Statistical technique for assessing trends in drought occurrence – a case study

<u>S. Rahman</u>^a, L. N. N. Jayasuriya^b and M. A. Bhuiyan^b

^a PhD Student, School of Engineering, RMIT University, Melbourne, Australia ^b Senior Lecturer, School of Engineering, RMIT University, Melbourne, Australia E-mail: <u>s3514429@student.rmit.edu.au</u>

Drought is an extreme weather event affecting human wellbeing and the natural environments. Abstract: In addition, climate change is exacerbating the situations by making drought conditions more common. Identifying and quantifying trends in drought events thus pose a real challenge in managing water resources systems. This study evaluated the trends in the frequency of occurrence of drought events in the Yarra River catchment in south-east Victoria, Australia. The Standardized Precipitation Index (SPI) and the Standardized Hydrological Drought Index (SHDI) were applied for 6-months timescale to identify meteorological as well as hydrological drought events, respectively. Drought indices are the most common and effective tools for the detection of drought periods. This study considered a drought event to commence as soon as the SPI/SHDI value becomes less than zero. Generally, trends in any data series are investigated by non-parametric statistical techniques. However, the time series of rare and extreme weather events are not suited to these traditional approaches due to the presence of zero values and non-normality of data. In this study, the changing pattern in the inter-arrival time of successive drought events were assessed with a statistical parametric method resulting from the concepts of Poisson Process and standard linear regression technique. The SPI and the SHDI were employed to detect drought incidences for two rainfall and two streamflow stations, respectively, within the study area. Overall, all the stations showed statistically insignificant decreasing trends in the rate of inter-arrival times of drought events which indicates that the drought events are becoming more frequent. This study will assist water managers to assess and develop appropriate mitigation strategies to overcome the future drought impacts.

Keywords: Drought, trend, frequency, parametric method, linear regression

Rahman et al., Statistical technique for assessing trends in drought occurrence-a case study

1. INTRODUCTION

A changing climate leads to the changes of characteristics of discrete and extreme climatic events such as droughts, floods, bushfires and cyclones. The consequences of such extreme events have profound impact on human lives and the natural environments. The Intergovernmental Panel on Climate Change (IPCC, 2015) have identified the variability and trends in climatic extremes as crucial aspects towards an improved understanding of past and potential future global change. The changes in the hydro-climatic data series are usually detected by the parametric and non-parametric tests (Rahmat et al., 2015, Rahmat et al., 2012). Conventional statistical non-parametric techniques like Mann-Kendall (MK), Sequential Mann-Kendall (SMK), Spearman's Rho (SR), are, however, often inappropriate for testing trends in rare weather events as the time series data are nonstationary (Frei and Schär, 2001, Rahmat et al., 2016). The population and sampling distribution generally govern non-parametric methods. Nevertheless, limited information about both distributions make nonparametric tests less powerful than the parametric tests. Based upon order statistics (i.e., an ordering of the data from smallest to largest or vice versa), the non-parametric techniques use empirical plotting formulas to assign exceedance probability where the rejection probabilities are highly uncertain. Hence, non-parametric techniques are unable to deal with the occurrence rate of rare climatic events successfully. Many researchers proposed several parametric tests to accommodate the specific nature of discrete events in the trend test (Katz, 2010, Katz, 2002). However, analysis of time series of rare events often addresses problems with parametric techniques. When common statistical techniques are applied for trend detection in randomly occurring events, basic hypothesis of the techniques are compromised because of the application of the enumerated data, the presence of zero values and non-normality of data, where little is known about the distribution of the underlying dataset (Ross, 2014). Due to such limitations in these traditional approaches, an alternative method was selected for analysing trend variability. The Standard Linear Regression method, which is very common in hydrology, was used for assessing the changing pattern of hydro-climatic data series.

When a truncation level is applied to the time series of rare climatological occurrences, the series transforms into a stochastic counting process (Keim and Cruise, 1998). In this process, a running enumeration of number of events occurring in a preselected time period is observed (Ross, 1993). In a counting process, the rate of occurrence of events usually follow Poisson distribution (Deng et al., 2016, Keim and Cruise, 1998). Eventually parametric methods can be applied to Poisson distributed data series which are initially produced from the counting process. Keim and Cruise (1998) tested the precipitation time series data by Poisson process and detected the trend of arrival rate of heavy rainfall using linear regression technique. They also suggested that the methods used for trend detection in extreme rainfall events could be adopted for the analysis of temporal and spatial frequencies of any random discrete event. Many researchers used the statistical technique of Poisson regression to identify the trends in the frequency of occurrence of extreme events (Katz, 2010, Deng et al., 2016, Rahmat et al., 2016).

Australia is one of the driest continents in the world. The occurrence of extreme weather events in Australia is very common due to its variable spatial and temporal hydroclimatic characteristics. The recent droughts have severely stressed water resources systems introducing the need for an appropriate drought analysis method (Rahman et al., 2016). The major objective of this study was to identify the trends in the frequency of drought events in the Yarra River catchment using a parametric method. The arrival rate was Poisson distributed and data were tested for trends by adopting a linear regression approach. Drought events were initially determined using widely accepted drought indices method. In this study, the Standardized Precipitation Index (SPI) and the Standardized Hydrological Drought Index (SHDI) were applied to identify meteorological as well as hydrological drought events, respectively. According to McKee et al. (1993), the SPI can detect and monitor drought events at multiple timescales. Additionally other researchers suggested that hydrological drought events (river discharges or reservoir storages) could be assessed by time scale started from 6 months or above López-Moreno et al., 2013). However, this study adopted a 6-month scale as a sound option for identifying both hydrological and meteorological drought events. To better understand the recent climatic fluctuations in Australia, it is worthwhile to investigate the temporal variability of extreme events. This study aimed to find out the inter-arrival times between the drought events and to determine if they are becoming frequent.

2. DATA AND METHOD

The study area was the Yarra River catchment in south-east Victoria, Australia (Figure 1). This catchment is used for supplying water to the greater Melbourne region for more than 100 years. It supports a wide range of water uses such as: urban water supply for 4 million people, agricultural, horticultural and industrial demand

and flow requirements for maintaining environmental flows to protect the ecology of the river (Melbourne Water, 2016). However, the recent droughts have severely pressurized the water supply systems, enhancing the urgency of drought analysis in this catchment. Moreover, the consequences of climate changes also intensify the drought effects and accelerate the need for study of drought frequency. Thus, for analysing trends in drought frequency, four stations (two rainfall and two streamflow) were selected in the upper reach of the Yarra River. Monthly rainfall and streamflow data were collected from the Bureau of Meteorology Australia website (http://www.bom.gov.au). The



Figure 1. Study Area

description of the stations is provided in Table 1. For analysis purposes, the data from the period between 1979 and 2014 were selected for all the stations. The time interval between successive events generally follow nonhomogeneous Poisson process and the arrival rate is not always constant. In Poisson process, the mean and variance of the counts, represented by the parameter λ , are equal or nearly equal. The Poisson process highly depends on two assumptions: (1) the number of events occurring in time is not controlled by its previous history and (2) only one event can occur in short interval of time. Another very important property of the Poisson process is that the interarrival times between realizations of successive events are known to be exponentially distributed (Ross 1993). In this study, the Poisson process represented the discrete drought events below a specified threshold value. The Poisson distributed data were then tested for determining the trends in the interarrival time of drought events.

Rainfall stations	Number of years of data	Beginning Year	Ending Year	Elevation (m)	Mean Annual Rainfall (mm)
Reefton	46	1970	2015	321	1086
Warburton	102	1915	2016	240	1390
Streamflow station	Number of years of data	Beginning Year	Ending Year	Catchment Area (km²)	Annual Average (Flow) (GL)
Aldermans Creek	35	1979	2014	24	8.33
Walshes Creek	35	1979	2014	55	15.7

Table 1. Descriptions of stations

3. RESULTS

3.1. Inter-Arrival Time of Drought Events

The trends of inter-arrival time of the drought events for the four stations selected across the Yarra River catchment were analyzed. In this study, the drought events were identified as soon as the SPI/SHDI reached negative values and continued until they attained positive values. The inter-arrival time is defined as the time period from the end of one drought event to the commencement of the next event (Shiau, 2006). According to Cox and Lewis (1966), the interarrival time series should start with the commencement of the event rather than from the point from which data are first available. This is 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 results of interarrival time of drought events of the selected four stations are depicted in Figure 2. In this section, the data period was taken between 1979 and 2014 for comparison purposes. For the two rainfall stations (Reefton and Warburton), the SPI 6-month scale identified 27 and 24 meteorological drought events within the specified time period. Twenty (20) and twenty-three (23) hydrological drought events were identified for Aldermans and Walshes creeks, respectively by the SHDI 6-month scale. The number of meteorological drought events in two rainfall stations were higher than the hydrological drought events found in two streamflow stations. Meteorological drought can be recovered by rainfall experienced during a drought event. Yet, due to the delay in time between the occurrence of rainfall and the water flowing into a stream, hydrological drought continues to occur even after termination of meteorological drought (Liu et al., 2012). Moreover, within this period, the next meteorological drought may occur while the preceding hydrological drought still continues to progress. It can be seen from Figure 2 that the inter-arrival time of drought events have high variability. At the Reefton station, the inter-arrival times of meteorological drought events ranged from 3 months to 29 months. However, for the other selected rainfall station (Warburton), the inter-arrival time of drought events was fluctuating between 4 months and 30 months for the period between 1979 and 2014. The inter-arrival time of hydrological drought events in the two streamflow stations (Aldermans Creek and Walshes Creek) varied from 4 months to 57 months and 5 months to 55 months, respectively. Hydrological drought events exhibited high variability on inter-arrival times compared to meteorological drought events for the same time period.



Figure 2. Inter-arrival time of drought events (a) Reefton, (b) Warburton, (c) Aldermans, (d) Walshes

3.2. Trend Analysis by Standard Linear Regression

In this study, the trends in drought frequency were determined for the selected stations. According to Cox and Lewis (1996), the gradual trend in the rate of occurrence of discrete events could be identified by the standard liner regression method. The advantages of using this method are that it employs Poisson distributed data in which successive events are known to be independently distributed and the probability distribution is moderately non-normal. Once data are identified as Poisson, the inter-arrival times are known to follow the exponential distribution. These data were then converted into the natural logarithms (ln) so that they follow the normal distribution (Rahmat et al., 2016). This study used the natural log of the intervals between consecutive events as the dependent variable (Yi) (in months) and the cumulative frequency of months from the occurrence of the first event as the independent variable (Xi) to describe the trends in drought frequency.

As mentioned earlier, the SPI and SHDI 6-month timescale were used to identify meteorological and hydrological drought events, respectively. The SPI 6-month series in Reefton (rainfall station) and the SHDI 6-month scale in Aldermans Creek station (streamflow) were plotted in Figure 3 (a & b). Both indices could identify several major drought events including the 1982-1983, 1997-1998 and 2002-2003 events. These events were also well documented by the Bureau of Meteorology (BoM, 2003). The inter-arrival times between meteorological drought events at Reefton station using the linear regression method is presented in Figure 4. The events in successions were plotted in the x axis in Figure 4(a) and the number of months between the end of one event to the staring of another event in sequence was designated in the y axis. In Figure 4(b), the time periods in months from the occurrence of the first event to the occurrence of another event were displayed in the x axis and values in y axis are the same as in Figure 4(a). The log normal values of the intervals between each drought event were plotted on y-axis and the times (in months) from the starting of the first event were displayed in the x axis in Figure 4 (c). The data were plotted with the linear least square fit (regression line) as shown in Figure 4(c). This station exhibited high variability in the time period between each drought event which was highlighted in Figure 4 (c) with the shortest interval time of 3 months (22^{nd}) inter-arrival time in Figure 4(a) from November 2008 to January 2009. Nevertheless, the longest inter-arrival time of 29 months occurred between 2010 and 2012 for the same station (24th inter-arrival time in Figure 4(a)). Warburton station disclosed the highest inter-arrivals times of 30 months between 1992 and 1995 (10th inter-arrival time). Within this period, there were no trace for drought occurrences. In Warburton station, the 17th event had the shortest inter-arrival time occurred from October 2005 to January 2006. After 14 years (1996-2009) of below average rainfall in 2010 and 2011 the drought was broken through back-to-back La Niña events in 2010 and 2011

giving high inter-arrival times for both rainfall stations between 2010 and 2012. In the Aldermans Creek station, the inter-arrival time of drought events was highest between 1992 and 1997, indicating that there was no drought event within this period. After 1997 during the millennium drought period (1996-2009), there were a number of frequent drought events. The interval was shortest between 2005 and 2006. Similar results were obtained from the other streamflow station. As like rainfall stations, two streamflow stations also exhibited high inter-arrival times between 2010 and 2012 due to La Niña effect.



Figure 3. Time series of (a) SPI 6-month scale of Reefton station, (b) SHDI 6-month scale of Aldermand Creek station



Figure 4. Inter-arrival time of drought events in Reefton station

The linear regressions of the log transformed inter-arrival time value for the other three stations were depicted in Figure 5. Table 2 presented the sample size and significance level of the linear regression analysis where *r* is defined as the Pearson Correlation Coefficient and *p* stood for significance of the slope of the regression line. A positive slope indicates the time interval between events are increasing and that the events are becoming less frequent. In contrast, a negative slope represents that the inter-arrival times of the events are becoming shorter and that the frequency of the events is increasing. Hence, all the stations exhibited insignificant negative slopes. That means the intervals between drought events are decreasing and the drought events are becoming more frequent in the study catchment. Yet, Walshes Creek station showed no precise trend in inter-arrival times of drought events at all as it possessed very small slope value. The slopes were signified by the χ^2 test (at a 0.05 level). Due to high variability in inter-arrival times of drought events for each of the stations, the slopes, (depicted in Figure 4(c) and Figure 5(a, b and c)) were insignificant. The sign of the *r* represents the association between λ (variance) and time. For Reefton and Warburton station, the Pearson Correlation coefficient values were -0.1 and -0.15 respectively. In case of the Aldermans Creek station and Walshes Creek, the *r* values were -0.13 and -.005, respectively. To summaries, although there is no relationship between the inter-arrival time and starting of the first event, the negative slopes of the regression lines for all the stations indicated that the drought events are becoming more frequent in the selected catchment. Hence, proper drought management is needed to avoid any subsequent devastation associated with drought events.



Figure 5. Linear Regression of drought events (a) Warburton, (b) Aldermans Creek, (c) Walshes Creek

S	Events	r	р			
Rainfall	Reefton	27	- 0.1	0.63		
(SPI)	Warburton	24	-0.15	0.49		
Streamflow	Aldermans creek	20	-0.13	0.60		
(SHDI)	Walsesh Creek	23	-0.005	0.98		
~						

Table 2. Results of liner regression

r= Pearson Correlation coefficient, p= significance of the slope

In general, the study of trend analysis in time series data always relate the statistics to the underlying physics behind the observed patterns. This usually could be the ultimate purpose behind statistical analyses of earth science data. Droughts are becoming more frequent in the study area which may be correlated to the shifting of climatic conditions in South-Eastern Australia. Since variability in inter-arrival times between drought events is moderate within the observed data, however, inferences made beyond the periods of record are limited. Therefore, the slope component of the regression models and their probabilities of significance are considered valid for the period of record under examination only. This study considered only four stations in the selected catchment. However, in future, a large number of stations could be engaged to describe the trends in drought frequency in a more granular manner.

4. CONCLUSIONS

This study presented a statistical framework and methodology for analysing the changing pattern of drought frequencies across the Yarra River catchment. The limitation of parametric approaches was overcome by the use of counted inter-arrival data series which followed a Poisson distribution. The trends in arrival rate of drought events were then tested through linear regression methods. The inter-arrival times of drought events have high variability. For the two rainfall stations, the interval between discrete drought events varied between 3 months to 30 months. In contrast, the inter-arrival time of such extreme events fluctuated between 4 months to 57 months at two streamflow stations. However, hydrological drought events revealed high variability on inter-arrival times compared to meteorological drought events for the same time period. Due to La Niña events in 2010 and 2011, there were high inter-arrival times for both rainfall and streamflow stations between 2010 and 2012. The results of meteorological and hydrological drought events were consistent with the SPI and

Rahman et al., Statistical technique for assessing trends in drought occurrence-a case study

SHDI 6-month series, respectively. All stations exhibited insignificant decrease in intervals between successive drought events. That means, droughts are becoming more frequent within this catchment. This knowledge of changing patterns of drought occurring in this catchment will provide vital insight to water authorities before planning proper drought mitigation strategy.

ACKNOWLEDGMENTS

The authors would like to acknowledge Bureau of Meteorology in Australia, for providing required rainfall and streamflow data and the Government of Australia for providing the Scholarship to the principal author for conducting this research.

REFERENCES

- Bureau of Meteorology (BoM) (2016). <u>http://www.bom.gov.au/water/hrs/#id=229650A</u> (accessed February 2017).
- Deng, Z., Qiu, X., Liu, J., Madras, N., Wang, X. and Zhu, H. (2016). Trend in frequency of extreme precipitation events over Ontario from ensembles of multiple GCMs. *Climate dynamics*, 46, 2909-2921.
- Cox, D. R., and Lewis, P. A. W. (1966): *The Statistical Analysis of Series of Events*. John Wiley and Sons, 285 pp.
- Frei, C. and Schär, C. (2001). Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. *Journal of Climate*, 14, 1568-1584.
- Intergovernmental Panel on Climate Change, IPCC. (2015). Climate change 2014: mitigation of climate change, Cambridge University Press.
- Katz, R. W. (2002). Stochastic modeling of hurricane damage. Journal of Applied Meteorology, 41, 754-762.
- Katz, R. W. (2010). Statistics of extremes in climate change. Climatic Change, 100, 71-76.
- Keim, B. D. and Cruise, J. F. (1998). A technique to measure trends in the frequency of discrete random events. *Journal of Climate*, 11, 848-855.
- Liu, L., Hong, Y., Bednarczyk, C. N., Yong, B., Shafer, M. A., Riley, R. and Hocker, J. E. (2012). Hydroclimatological drought analyses and projections using meteorological and hydrological drought indices: a case study in Blue River Basin, Oklahoma. *Water resources management*, 26, 2761-2779.
- López-Moreno, J., Vicente-Serrano, S., Zabalza, J., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C. and Morán-Tejeda, E. (2013). Hydrological response to climate variability at different time scales: A study in the Ebro basin. *Journal of hydrology*, 477, 175-188.
- Mckee, T. B., Doesken, N. J. and Kleist, J. (1993) The relationship of drought frequency and duration to timescales. *Proceedings of the 8th Conference on Applied Climatology*. American Meteorological Society Boston, MA, 179-183.
- Melbourne Water. (2016). www.melbournewater.com.au (access in JUNE, 2016).
- Rahman, S., Jayasuriya, N. and Bhuiyan, M. (2016). Evaluation of hydrological drought severity: A case. Water Infrastructure and the Environment, New Zealand Hydrological Society, 1-8.
- Rahmat, S., Jayasuriya, N., Bhuiyan, M. and Adnan, M. (2016). Temporal Trends of Discrete Extreme Events-A Case Study. IOP Conference Series: Materials Science and Engineering. IOP Publishing, 012085.
- Rahmat, S. N., Jayasuriya, N. and Bhuiyan, M. (2012). Trend analysis of drought using Standardised Precipitation Index (SPI) in Victoria, Australia. Hydrology and Water Resources Symposium 2012, Engineers Australia, 441.
- Rahmat, S. N., Jayasuriya, N. & Bhuiyan, M. A. (2015). Precipitation trends in Victoria, Australia. Journal of Water and Climate Change, 6, 278-287.
- Ross, S. M., (1993). Introduction to Probability Models. Academic Press, 556 pp.
- Ross, S. M. (2014). Introduction to probability models, 11th Edition. Elsevier, San Diego, California, pp.759.
- Shiau, J. (2006). Fitting drought duration and severity with two-dimensional copulas. *Water resources management*, 20, 795-815.