded that the methods provided in the paper to examine trends in rainfalls could also be applied to the temporal and spatial frequencies of other



extreme events or any discrete random event. The Poisson distribution has also been fitted to the frequency of hurricanes by [1,7,8].

The trend of Standardised Precipitation Index (SPI) was investigated for more than 100 years of data using non-parametric trend technique Mann Kendall (MK) to detect wet (increasing) and dry (decreasing) periods across Victoria, Australia [1]. Out of 70 stations, 21 stations showed significant decreasing trend, indicating an increase in magnitude of dry periods. This finding highlights the need to look at the patterns as to whether the intervals between extreme (drought) events are becoming shorter and frequent. Moreover, Australia's climate has always been variable and in particular, prone to drought.

Using a method proposed by Keim and Cruise [6], this present study will test for any gradual trend in the rate of occurrence of drought through an analysis of processed inter-arrival times using parametric method, Poisson process and standard linear regression.

2. Data and Methods

The locations of the 21 rainfall stations in Victoria, Australia and their characteristics for this study are shown in Figure 1 and Table 1. The number of years of available data varied from one station with a minimum of 60 years to two stations with a maximum of 160 years of monthly data.

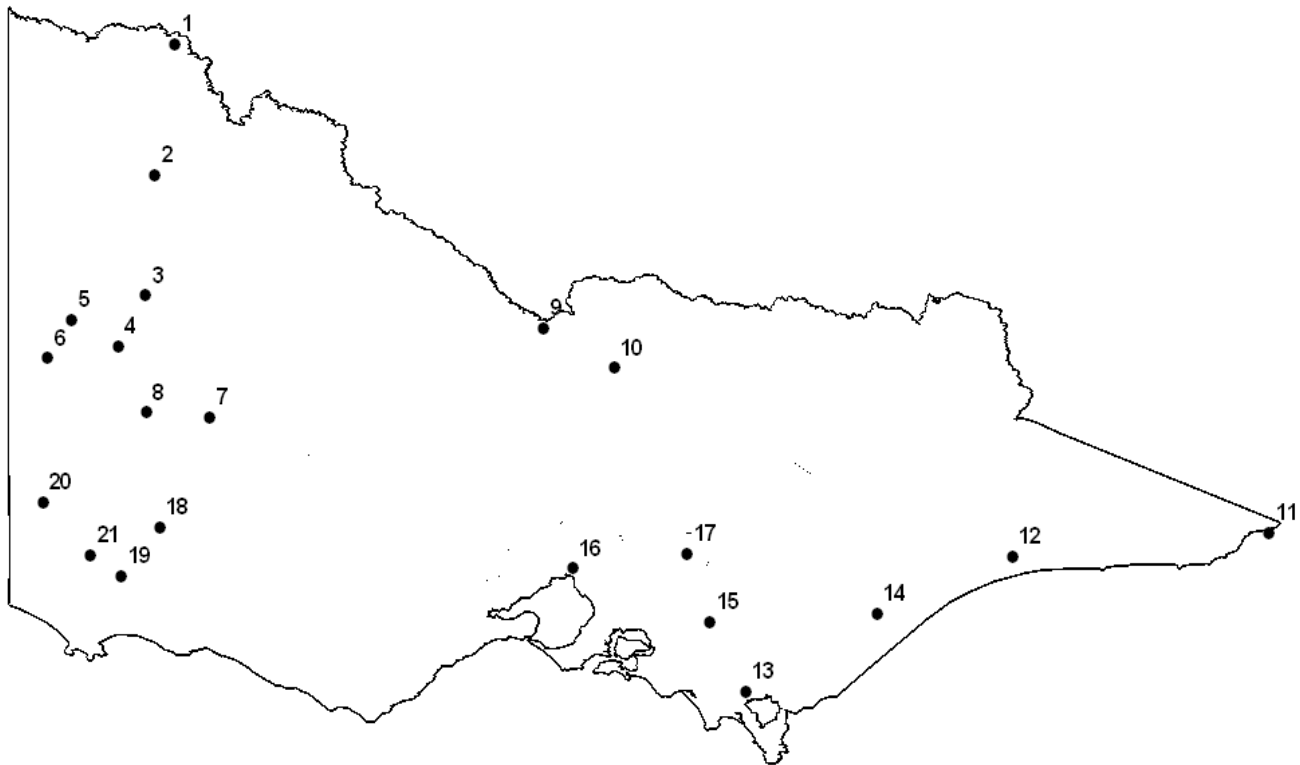


Figure1. Locations of the study sites.

2.1 The Poisson process

A Poisson process is intended to represent events that occur at random or in more general terms, a stochastic model for a series of events is called a counting process [3]. A counting process is generally a stochastic process by which a running count of the number of events occurring in a given time increment is observed [5]. In the Poisson distribution, there is equality of mean and variance represented by the parameter λ . If the arrival rate remains constant with time, then the process is homogeneous. Another very important property of the Poisson process is that the inter arrival times between realizations of successive events are known to be exponentially distributed [4].

Table 1. Description of selected rainfall stations for the analysis.

Stn. No.	Stn. Name	No of years of data
1	Irymple	103
2	Walpeup Research	74
3	Rainbow	115
4	Nhill (Woorak)	82
5	Yanac North	115
6	Kaniva	130
7	DrungDrung	83
8	Natimuk	106
9	Echuca Aerodrome	132
10	Tatura	71
11	Gabo Island Lighthouse	146
12	NowaNowa	64
13	Foster (Post Office)	128
14	East Sale Airport	69
15	Warragul	124
16	Melbourne Regional Office	157
17	Warburton	97
18	Cavendish (Post Office)	128
19	Branxholme (Bassett)	67
20	Dergholm (Hillgrove)	113
21	Merino	114

3. Results and Analysis

3.1 Temporal trends in drought events

In the present study, the arrival rate of drought events for all stations were tested to classify its distribution. As mentioned earlier, the most important counting process is the arrival rate of events follows a Poisson distribution which is, mean and variance must be equal or nearly equal. However, arrival rates are not always temporally constant and always showing non homogeneous Poisson process. The Poisson process relies upon two underlying assumptions; the number of events occurring in time is independent of its previous history, and no more than one event can occur in a short interval of time [9,10]. In this case, the Poisson model represents the number of randomly occurring infrequent storm events over a specified threshold.

Time series of the SPI at 21 stations selected across Victoria, are analyzed for temporal trends. The droughts events and the inter-arrival time were identified based on the SPI value, which is continuously negative and reaching a value of -1.0 or less. As for examples, the results for five stations are plotted in Figure 2. As proposed by Cox and Lewis [11], the series of inter-arrival times begins with the onset of an event because the length of time from some arbitrary point to the first event does not belong to the same distribution of intervals between events. The number of droughts events recorded varies due to the time series selected which ranging from 14 to 24. It can be seen that for each station inter-arrival times have high variability. For Irymple station, the inter-arrival times range from 10 to 99 months, while for Rainbow, Edenhope, Dookie and MRO, the inter-arrival times range from 6 to 144, from 5 to 135, from 12 to 140 and from 10 to 194 months, respectively.

3.2 Trend analysis of inter-arrival times of droughts using linear regression

In the present study, the natural log of the intervals between successive events (Y_i), or successive groups of events (in months), on the independent variable (X_i), which is the cumulative frequency of months from the occurrence of the first event was identified. Advantages of using linear regression on transformed interarrival time data generated by a Poisson process include successive events are known to be independently distributed, and the probability distribution is moderately non-normal [5,6].

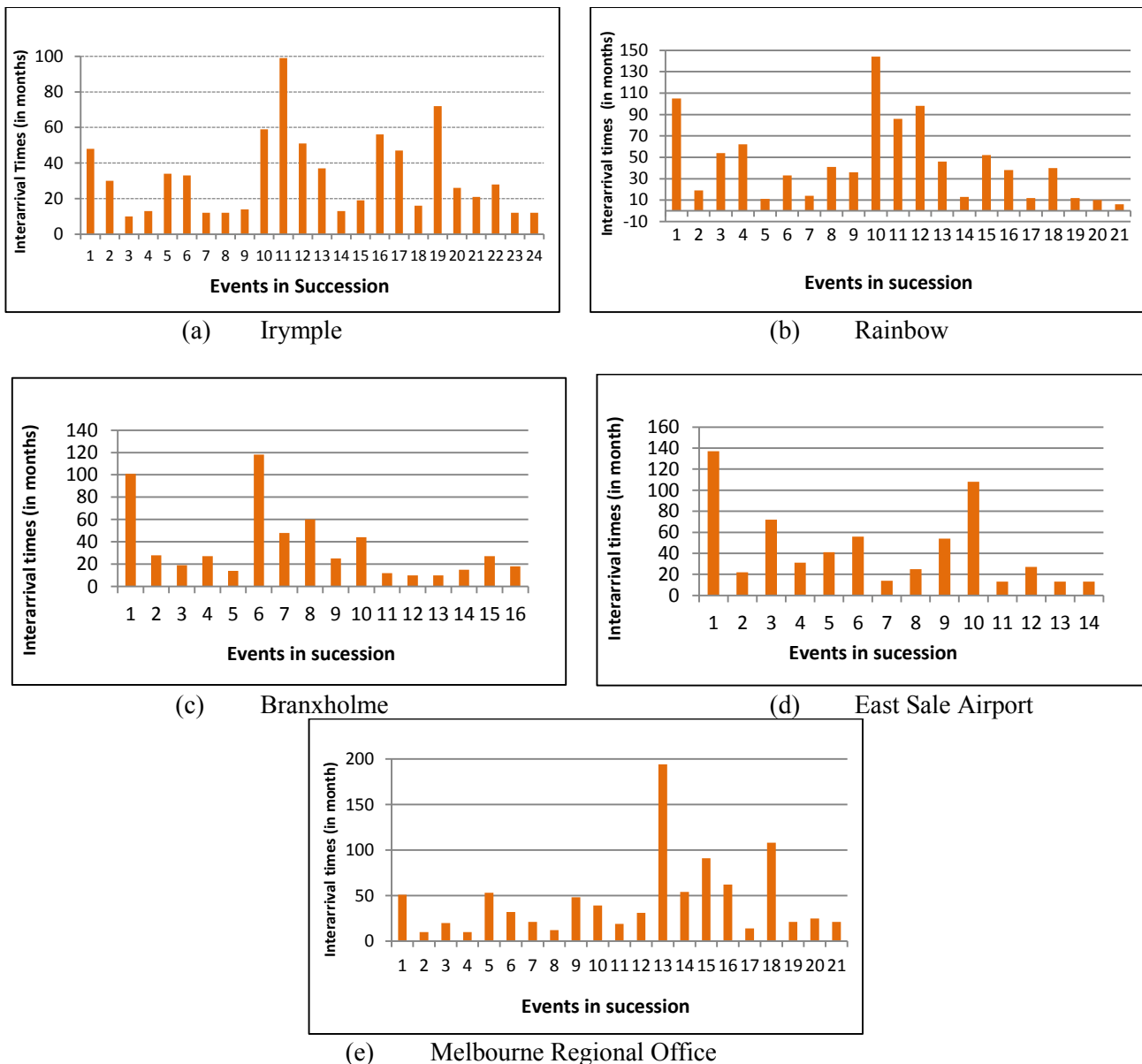


Figure 2. Recurrence intervals of droughts for all stations.

As for example, the SPI values for 12-month time scale and linear regression analysis for Irymple station are given in Figure 3. The identification of the drought events was based on SPI drought class threshold ($SPI = -1$ and less) (see Figure 3a). A number of major drought events were identified including the World War II drought, the 1982-1983 drought and the recent drought (Big Dry). Figure 3b displays the number of months between each drought event. Irymple station shows variability of inter-arrival times between events, with intervals 11 and 19 being the longest of the series. Those two events were in 1951 and 1985, respectively. Events in succession are plotted across the x axis and the number of months from the end of one event to the beginning of the next event (in sequence) is shown on the y axis. The x axis in Figure 3c shows the intervals from the occurrence of the first event in the series. In Figure 3d, the graph shows the number of months transformed into $\ln$ and then plotted on the y axis. The data are then plotted with the linear least squares fit. A similar approach was applied to other four stations and the linear regressions of the log transformed intervals are shown in Figure 4. Table 2 shows the sample sizes and significance levels of the linear regression analysis. A positive sign for the Pearson, r signifies that the time between events is increasing and those events are becoming less frequent. In contrast, a negative sign indicates that the intervals between events are becoming shorter and that the frequency of events is temporally increasing. As shown in Table 2, the slope ($r = 0.04$)

depicts that the events are becoming less frequent but is insignificant. All the other three stations (Edenhope, Dookie and MRO) are similar, showing decreasing patterns of frequencies with r values of 0.01, 0.21 and 0.28, respectively. In contrast, the Rainbow station ($r = -0.31$) has an insignificantly increasing pattern of frequencies. This trend analysis was then applied for the rest of the 16 stations and the results are given in Table 2. Overall, of the 21 stations, 15 show statistically significant increasing frequency and those events are becoming more frequent. The stations showing significantly increasing frequency are Walpeup Research, Nhill, Kaniva, Drung Drung, Natimuk, Echuca, Tatura, Gabo Island, Nowa Nowa, Foster, East Sale, Warragul, Avenel, Heathcote, Newstead, Cavendish, Branhholme and Dergholm. The remaining 5 stations (Irymple, Yanac North, Melbourne Regional Office (MRO), Warburton and Merino) show the time between events is increasing indicates droughts are becoming less frequent.

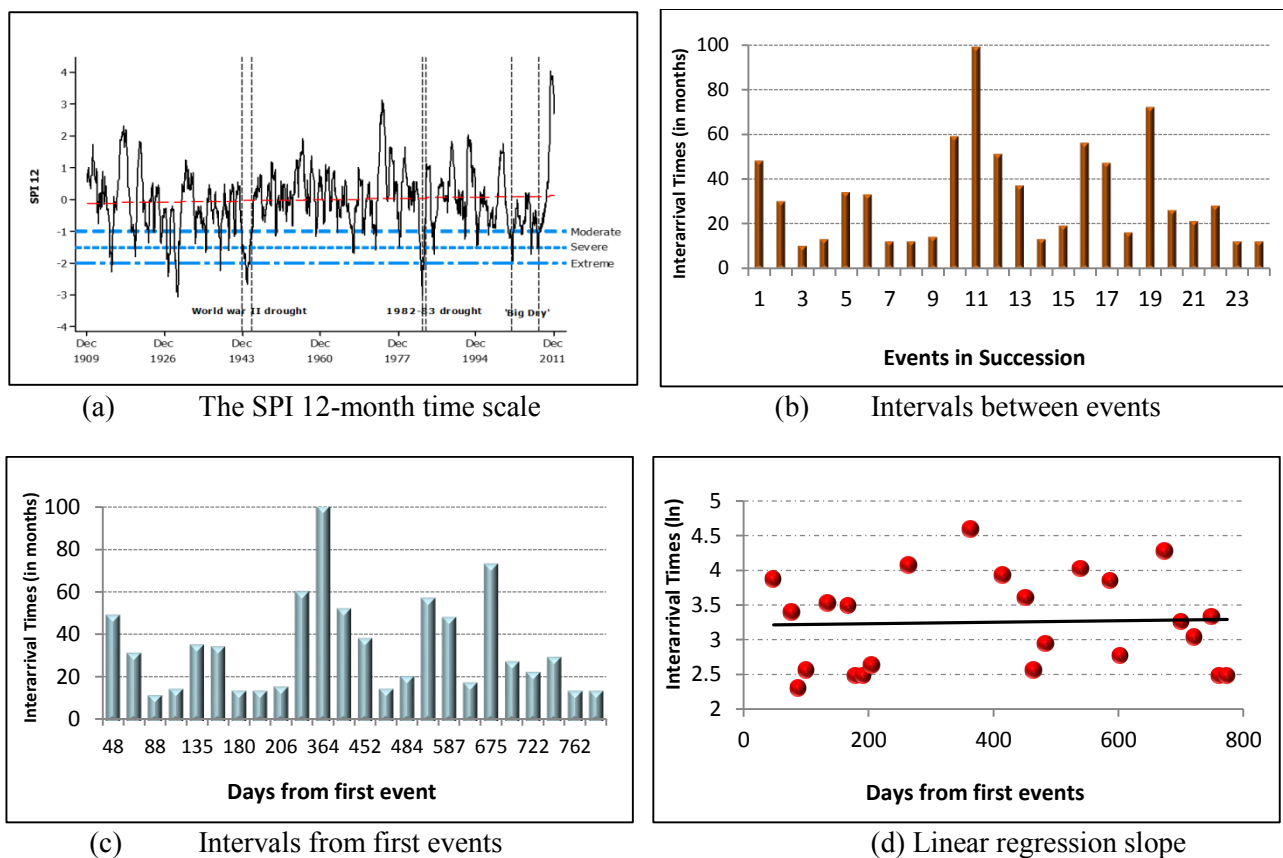


Figure 3. Time series of droughts for Irymple station.

Overall, the following steps need to be carried out to identify the trend for any discrete random events:

- 1) The arrival rate of drought events for all stations are tested to classify its distribution.
- (2) Once data are identified as Poisson, the inter-arrival times are known to follow the exponential distribution.
- (3) These data are transformed by converting the raw values into natural logarithms ($\ln$) so that they follow the normal distribution.
- (4) The data are then plotted with the linear least squares fit.
- (5) Pearson's correlation coefficient, r is identified to indicate whether the time between events is increasing, or decreasing.

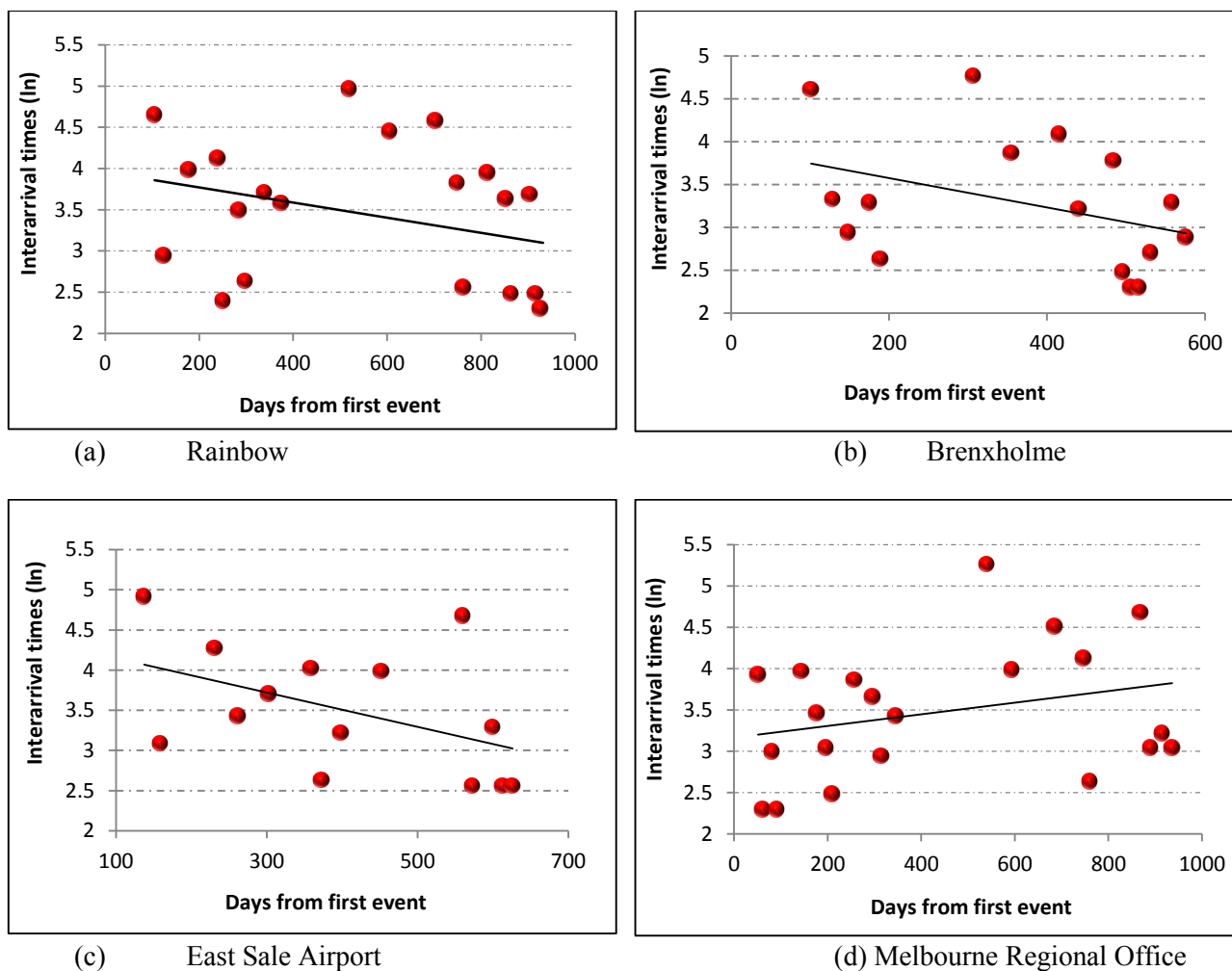


Figure 4. Linear regression of drought events at all stations.

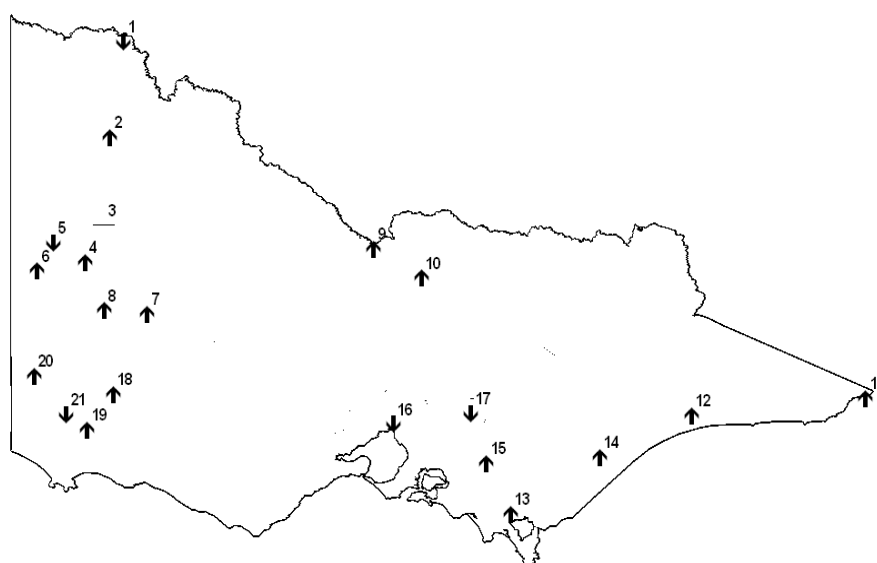


Figure 5. Trend analysis for all 21 stations (- : No trend; ↑: Increasing frequency; ↓: Decreasing frequency).

Table 2. Results of linear regression analysis, r = Pearson's correlation coefficient, p = significance of slope

No	Station	Events	r	p
1	Irymple	24	0.04	0.8
2	Walpeup Research	15	-0.34	0.01
3	Rainbow	21	-0.31	0.2
4	Nhill (Woorak)	18	-0.12	0.04
5	Yanac North	24	0.11	0.17
6	Kaniva	33	-0.03	0.03
7	Drung Drung	20	-0.04	0.05
8	Natimuk	25	-0.06	0.03
9	Echuca Aerodrome	35	-0.01	0.02
10	Tatura	14	-0.65	0
11	Gabo Island	42	-0.2	0
12	NowaNowa	14	-0.45	0.01
13	Foster (Post Office)	32	-0.2	0
14	East Sale Airport	14	-0.46	0
15	Warragul	30	-0.05	0.01
16	Melb. Reg. Office	21	0.28	0.2
17	Warburton	23	0.24	0.36
18	Cavendish	36	-0.04	0.01
19	Branxholme	16	-0.38	0
20	Dergholm	27	-0.11	0.01
21	Merino	30	0.08	0.05

*Results in boldface indicate significant increasing frequency.

4. Conclusions

Measuring trends of extreme events or disasters are real challenges in water resources management. However, it is essential as extreme events are becoming more common due to the impact of climate change. Also, the ability to detect long term variations and trends is critical to begin to detect any observed changes.

The trend of the SPI was investigated to detect trends for 70 selected stations across Victoria, Australia. 21 stations showed significant decreasing trend, showing an increase in magnitude of dry periods. Further, the trends were identified to measure the intervals between extreme (drought) events using parametric method. The parametric method, Poisson process is found to be appropriate for measuring trend in discrete random events. The advantage to using parametric method is that the test is generally more efficient.

Overall, of the 21 stations, 15 showed statistically significant increasing frequency and those events are becoming more frequent and one station gave insignificant value. The remaining 5 stations showed the time between events is increasing indicate that droughts are becoming less frequent. Based on these results, some of the study areas in Victoria may be facing frequent drought. Therefore, it is essential to use an appropriate methodology to develop suitable strategies to mitigate the impacts of future droughts and properly understand past droughts to be able to forecast the future wherever possible.

Acknowledgements

The authors are grateful for the funding provided by the Ministry of Higher Education under the Fundamental Research Grant Scheme (Vote No. 1232).

References

- [1] Otto, M. (1998). Multivariate Methods: In: R. Kellner, J. M. Mermet, M. Otto and H. M. Widmer, Eds., Analytical Chemistry. Wiley-VCH, Weinheim, pp.916.

- [2] Tabari, H., S. Marofi, A. Aeini, P. H. Talaee and K. Mohammadi (2011). Trend analysis of reference evapotranspiration in the western half of Iran. *Agricultural and Forest Meteorology*, 151(2), pp.128-136.
- [3] Allen, L. J. S. (2010). *An introduction to stochastic processes with applications to biology*. Chapman & Hall/CRC Press, Boca Raton, Florida, pp.461.
- [4] Ross, S. M. (2014). *Introduction to probability models*. 11th Edition. Elsevier, San Diego, California, pp.759.
- [5] Katz, R. W. (2002). Stochastic Modeling of Hurricane Damage. *Journal of Applied Meteorology*, 41(7), pp.754-762.
- [6] Keim, B. D. and J. F. Cruise (1998). A Technique to Measure Trends in the Frequency of Discrete Random Events. *Journal of Climate*, 11(5), pp.848-855.
- [7] Bove, M. C., J. J. O'Brien, J. B. Eisner, C. W. Landsea and X. Niu (1998). Effect of El Niño on U.S. Landfalling Hurricanes, Revisited. *Bulletin of the American Meteorological Society*, 79(11), pp.2477-2482.
- [8] Elsner, J. B. and B. H. Bossak (2001). Bayesian Analysis of U.S. Hurricane Climate. *Journal of Climate*, 14(23), pp.4341-4350.
- [9] Ashkar, F. and J. Rousselle (1983). The effect of certain restrictions imposed on the interarrival times of flood events on the Poisson distribution used for modeling flood counts. *Water Resources Research*, 19(2), pp.481-485.
- [10] Bezak, N., M. Brilly and M. Šraj (2014). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5), pp.959-977.
- [11] Cox, D. R. and P. A. W. Lewis (1996). *The Statistical Analysis of Series of Events*. Springer, London and Colchester, pp.279.