

Research papers

Multi-temporal clustering of continental floods and associated atmospheric circulations



Jianyu Liu ^{a,b}, Yongqiang Zhang ^{b,*}

^a Department of Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China

^b CSIRO Land and Water, GPO Box 1700, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:

Received 3 May 2017

Received in revised form 20 October 2017

Accepted 27 October 2017

Available online 28 October 2017

This manuscript was handled by Tim R. McVicar, Editor-in-Chief, with the assistance of Sergio M. Vicente-Serrano, Associate Editor

Keywords:

Climate variability

Clustering of floods

Multi-temporal scale

Cox regression

Kernel estimation

Atmospheric circulation

ABSTRACT

Investigating clustering of floods has important social, economic and ecological implications. This study examines the clustering of Australian floods at different temporal scales and its possible physical mechanisms. Flood series with different severities are obtained by peaks-over-threshold (POT) sampling in four flood thresholds. At intra-annual scale, Cox regression and monthly frequency methods are used to examine whether and when the flood clustering exists, respectively. At inter-annual scale, dispersion indices with four-time variation windows are applied to investigate the inter-annual flood clustering and its variation. Furthermore, the Kernel occurrence rate estimate and bootstrap resampling methods are used to identify flood-rich/flood-poor periods. Finally, seasonal variation of horizontal wind at 850 hPa and vertical wind velocity at 500 hPa are used to investigate the possible mechanisms causing the temporal flood clustering. Our results show that: (1) flood occurrences exhibit clustering at intra-annual scale, which are regulated by climate indices representing the impacts of the Pacific and Indian Oceans; (2) the flood-rich months occur from January to March over northern Australia, and from July to September over southwestern and southeastern Australia; (3) stronger inter-annual clustering takes place across southern Australia than northern Australia; and (4) Australian floods are characterised by regional flood-rich and flood-poor periods, with 1987–1992 identified as the flood-rich period across southern Australia, but the flood-poor period across northern Australia, and 2001–2006 being the flood-poor period across most regions of Australia. The intra-annual and inter-annual clustering and temporal variation of flood occurrences are in accordance with the variation of atmospheric circulation. These results provide relevant information for flood management under the influence of climate variability, and, therefore, are helpful for developing flood hazard mitigation schemes.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Temporal clustering of extreme events, such as heavy precipitation, windstorms, and floods, has attracted considerable attention over the last several years since these events have great social, economic and ecological impacts and disruption (Gu et al., 2016b, 2017; Mediero et al., 2015; Merz et al., 2016; Mumby et al., 2011; Pinto et al., 2013; Villarini et al., 2011, 2013b; Wolff et al., 2016). The clustering of disaster events also affects the insurance industry, for insurance premiums and conservation planning (Mumby et al., 2011). Flood clustering, which means that flood occurrences concentrate in a short time span, has profound

impacts on flood estimation, flood design and flood risk management (Merz et al., 2014, 2016). The clustering of floods may cause correlation in flood series and failure of independence hypothesis (Merz et al., 2016), leading to the increased uncertainty in flood quantile estimation (Koutsoyiannis, 2005) and biased flood design. The impacts of flood clustering on flood estimation depend on proportion of oscillation cycle and time series length (Merz et al., 2016), which can be negligible if the oscillation cycle is distinctly shorter than the series length and vice versa (Jain and Lall, 2001; Hirschboeck, 1988). Therefore, it is necessary to assess the flood clustering at different temporal scales.

Recent studies examined the temporal clustering of floods in several regions, and considered spatial differentiation in the strength of clustering (Gu et al., 2016b; Mediero et al., 2015; Merz et al., 2016; Villarini et al., 2013b). As summarised in Table 1, all reported studies are confined to the Northern Hemisphere and

* Corresponding author at: CSIRO Land and Water, Clunies Ross Street, Canberra 2601, Australia.

E-mail address: yongqiang.zhang@csiro.au (Y. Zhang).

Table 1

Summary of the studies focusing on the temporal clustering of floods. In the 'Main results' column five components, which are related to the main content including in the aims of this study, are identified by the code: (1) investigating intra-annual clustering related to climate index; (2) identifying the time when flood cluster occurs within the year; (3) assessing inter-annual clustering of floods; (4) identifying flood-rich and flood-poor periods; (5) examining the strength variation of inter-annual clustering with the flood magnitude or time scale; and (6) exploring the reasons for flood clustering. The components of (1) and (2) in this table correspond to objectives (1) in the main text; the components (3) and (4) correspond to the objective 2; the components (5) and (6) correspond to the objectives (3) and (4), respectively. NA is used to follow the code if the component does not assess.

Continent	Study	County	River/catchments	POT threshold	Temporal scales	Methodology	Main results
Europe	Mediero et al. (2015)	25 countries	103 rivers	On average 3 floods/year	Seasonal; Inter-annual	Monthly frequency; Dispersion index	(1) NA. (2) Different flood-rich months are found in five regions across Europe. (3) Inter-annual clustering is found in the Atlantic and Continental regions. (4) NA. (5) NA. (6) NA
	Merz et al. (2016)	Germany	68 catchments	On average 1, 3 floods/year and 1 flood/3- and 5-years	Inter-annual; different annual	Dispersion index; Kernel estimation	(1) NA. (2) NA. (3) Most stations show significant clustering; (4) floods in Germany are organised in flood-rich and flood-poor periods. (5) More remarkable for smaller flood threshold and smaller time scales. (6) hypothesis catchment memory is the main reason for flood clustering
Asian	Gu et al. (2016b)	China	Tarim River	On average 2.4–3 floods/year for different stations	Intra-annual; Seasonal; Inter-annual	Cox regression; Monthly frequency; Dispersion index;	(1) NAO and AO are the key factors to intra-annual clustering. (2) Flood-rich months are found between June to August. (3) Significant inter-annual clustering is found in three out of eight stations. (4) NA. (5) NA. (6) NA
	Liu et al. (2017)	China	Poyang Lake basin	On average 2.4–3 floods/year for different stations	Intra-annual; Inter-annual	Monthly frequency; Dispersion index	(1) NA. (2) Flood-rich months are found between May to July. (3) NA. (4) Significant inter-annual clustering is found in Rao River. (5) NA. (6) NA
North America	Villarini et al. (2013a,b)	America	41 catchments in Iowa	On average 2 floods/year	Intra-annual	Cox regression model	(1) Flood occurrence is affected by NAO and PNA; Dependence of flood occurrence on climate indices indicates the existence of clustering. (2) NA. (3) NA. (4) NA. (5) NA. (6) NA
	Villarini (2016)	America	755 catchments across the US	Annual maximum peak discharge	Seasonal	Circular statistics	(1) NA. (2) A strong seasonality of floods is found across the US and varies in different areas; urbanization and regulation weak the seasonality of floods. (3) NA. (4) NA. (5) NA. (6) NA
Oceania	This study	Australia	413 unregulated catchments across Australia	On average 0.5, 1, 2 and 3 floods/year	Intra-annual; Seasonal; Inter-annual; different annual	Cox regression; Monthly frequency; Dispersion index; Kernel estimation	(1) ENSO and SAM are the dominant factors to intra-annual clustering. (2) flood-rich months are from January to March and July to September for northern and southern Australia, respectively. (3) Significant inter-annual clustering is identified over southeastern Australia. (4) 2001–2006 is a flood-poor period for most catchments. (5) clustering decreases with increasing magnitude and increases with timescale for larger floods. (6) the variation of atmospheric circulation mainly causes the seasonal, inter-annual clustering and time-variation of flood occurrence

most consider one to two temporal scales. Little attention has been paid to investigate the temporal clustering at different thresholds and different annual scales (Merz et al., 2016). In addition, most studies have only focused on existence of the flood clustering (Villarini et al., 2013b; Gu et al., 2016b; Mediero et al., 2015; Liu et al., 2017), but not assessing the time of flood clustering, i.e., flood-rich/flood-poor period. Hall et al. (2014) emphasised that future studies should pay more attention to identify flood-rich and flood-poor periods rather than only examining the trends in flooding. Flood frequency is affected by the precipitation and antecedent soil moisture conditions (Norbiato et al., 2008). The frequent floods during a short time span (i.e. flood-rich period) may cause large economic and social losses. For example, Queensland was hit by several disastrous floods during 2010–2011, and suf-

fered AUD 2 billion infrastructures damages and 35 fatalities (Johnson et al., 2016), which was caused by a strong La Niña event in early 2010 (Fasullo et al., 2013; Boschat et al., 2015). However, scarce floods for a long time (i.e. flood-poor period) would also cause severe associated environmental disasters. For instance, since there was no flooding for 10 years, large numbers of trees developed health problems around 2004, and large floodplains and wetlands were under threat in the Lower Murray region (Murphy and Timbal, 2008). It is worth mentioning that this period was during one of the worst droughts (identified as "the Millennium drought") in Australia, which has played an important role on global water and carbon cycles (Zhao and Running, 2010; Fasullo et al., 2013). Therefore, it is pressing to investigate the clustering of floods across Australian continent. In this study we

incorporate a series of methods into an integrated framework to address whether and when the flood clustering occurs for Australian catchments at different temporal scales (Table 1).

Cox regression built by Smith and Karr (1986) is used to examine whether intra-annual variation of flood occurrence rates depends on covariate processes to identify whether the intra-annual flood clustering exists. Using the Cox regression to investigate the flood clustering in Iowa, Villarini et al. (2013b) found the occurrence of floods in Iowa is not independent and affected by climate indices, indicating the possible clustering for the flood occurrence. It is well established that the climate system of Australia is affected by climate indices related to the Pacific and Indian Oceans, such as El Niño-Southern Oscillation (ENSO), Indian Ocean Dipole (IOD), Inter-decadal Pacific Oscillation (IPO) and Southern Annular Mode (SAM) (Ishak et al., 2013; Min et al., 2013; Yilmaz et al., 2017). Summarising recent studies on Australian floods, Johnson et al. (2016) indicated that the global climate indices, such as ENSO, SAM, and IPO, have a strong influence on the variation of occurrence and magnitude of floods at intra-annual and inter-annual scales. Cai et al. (2012) also pointed out that IOD plays an important role in changing the Australian climate. However, little attention has been paid to the climate indices impacts on the flood occurrence time, which is related to intra-annual flood clustering. Investigating the relationships between seasonal flood frequency and climate indices over the central United States, Mallakpour and Villarini (2016) concluded that future work should focus on sub-seasonal scale to address how flood events manifest themselves within one year, and further pointed out that this work can be achieved by using the Cox processes to assess the flood temporal clustering. In this study, we used the Cox regression with these climate indices as covariates to investigate whether the flood clustering exists within one year. To further examine when the flood clustering occurs within one year, this study uses monthly frequency, one of the most common methods used to characterise flood seasonality (Mediero et al., 2015), to examine the flood-rich and flood-poor months for seasonal floods, which are vital for flood management (Gu et al., 2016b).

Inter-annual clustering of floods can be identified by dispersion index that assesses the departure from a homogeneous Poisson process and is often applied in clustering of tropical cyclone events (Mumby et al., 2011). Moreover, to examine whether the flood clustering changes with different annual time scale, we selected aggregation time windows of 1, 2, 3 and 4 years. Using the dispersion index, many studies examine whether there exists the clustering (Mumby et al., 2011; Pinto et al., 2013; Mediero et al., 2015; Gu et al., 2016b), but no study has attempted to identify when the clustering occurs. This study uses the Kernel occurrence rate estimation method to smooth the flood time series and then identifies the flood-rich and flood-poor periods at inter-annual scale.

In this study, we conduct a comprehensive flood clustering study at intra-annual and inter-annual scales, and further examine the flood-rich/flood-poor months and periods. We also investigate the difference of inter-annual clustering between different thresholds and different annual scales from one to five years. Moreover, we try to address the scientific question “why temporal clustering exists for floods”, which remains unclear but highlighted in several recent studies. For instance, Wadey et al. (2014) summarised that ‘Further work will assess clustering mechanisms’. Merz et al. (2016) also concluded that it is necessary to have further mechanism analysis to understand how and to which extent flood clustering is governed by atmospheric variation. We address this question by examining the spatiotemporal variation of atmospheric circulation at intra-annual and inter-annual scales, and the entire temporal series over Australia.

In summary, the main objectives of this study are to:

- 1) Investigate whether the flood clustering exists at intra-annual scale and when it occurs within one year;
- 2) Examine whether the flood clustering exists at inter-annual scale and identify when the flood-rich/flood-poor periods occur;
- 3) Examine the strength variation of flood clustering with different flood thresholds and annual time scales; and
- 4) Explore the underlying mechanism for the formation and variation of flood clustering.

This study presents a generic framework combined with a series of methods to systematically analyse the temporal clustering of flood occurrences at different time scales, which sheds new light on the spatiotemporal variation and possible causes of flood clustering across Australia, and provides a reference for the comprehensive examination of flood clustering and its causes.

This paper is organised as follows. Section 2 shows the data; Section 3 describes the statistical methods used to examine the intra-annual and inter-annual flood clustering; Section 4 summarises main results; Section 5 discusses the results; Section 6 draws conclusions.

2. Data

Daily streamflow data covering the period of 1975–2012 for 780 unregulated catchments were collated by Zhang et al. (2013). To further ensure the quality of our analysis, the catchments with more than 10% missing data and the years with more than 15% missing data were deleted. Finally, streamflow data covering 1976–2010 for 413 catchments with 2.57% missing rate were selected (Figs. 1, 2a). There are thirteen drainage divisions across the Australian Continent. No catchments are available at the SWP and this drainage division is thus not included in our analysis. The catchment attributes for each drainage division are shown in Table 2.

For each catchment, missing values were gap-filled by the best simulation obtained from three calibrated hydrological models: Xinganjiang, SIMHYD and AWRA models (Zhang and Chiew, 2009; Zhou et al., 2013; Zhang et al., 2016). For each catchment, the best model was selected using the highest Nash-Sutcliffe Efficiency (NSE) (Fig. 2b). The “best” model simulation performs well (Zhang and Chiew, 2009; Li and Zhang, 2017), with 10th, 50th, and 90th NSE of daily runoff for the 413 catchments being 0.43, 0.67, and 0.81, respectively.

It is worth noting that the streamflow data for pre-1975 period were also collected by Zhang et al. (2013). But in this period, there are about 40% of the 413 catchments without observations. To meet uniform streamflow data length for the 413 catchments, this study only used the data for the 1976–2010 period, which guarantees that enough dataset is available for space-time contrastive analysis. Correspondingly, climate indices (i.e., ENSO, IOD, IPO, SAM) covering the same period were extracted from Earth System Research Laboratory (<http://www.esrl.noaa.gov/psd/data/climateindices/list/>) and Marshall (2003) (<https://legacy.bas.ac.uk/met/gjma/sam.html>).

3. Methods

Various methods are used in this study to understand the flood clustering at intra-annual and inter-annual scales, and the entire temporal series over Australia. We use POT sampling with four thresholds to extract different severities of floods series. Cox regression and monthly frequency methods are respectively

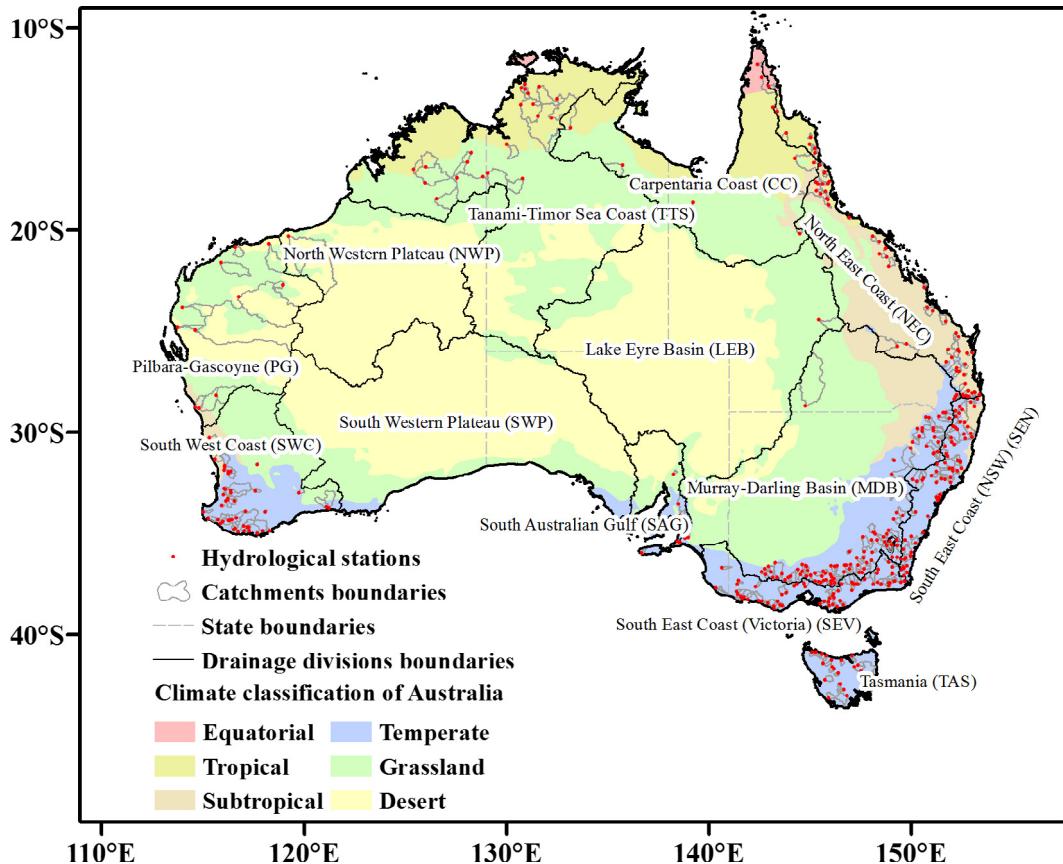


Fig. 1. Location of 413 unregulated catchments, climate regions and Australian river drainage divisions. The abbreviation for each drainage division is in the bracket.

applied to examine whether there exists the intra-annual clustering of floods and when it occurs. To investigate the inter-annual clustering of floods and identify the flood-rich/flood-poor periods, dispersion index and Kernel occurrence rate estimation methods are used. These methods are described in detail in the following sections.

3.1. POT sampling

The time series of floods were extracted by the Peaks-Over-Threshold (POT) sampling technique (Lang et al., 1999). POT sampling counts the flood events over a threshold with more than two-week time interval between two flood events, which avoids counting the same flood events repeatedly, and overcomes the insufficiency of annual maximum (AM) sampling regarding only one flood event per year (Mallakpour and Villarini, 2016). The major inconvenience of POT sampling is that the extracting process of POT series is more complicated than AM series, particularly considering the selection of different thresholds (Mediero et al., 2015). Stedinger (1993) recommended selecting the threshold with an average of 1.65 flood events per year. Mallakpour and Villarini (2016) selected the threshold with an average of 2 flood events per year. There are also several papers that used other thresholds with an average of 2.4–3, 3 and more than 3 flood events per year (Cunderlik et al., 2004; Gu et al., 2016a; Mediero et al., 2015). To comprehensively understand the influence of different thresholds on the flood temporal clustering, we use 4 thresholds to extract the flood series, i.e., POT3 (with average of 3 flood events per year), POT2 (with average of 2 flood events per year), POT1 (with average of 1 flood event per year) and POT0.5 (with average of 0.5 flood event per year). Therefore, the count and temporal data of flood

occurrence were attained for different thresholds. In addition, two different types of temporal data were considered. One is the Julian time data within the year (temporal data II) (Villarini et al., 2013b), which is used analysing the intra-annual clustering by the Cox regression model. Another is the inter-annual time of flood occurrence in the total days of the streamflow time series (temporal data I) (Silva et al., 2012), which is used analysing the inter-annual clustering by dispersion index.

3.2. intra-annual clustering analysis

3.2.1. Cox regression model

Point process of flood events can be characterised by a counting process representation (Villarini et al., 2013a):

$$N_i(t) = \sum_{j=1}^{M_i} \mathbf{1}(T_{ij} \leq t), \quad (1)$$

where M_i is the number of flood occurrences during the year i ; T_{ij} is the time of j th flood event in year i ; $t \in [0, T]$, 0 and T are the start and end time of the year, respectively. $N_i(t)$ is a time series of cumulative sum of flood counts from 0 to T in year i , which can be assumed as a Poisson process:

$$\Pr\{N_i(t) = k\} = \frac{1}{k!} \exp\left\{-\int_0^t \lambda(u) du\right\} \left[\int_0^t \lambda(u) du\right]^k, \quad (2)$$

where $\lambda(u)$ is a non-negative time function representing the time-varying of flood occurrence rate, $u \in [0, T]$. In the case of non-seasonal flood occurrence, the $\lambda(u)$ is a constant. However, if $\lambda(u)$ is not independent for the interval $u \in [0, T]$ and depends on the external physical processes, then the flood occurrence process tends

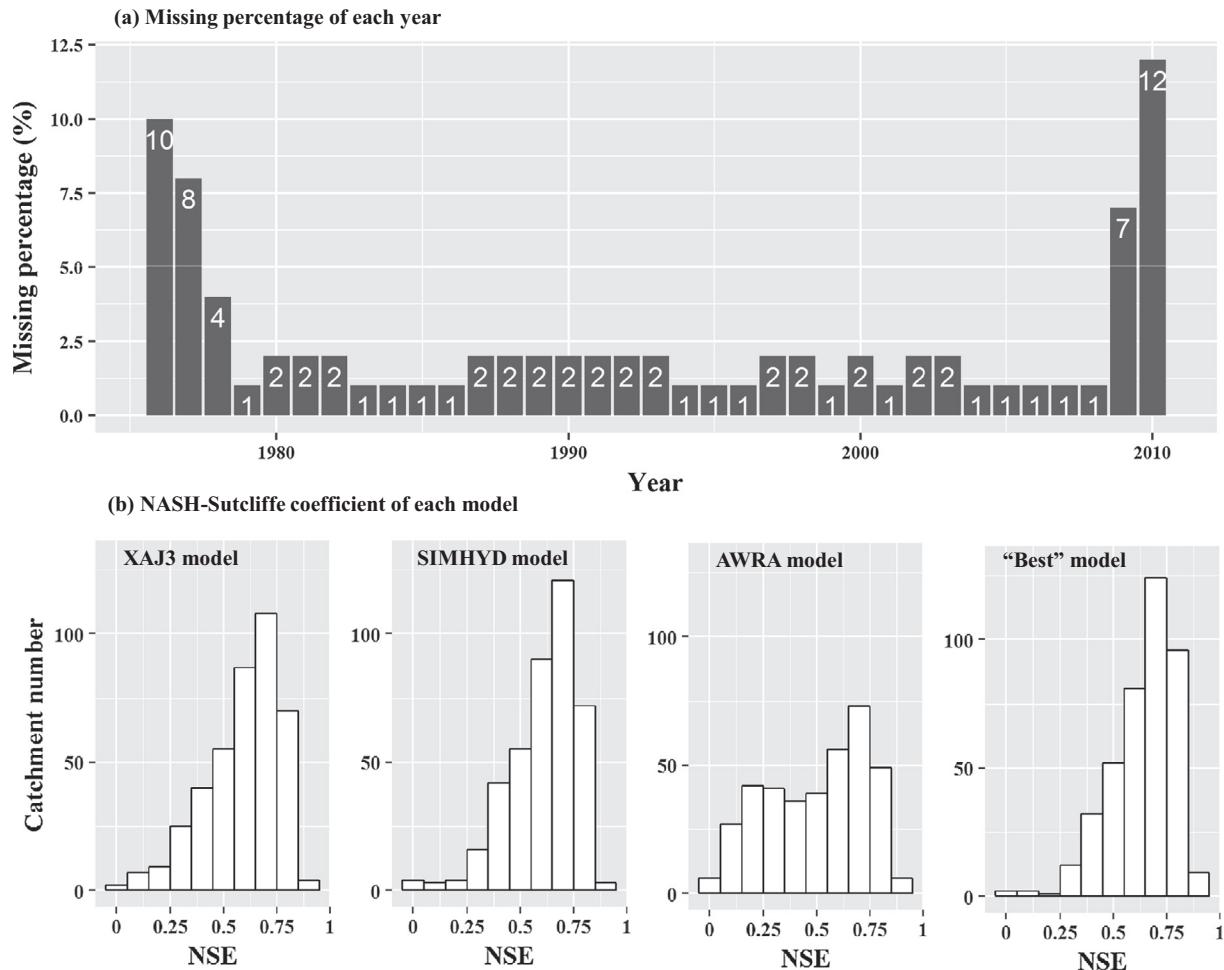


Fig. 2. (a) Histogram showing the total missing percentage in each year. (b) Histogram showing the Nash-Sutcliffe Efficiency (NSE) for different models. XAJ3, SIMHYD and AWRA refer to Xinganjiang, SIMHYD and AWRA models, respectively. For each catchment, the "Best" model is selected from one of the three models with the largest NSE.

Table 2

Statistical summary of catchment attributes in each drainage division.

Location	Drainage division	Abbreviation	Number of catchments	Mean area (km ²)	Min area (km ²)	Max area (km ²)	Mean elevation (m)	Mean precipitation (mm)
Northern Australia	Pilbara-Gascoyne	PG	11	20,940	1419	72,902	351	321
	North Western Plateau	NWP	3	21,186	3732	53,323	335	399
	Tanami-Timor Sea Coast	TTS	21	8088	165	47,651	275	1043
	Lake Eyre Basin	LEB	1	5792	5792	5792	434	556
	Carpentaria Coast	CC	11	4274	170	12,652	365	1016
	North East Coast	NEC	59	979	55	15,851	380	1387
Southern Australia	South East Coast (NSW)	SEN	68	1182	53	16,953	598	1044
	South East Coast (Victoria)	SEV	60	444	53	1974	430	897
	Tasmania	TAS	21	630	87	3285	479	1440
	Murray-Darling Basin	MDB	107	860	61	22,885	660	882
	South Australian Gulf	SAG	6	576	56	2464	329	618
	South West Coast	SWC	45	1099	53	6773	237	685

to exhibit in clustering pattern during a specific time within the year (Gu et al., 2016b; Villarini et al., 2013b). The Cox regression supplies a robust framework to identify whether the flood occurrence process depends on external covariate process (Villarini et al., 2013b).

In the Cox regression model, the distribution of $\lambda(u)$ is a specific function which is depended on the process of covariate variate:

$$\lambda_i(t) = \lambda_0(t) \exp \left[\sum_{j=1}^m \beta_j Z_{ij}(t) \right], \quad (3)$$

where $\lambda_i(t)$ is known as conditional intensity function or hazard function and is the rate of occurrence process in the year i . $\lambda_0(t)$ is a nonnegative time function regarded as a baseline hazard function. $Z_{ij}(t)$ is the j th covariate function in the year i . β_j is the coefficient for the j th covariate.

To select the optimal covariate, we use the stepwise method related to the Akaike Information Criterion (AIC) (Akaike, 1974). The AIC is often used to assess the relative quality of a statistical model, in agreement with the parsimony principle and to avoid model overfitting. The chi-square test is applied to check whether

the final model can describe the occurrence process of floods adequately and satisfy the assumption of the proportion-hazards model (PHM) (Thomas and Reyes, 2014). The null hypothesis is that the Cox regression model satisfies the PHM assumption. The p value of chi-square larger than .05 represents that the assumption is not rejected, and the final Cox regression model can well characterise the flood occurrence. The calculative processes of the Cox regression model are carried out by the freely available *survival* package in R language (Therneau and Lumley, 2011).

3.2.2. Monthly frequency

To assess the intra-annual distribution of floods, the monthly frequency method is used (Macdonald et al., 2010). Non-seasonal floods likely have the same monthly frequency for different months, while seasonal floods with intra-annual clustering pattern usually exist in particular months. Monthly frequencies of floods can be calculated by Cunderlik et al. (2004):

$$FF_m = \frac{F_m}{N} \frac{30}{n_m}, \quad (4)$$

where FF_m denotes the monthly flood frequency in m month, F_m is the annual accumulative flood events occurred in m month, N is total flood events for POT series, n_m is the total days in m month. In non-seasonal floods following the uniform distribution, FF_m belongs to the range from lower bound to upper bound calculated by Eqs. (5) and (6) (Mediero et al., 2014):

$$L_U^N = \frac{N + 11.491}{0.048 \times N^{1.131}}, \quad (5)$$

$$L_L^N = \frac{N - 27.832}{0.199 \times N^{0.964}}, \quad (6)$$

where L_U^N and L_L^N are the lower and upper bounds, respectively, for the non-seasonal floods following a uniform distribution alone the year at a confidence of 95%. If FF_m is larger (smaller) than the upper (lower) bound, then a significant flood-rich (flood-poor) month is detected at 95% confidence level.

3.3. Inter-annual clustering analysis

3.3.1. Dispersion index

The annual number of flood occurrences is also assumed to follow the Poisson distribution characterised by equality of variance and mean: $\text{Var}(M) = E(M) = \mu$ (Mediero et al., 2015; Villarini et al., 2011). The levels of inter-annual flood clustering related to Poisson process are assessed by the dispersion index (DI) (Vitolo et al., 2009):

$$DI = \frac{\text{Var}(M(T))}{E(M(T))} - 1, \quad (7)$$

where $M(T)$ is the series of flood occurrence by T time window. One year is used as the time window T to detect the inter-annual flood clustering. To examine whether the clustering changes with time scales, the time windows in 2, 3, 4, 5 years are also analysed. A positive DI indicates over-dispersion of flood occurrences, which indicates that the flood clustering exists at inter-annual scale. In contrast, a negative DI means under-dispersion, indicating that the flood series shows a more regular pattern than the case in the Poisson process.

The statistical significance of DI is assessed by the Lagrange multiplier (LM) statistic at a 95% confidence level (Greene, 2003):

$$\text{LM} = 0.5 \times \frac{\sum_{i=1}^k (M_i - \hat{\lambda})^2}{k\hat{\lambda}^2}, \quad (8)$$

where $\hat{\lambda}$ is the estimated mean of Poisson distribution for the flood occurrences with length k .

3.3.2. Kernel occurrence rate estimation

Kernel estimation can be used to smooth point process data and assess the temporal variation of the point process. In this study, the Kernel estimation is used to identify the flood-rich and flood-poor periods. The occurrence rate of floods can be calculated by Diggle (1985):

$$\hat{\lambda}_i(t) = h^{-1} \sum_{j=1}^m K\left(\frac{t - T_j}{h}\right), \quad (9)$$

where T_j is the occurrence time (time data I) of the j th flood; K is the Kernel function; h is the bandwidth; $\hat{\lambda}_i(t)$ is the occurrence rate of the flood at time t . The Gaussian Kernel, which can evaluate the efficacy of $\hat{\lambda}(t)$ in the Fourier space, is used to smooth flood occurrence rates (Mudelsee et al., 2004):

$$K(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right), \quad (10)$$

The occurrence rate $\hat{\lambda}_i(t)$ near boundary may be underestimated since there are no data outside the boundary of observed period. To reduce the boundary effect, we use a straightforward method named *reflection* to produce the pseudodata (pT) out of the time interval $[t_0, t_n]$ (Xiao et al., 2013). On the left side of the boundary, i.e., $t < t_0$, $pT(i) = t_0 - [T(i) - t_0]$, covering an amplifying series of 3 times h before t_0 ; likewise for the right side of the boundary. It should be mentioned that the pseudodata is produced by empirical distribution of occurrences nearby the boundary, hence the occurrence rate near the boundary should be analysed cautiously (Mudelsee et al., 2004; Silva et al., 2012).

To examine whether the flood occurrence rate significantly rejects the null hypothesis, i.e., the flood occurrence rate following the Poisson process, the bootstrap resampling technique is used to confirm the confidence intervals (Mudelsee et al., 2004). A significant flood-rich period is identified if the low bound of confidence interval larger than the mean occurrence rate in the homogeneous Poisson process. On the contrary, a significant flood-poor period is identified when the upper bound of confidence interval is smaller than the constant occurrence rate (Merz et al., 2016). Fig. 3 displays an example of the Kernel occurrence rates with 95% confidence interval, which demonstrates a clear distinction in the occurrence rates with and without the pseudodata generation at the beginning and end of the time series. In this case, a significant flood-rich period can be found around 1990 and a significant flood-poor period can be detected around 2004. It is noted that another

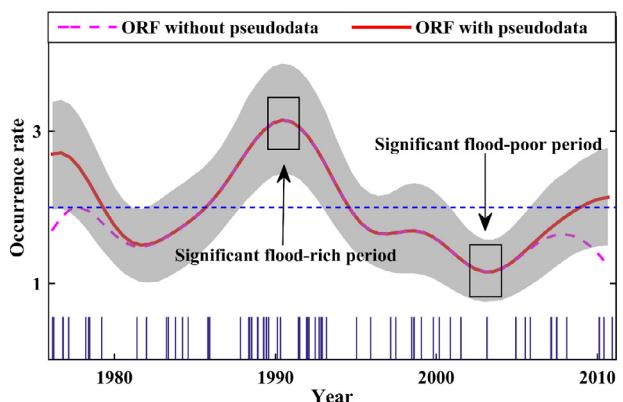


Fig. 3. An example for flood occurrence rates for POT2 with 95% confidence intervals (shaded area). The solid (dashed) curve represents the occurrence rates with (without) pseudodata generation. The blue dashed line represents the long-term mean occurrence rate. The period labelled by a black rectangle above (under) the blue dashed line represents flood-rich (flood-poor) period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

flood-poor period at the end of time series cannot be confirmed and not counted because of the boundary effect.

Another key step for the Kernel occurrence rate estimation is to select a suitable bandwidth, which affects the bias and variance properties of the Kernel estimation. A smaller bandwidth tends to bring a randomness of $\hat{\lambda}_i(t)$, resulting in a higher variance with a lower bias; In contrast, a larger bandwidth leads to an over-smoothing $\hat{\lambda}_i(t)$, leading to a small variation with larger bias. Hence the choice of the bandwidth can be regarded as a compromise between these two cases. Therefore, the cross-validation bandwidth selector is applied to confirm the suitable bandwidth (Mudelsee, 2010).

4. Results

4.1. intra-annual clustering of floods

The flood clustering is often generated by external physical processes represented by covariates (Villarini et al., 2013b). Therefore,

we use the Cox regression to investigate the relationship between intra-annual time of flood occurrence (i.e. time data II) and climate indices such as ENSO, IOD, IPO and SAM. Fig. 4 shows that the *p* value of chi-square test in all catchments are larger than .05, implying that the models comprised by the climate indices can meet the PHM assumption and the Cox regression model can well characterise the flood occurrence. ENSO and SAM have larger influences on the time variation of flood occurrences in the year. There are 191 and 214 out of 413 catchments significantly influenced by ENSO and SAM for the POT3, respectively. However, the impacts of ENSO and SAM tend to weaken with the increase of threshold, with 147 and 189 catchments for POT0.5. The impacts of IOD and IPO on the intra-annual variation of flood occurrences are relatively smaller, but tend to spread with the threshold increase. Nevertheless, IOD and IPO are important predictors for more than 73 catchments. These results indicate that the flood occurrence is non-independent, and temporal clustering exists within one year.

The intra-annual clustering results shown in Fig. 4 do not offer any information on clustering months. To assess the flood-rich/

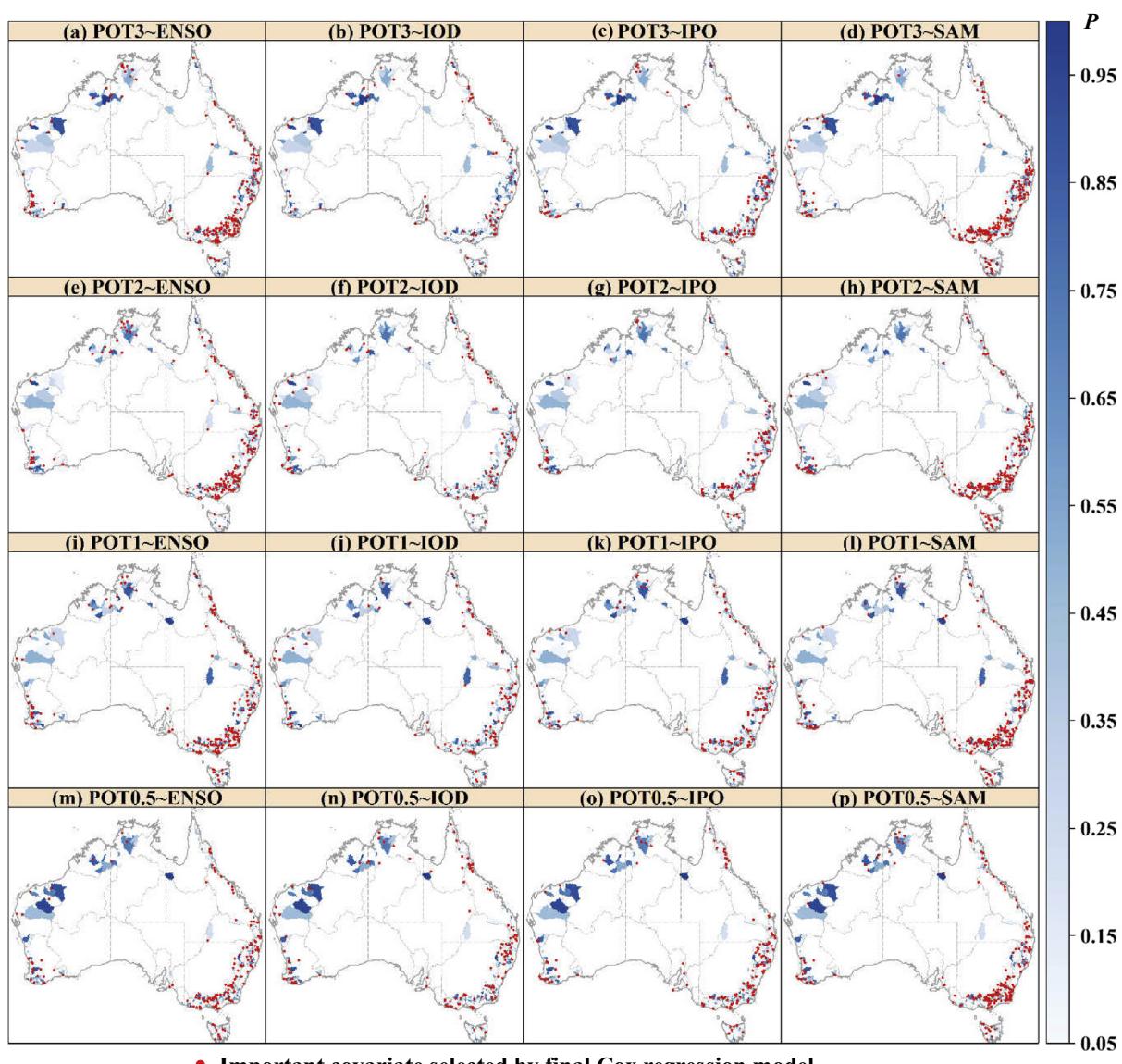


Fig. 4. Maps showing the *p* value (shaded areas) of chi-square for the Cox regression with the climate index as a covariate for different flood thresholds across Australia. Red points represent the catchments regarding climate index as an important covariate selected by the final Cox regression model, indicating that flood occurrences in the corresponding catchments are significantly impacted by the particular climate index. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

flood-poor months within one year, we further analyse the seasonal characteristics of flood occurrence frequency (i.e. count data) by monthly frequency method (Figs. 5 and 6). Fig. 5 shows the spatial distribution of monthly flood frequencies. A distinctive difference for the flood occurrences can be observed among different months, including that northern Australia is characterised by the high frequencies of flood occurrences concentrating between January and March (Fig. 5a–c), and southwestern corner and southeastern Australia are featured by the high flood occurrences occurring between July and September (Fig. 5g–i). Fig. 6 further summarises flood frequencies for each of the 12 drainage division boundaries (Fig. 1). The highest flood frequencies occur during February over five drainage divisions located across northern Australia, including the PG, NWP, TTS, CC and NEC (Fig. 6a–f). Besides, a longer duration of high frequencies is identified between January and March and the frequencies in most catchments exceed the upper bounds at 95% confidence level and for the non-seasonal floods, indicating that January to March are the flood-rich months

across northern Australia. In addition, a long duration dominated by flood-poor months can be observed between June to October for the drainage divisions in northern Australia, with monthly frequency closing to zero or even equal to zero in major catchments.

Over southwest and southeast of Australia, three significant flood-rich months are identified from July to September across the SEV, TAS, MDB, SAG and SWC (Fig. 6h–l). Besides, longer flood-poor months are detected between January and May, with the frequencies smaller than the lower bounds for the non-seasonal floods at 95% confidence level. However, the monthly frequencies for SEN are relatively uniform, with two slightly flood-active months (Fig. 6g).

4.2. Inter-annual clustering of floods

To examine the inter-annual flood clustering, Dispersion index, DI , is applied with the flood occurrence frequency data (i.e., count data). The $DI > 0$ is observed in most catchments across Australia,

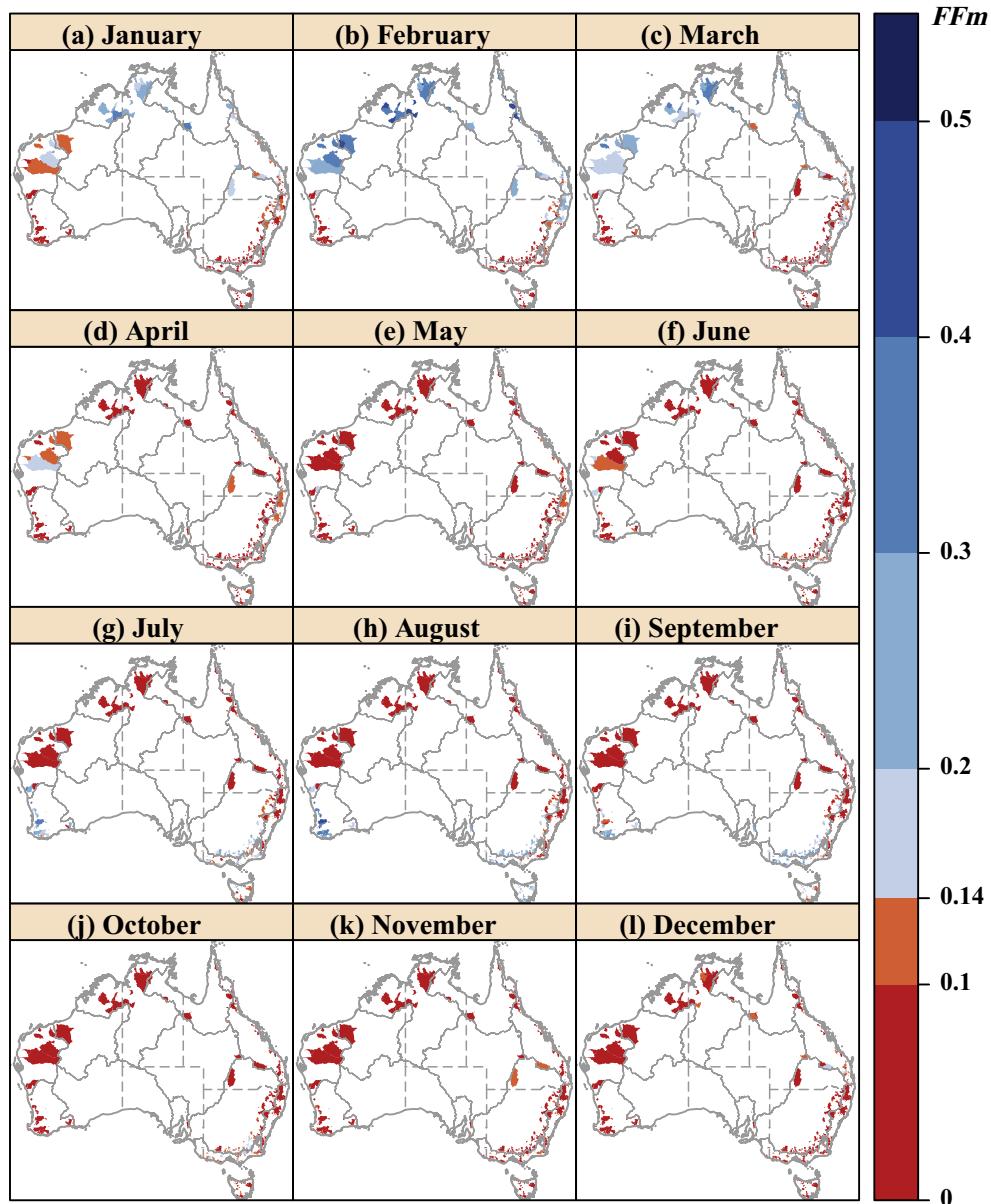


Fig. 5. Maps showing the monthly flood frequency of POT2 by different months. The colour bar represents the monthly flood frequency for each month. The blue colour areas, where the monthly frequency is larger than 0.14, are the catchments with significant flood-rich month. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

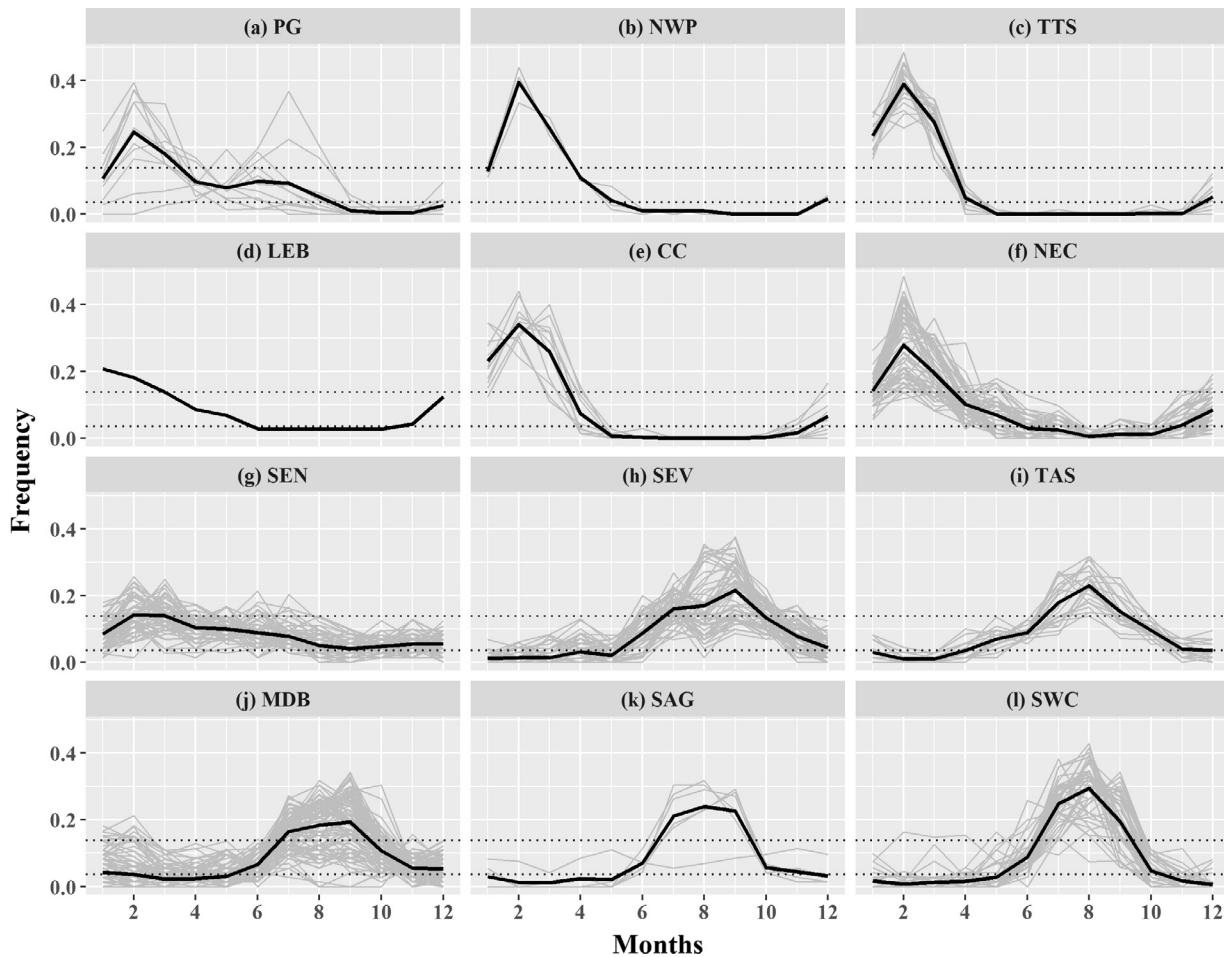


Fig. 6. Monthly flood frequency for POT2 by 12 drainage divisions. Grey lines represent the monthly frequency for each catchment. Solid black lines represent the regional mean. Dotted lines represent the range of confidence interval at 95% confidence level in the status of a non-seasonality pattern. PG locates in western Australia; NWP, TTS, LEB and CC locate in northern Australia; NEC locates in northeastern Australia; SEN, SEV, TAS, MDB, SAG located in southeastern Australia; SWC locates in southwestern Australia. See Fig. 1 or Table 2 for abbreviations of the drainage divisions.

indicating that clustering plays a role in flood occurrence (Fig. 7). Extensive significant clustering can be detected over southwestern corner and southeast of Australian continent with significant over-dispersion at 95% confidence level, resulting in a large variability of flood occurrences in regard to the Poisson process. Especially, large DI up to 2 can be observed over southeastern Australia for the POT3 series (Fig. 7a), indicating that during the flood-rich period the average flood events occur two times more often than those in the normal Poisson process. The strength of clustering is fading out from south to north of Australia. A more regular distribution of flood occurrences is mainly identified over northern Australia with DI smaller than 0. Nevertheless, there are few catchments showing statistically significant results in negative DI, which indicates that under-dispersion is not clear across Australia. Overall, the flood clustering is identified for most catchments with different thresholds, and the strength of clustering tends to decrease with the thresholds increase.

To analyse the effects of different temporal scales (from one to five years) on the flood clustering, we summarise the proportion of catchments with significant clustering at 95% confidence level, for seven drainage divisions with sufficient number of catchments and whole Australia (Fig. 8). The clustering is indeed basin related, and large percentages of catchments with clustering can be found in SEN, SEV and MDB (Fig. 8e, f, h). The percentage of catchments with significant clustering varies with the POT thresholds and time scales. Increasing thresholds tends to decrease the percentage of

catchments with significant clustering, indicated by the smallest percentage occurring with the highest threshold of POT0.5. For the whole Australia (Fig. 8a), in dependence of window size, the percentage of significant clustering is 62–72% in POT3, but drops to 27–53% in POT0.5. However, the significant clustering does not exhibit distinct spatial pattern with different window sizes.

To further identify the flood-rich and flood-poor periods for each drainage division, we use Kernel occurrence rate estimation to analyse the variation of inter-annual time of flood occurrence (i.e. temporal data I). Fig. 9 shows that the temporal variation of flood occurrence rate over the period of 1976–2010 are different for different drainage divisions, particularly between northern drainage divisions and southern drainage divisions. Nonetheless, flood occurrence rates for different catchments within the same drainage division have roughly similar variation characteristics. The period of 1987–1992 witnesses lower flood occurrence rate for most drainage divisions over northern Australia, particularly for the NWP, TTS (Fig. 9b–c), which are identified as significantly (at 95% confidence level) poor. On the contrary, in this period southern Australia has a higher flood occurrence rate except for TAS (Fig. 9g–l). The period of 2001–2006 is identified as with low flood occurrence rates in most drainage divisions across Australia, particularly for southeastern Australia, and all drainage divisions except for TAS exhibit flood poor significantly (at 95% confidence level) in this period (Fig. 5g, h, j and k). The flood occurrence rates in flood-rich period are noticeably larger than those in flood-poor

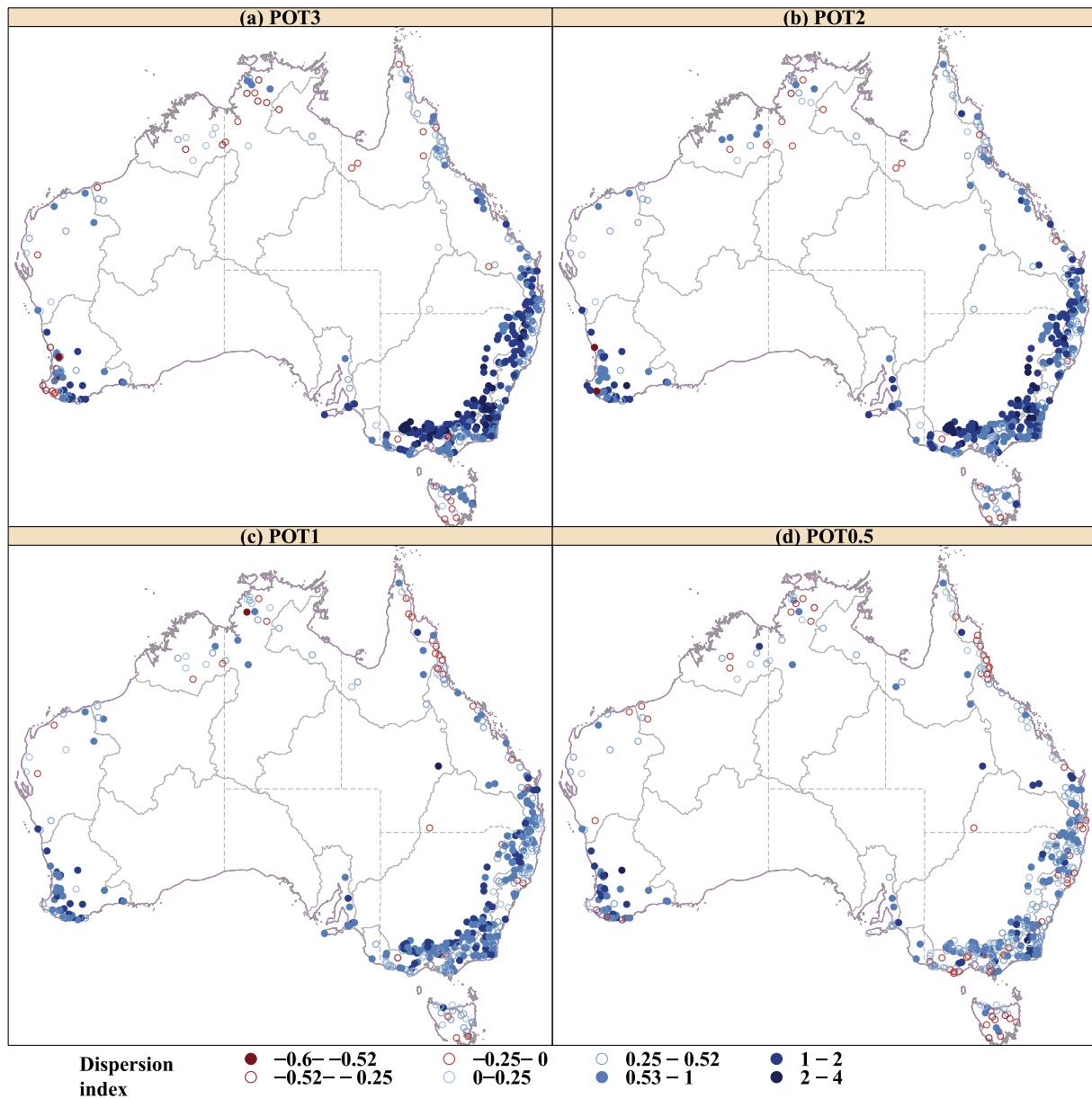


Fig. 7. Dispersion index of floods in one-year window size with different thresholds. Blue circles represent the clustering of floods, while red circles represent the under-dispersion of floods with the more regular inter-annual occurrence. Solid circles represent significant dispersion at 95% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

period. Taking the drainage division of SEN for an example, the peak flood occurrence rate in the flood-rich period is 2.6 times larger than the trough flood occurrence rate in the flood-poor period.

Fig. 9 also shows the linear trend of mean flood occurrence rates for each drainage division. The increasing trend in occurrence rates is observed across PG, NWP, TTS and LEB in northern Australia (Fig. 9a-d), with a high flood occurrence rates after 2000; in contrast, the decreasing trend is found in all the drainage divisions over southeastern Australia (Fig. 9g, h, i, k).

4.3. Atmospheric circulation associated with flood clustering

To understand the underlying causes of the flood clustering, we further analyse the spatiotemporal variation of atmospheric circulation at intra-annual and inter-annual scales and the entire series across Australia. This will be helpful to shed light on the physical

mechanisms behind the flood clustering. Based on the reanalysis data from National Centres for Environmental Prediction/National Centres of Atmospheric Research (NCEP/NCAR) (Kalnay et al., 1996), we extract atmospheric circulation data for the horizontal wind at 850 hPa and vertical wind velocity at 500 hPa during 1976–2010.

The results in Section 4.1 show that there exists temporal clustering for flood occurrences at intra-annual scale. February is the most flood-rich month over northern Australia but one of the flood-poor months over southern Australia. In contrast, August is the most flood-rich month over southern Australia, but one of the flood-poor months over northern Australia (Figs. 5 and 6). To address the underlying causes of intra-annual flood clustering, taking February and August as examples, we examine and compare the long-term mean of horizontal wind and vertical velocity during 1976–2010 (Fig. 10). In February (Fig. 10a), prominent northerly

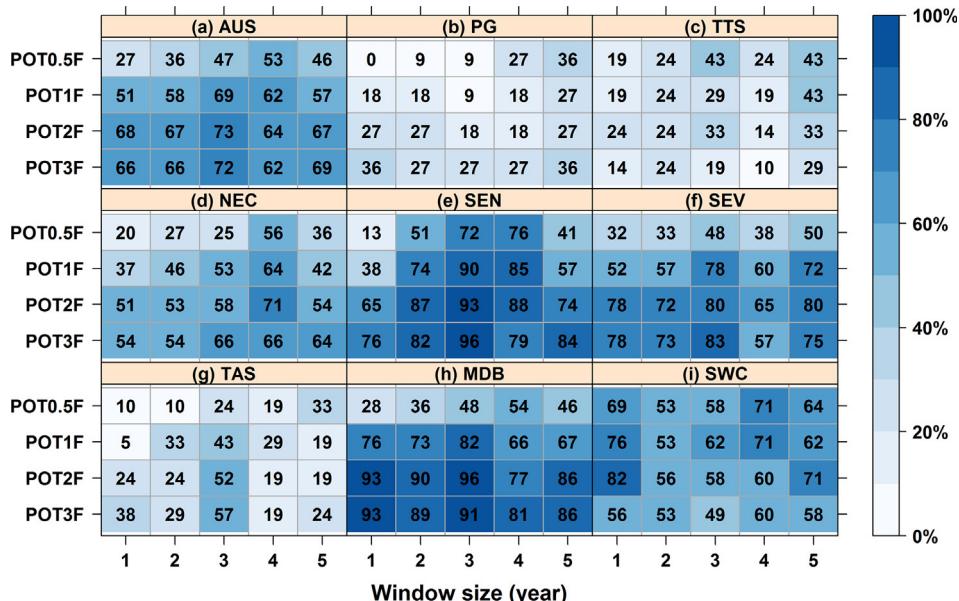


Fig. 8. Percentage of catchments in significant clustering of floods at 95% confidence level for different POT thresholds and inter-annual scales. These drainage divisions with more than 11 catchments are summarised, including 8 drainage divisions and the whole Australia (AUS). PG locates in western Australia; NWP, TTS, LEB and CC locate in northern Australia; NEC locates in northeastern Australia; SEN, SEV, TAS, MDB, SAG located in southeastern Australia; SWC locates in southwestern Australia. See Fig. 1 for abbreviations of the drainage divisions.

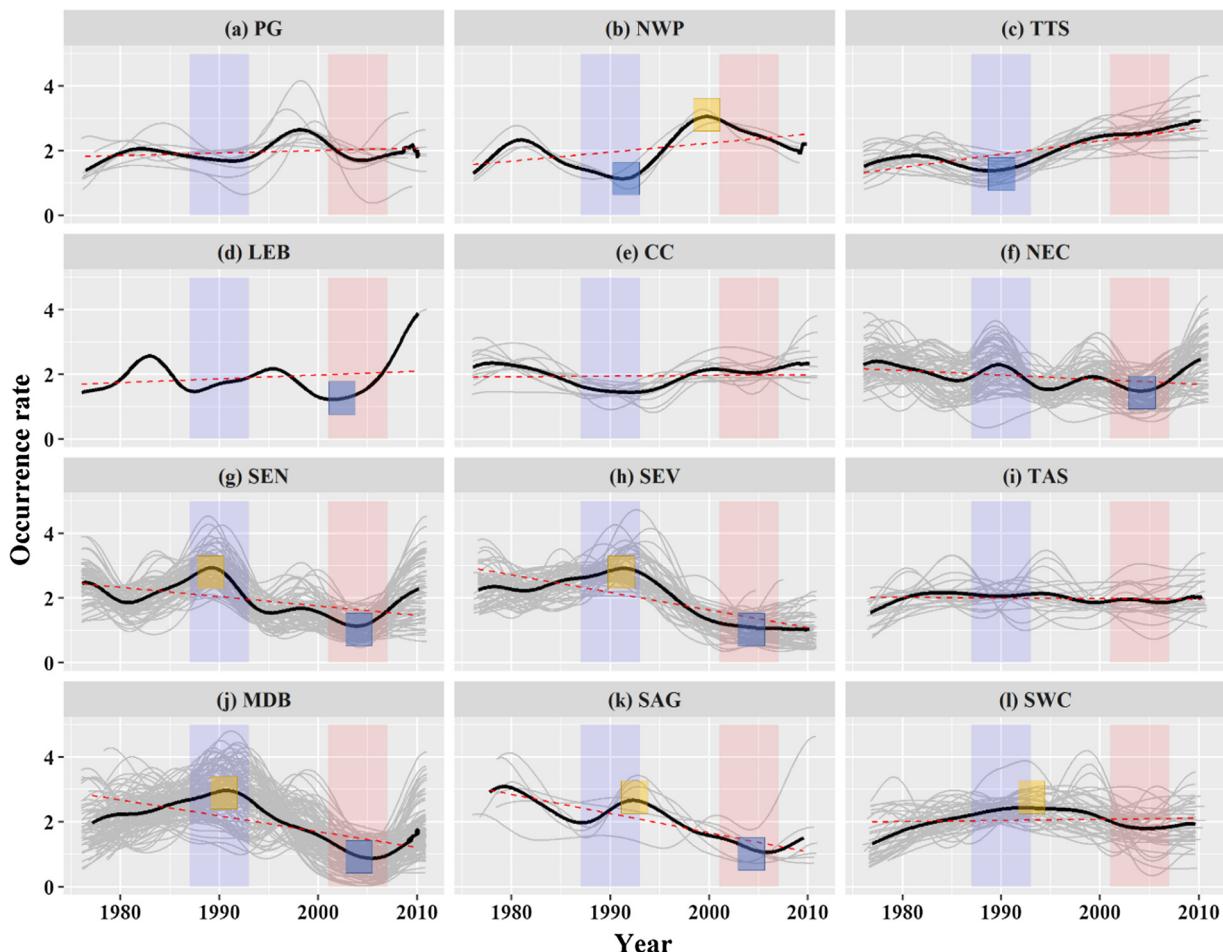


Fig. 9. Occurrence rates of floods in different drainage divisions. Gray lines show flood occurrence rates for different catchments; black lines represent the mean occurrence rates for each drainage division; red dashed lines show trends in regional mean. Peak values of occurrence rates with a significant (at 95% confidence level) flood-rich period are marked with yellow rectangle; trough values with a significant flood-poor period are marked with blue rectangle. PG locates in western Australia; NWP, TTS, LEB and CC locate in northern Australia; NEC locates in northeastern Australia; SEN, SEV, TAS, MDB, SAG located in southeastern Australia; SWC locates in southwestern Australia. See Fig. 1 for abbreviations of the drainage divisions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

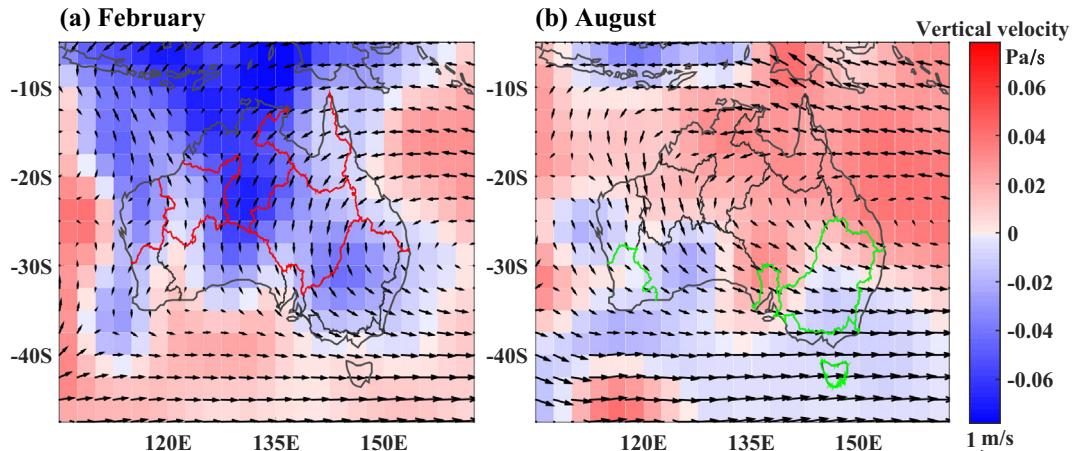


Fig. 10. Long-term mean of 850 hPa horizontal wind (vectors, unit: m/s) and 500 hPa vertical velocity (shading, unit: Pa/s) in (a) February and (b) August (b) in 1976–2010. The area in red shaded represents descending vertical wind velocity while the blue shaded area represents ascending vertical wind velocity. Red lines present the drainage division regarding February as the flood-rich month and green lines present the drainage division regarding August as the flood-rich month. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

winds control northern Australia; meanwhile, ascending vertical wind velocity, which indicates moisture convergence transferring from surface to upper tropospheric levels and brings more moisture condensation (Ummenhofer et al., 2015), can be found over northern Australia. As a result, more water vapour is transported and condensed over those regions, leading to more precipitation in February. In contrast, all catchments across southern Australia are dominated by northerly wind from inland; moreover, some areas in SWC, SAG and SEV over southern Australia are also controlled by descending vertical wind velocity. As a result, February becomes flood rich over northern Australia, but flood poor over southern Australia. In August (Fig. 10b), a subtropical ridge moves northward over northern Australia, where is controlled by trade winds instead of northerly winds, while southern Australia is dominated by prevailing westerlies. Meanwhile, northern Australia is characterised by descending motion and southern Australia is featured by ascending motion. As a result, there is less precipitation, resulting in flood poorer over northern Australia in August. In contrast, there is more precipitation, resulting in flood richer over southern Australia.

The results from Section 4.2 show that 1987–1992 is a significant (at 95% confidence level) flood-poor period for NWP and TTS in northern Australia (Fig. 9b–e), and a significant flood-rich period for all drainage divisions except TAS in southern Australia (Fig. 9g, h and j–l). 2001–2006 is a significant flood-poor period for LEB and NEC in northern Australia and all drainage division except TAS in southeastern Australia (Fig. 9d, f, g, h, j, h). Fig. 11 shows mean anomalies of horizontal wind and vertical wind velocity in February and August during 1987–1992 and 2001–2006 related to 1976–2010. In February, the major areas in NWP and TTS are featured by anomalies of descending motion and controlled by southerly wind from inland during 1987–1992 (Fig. 11a), leading to less precipitation and floods, which is roughly in line with the flood-poor period over these regions. During 2001–2006, the drainage divisions of LEB and NEC is controlled by the anomalies of southerly winds from arid inland (Fig. 11b), which makes these areas a significant flood-poor period. However, due to the anomalous cyclonic influence over west of Timor, more water vapour is transported to NWP and TTS, resulting in high flood occurrence rates in those regions (Fig. 9b–c). In August, there is an anomalous anticyclone over west of Australia and an anomalous cyclone over southwestern Australia during 1987–1992 (Fig. 11c), which is a north-south-oriented dipole pattern (Fig. 11c). Consequently, enhanced anomalies of northerly winds and of ascending motion control the SWC, resulting in the high flood occurrence rates during this

period. Nevertheless, southeastern Australia is characterised by the anomalies of easterly winds from the Pacific Ocean with lots of water vapour, resulting in increasing flood occurrence rates. During 2001–2006, an anomalous anticyclone can be found over southeast of Australia (Fig. 11d), most regions of which are dominated by anomalies of descending motion and distinct northerly wind from inland. Both lead to the decrease of flood occurrence rates. Overall, the flood temporal clustering shown in Section 3.2 is in line with the anomalies patterns of atmospheric variation.

To explain the general variation of flood occurrences rates, we further investigate the linear trends of horizontal wind and vertical wind velocity from 1976 to 2010 (Fig. 12). In February, there are two anomalous cyclones over western Timor Sea and southeastern Australia, respectively (Fig. 12a). As a result, west part of northern Australia is dominated by westerly winds from the Timor Sea with warm and moist water vapour; meanwhile, NWP is featured by significant ascending motion at 95% confidence level. Therefore, the increasing trend of flood occurrence rates is identified in February over these drainage divisions. In contrast, eastern coast of Australia over NEC and CC are characterised by westerly winds and dominated by significant (at 95% confidence level) descending motion, which leads to the decreasing trend in flood occurrence rates in February. In August (Fig. 12 b), south eastern Australia is controlled by a large-scale anticyclone, resulting in reducing precipitation and decreasing trend of flood occurrence rate across the drainage divisions in August.

5. Discussion

Lots of studies demonstrate that climate variation is significantly affected by climate indices. Ishak et al. (2013) indicated that Australian annual maximum flood is well related to climate indices including NSO, IPO and SAM. Johnson et al. (2016) stated that these climate indices also have a strong impact on magnitude and frequency of floods. Therefore, we first use the Cox regression model with these climate indices as covariates to examine the impacts of climate indices on flood occurrence time within the year across Australia. ENSO has a remarkable impact on the variation of precipitation in Australia (Cai et al., 2012). Pui et al. (2012) indicated that ENSO is the major impact index for precipitation change across eastern Australia. Risbey et al. (2009) found that ENSO plays a role in precipitation change over northern Australia. IOD and IPO mainly affect the variation of floods and extreme precipitation in

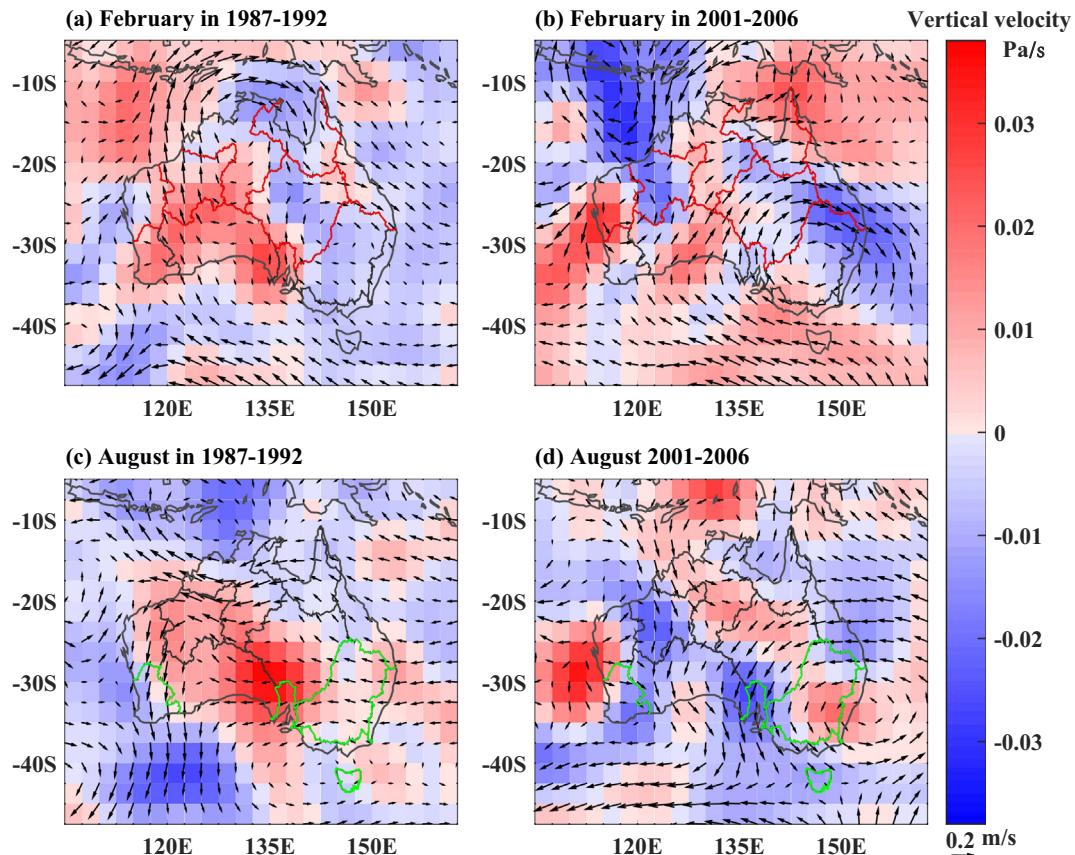


Fig. 11. Mean anomalies of 850 hPa horizontal wind (vectors, unit: m/s) and 500 hPa vertical velocity (shading, unit: Pa/s) in February (a–b), and August (c–d) by 1987–1992 minus 1976–2010 (a, c) and 2001–2006 minus 1976–2010 (b, d), respectively. The area in red shaded represents descending vertical wind velocity while the blue shaded represents ascending vertical wind velocity. Red lines present the drainage division regarding February as the flood-rich month and green lines present the drainage division regarding August as the flood-rich month. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

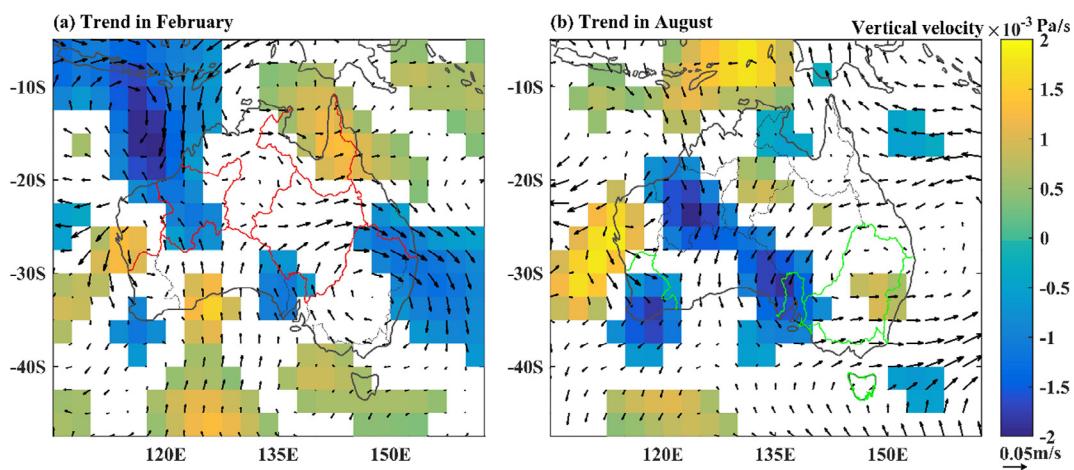


Fig. 12. Linear trends of 850 hPa horizontal wind (vectors, unit: m/s per year) and 500 hPa vertical velocity (shading, unit: Pa/s per year) from 1976 to 2010 in February (a) and August (b). The area in yellow shaded represents the significant increasing trend of ascending vertical wind velocity while the blue shaded represents the significant increasing trend of descending vertical wind velocity. Red lines present the drainage division regarding February as the flood-rich month and green lines present the drainage division regarding August as the flood-rich month. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

eastern Australia (Pui et al., 2011; Ishak et al., 2013; Pepler et al., 2014). SAM is adjudged as a dominant index of climate change in extratropical Southern Hemisphere (Thompson and Wallace, 2000). Hendon et al. (2007) found that SAM explains up to about 15% of rainfall variance over southwest and southeast of Australia. The previous studies have indicated the noticeable impacts of

climate on the variation of precipitation and floods in magnitude and frequency across Australia, with affecting area for each index similar to our findings (Hendon et al., 2007; Risbey et al., 2009; Cai et al., 2012; Pui et al., 2012; Ishak et al., 2013; Johnson et al., 2016). Nevertheless, our results further demonstrate that the climate indices also play an important role in controlling the

occurrence time of Australian floods. However, we did not explore how the climate indices influence flood occurrence time and why their impacts vary with the impact area and flood severities, since these would require analysing the influencing mechanism of climate indices on floods, which is beyond the scope of this study.

Using the monthly frequency method, we identify flood-rich months occurring from January to March over northern Australia, and from July to September over southern Australia. The precipitation in Australia is winter dominant in south, but summer dominant in north. As a result, the flood-rich/flood-poor months show a distinct north-south difference. It is noted that the similar monthly frequencies can be observed in SEC, which is dominant by the temperate climate with uniform rainfall and no dry season (Fig. 1). Even so, February and March are two slightly flood-rich months in SEC, which may be caused by large-scale precipitation from the cyclonic activity (Lavender and Abbs, 2013). Strong flood seasonality is also found in Europe, Asian and North America (Table 1). For instance, most floods over western and eastern United States occur between October to March (Villarini, 2016). For Europe, however, different regions exhibit different flood-rich months, reflecting the differences in flood generating mechanisms.

As for inter-annual scale, the dispersion index and kernel occurrence estimation are used to assess whether there exists the flood clustering. Flood-rich/flood-poor periods found in Australia is in line with the findings of Merz et al. (2016) in Germany (Table 1). 2001–2006 is a distinct flood-poor period over southeastern Australia where precipitation is very low from 1997 to 2009, and streamflow in the River Murray during 2001–2005 was the lowest on record, with 40% of annual average (Murphy and Timbal, 2008; Yang et al., 2017). Lavender and Abbs (2013) found that Australian precipitation has experienced significant changes over 1970–2009 with decreases in east and increases in northwest. Yilmaz and Perera (2015) indicated that there is a significant decreasing trend in extreme rainfall in long storm duration (i.e., 6, 12, 24 and 48 h) over southeastern Australia. These results are in accordance with our findings that flood occurrence rates decrease in southeastern Australia and increase in northwestern Australia.

Catchment characteristics can impact streamflow generation and play a critical role in controlling the lags in hydrological recovery (Yang et al., 2017). The catchment antecedent condition has a prominent impact on the frequency distribution of flood (De Michele and Salvadori, 2002). Persistent wet or dry conditions for several months or years may also play a role in the formation of the flood-rich or flood-poor period (Merz et al., 2016). Merz and Plate (1997) indicated that the catchment characteristics have larger impacts on smaller floods than severe ones. For the catchment in wet conditions, a moderate precipitation may generate a smaller flood event. Meanwhile, the moderate precipitation has a larger occurrence probability than the severity precipitation. As a result, the lags in the hydrological recovery play a more important role in flood clustering for the smaller floods with short time interval than the larger ones with long time interval. With the decrease of threshold, the POT sampling captures more flood events with shorter temporal interval and smaller magnitude. Therefore, the lower threshold tends to show higher clustering related to the stronger impacts by the catchments memory. This is consistent with the result of the strength of clustering, as well as the number of catchments showing significant clustering, decreasing with the increasing flood severity. A similar relationship between flood clustering and flood threshold is also found in Germany by Merz et al. (2016). However, we didn't examine whether the clustering exists in more severe floods (such as one flood for each 5 or 10 years) and multidecadal time scales, since there are only 36 years of observed streamflow data available.

The unregulated catchments selected in this study are generally located in headwaters with less intensive land use and without

subject to reservoir regulation. Therefore, the clustering of floods only reflects the impact from climate variability, which is in well-organised modes of inter-annual, inter-decadal and lower frequency (Barnston and Livezey, 1987) and has great effects on flood occurrence through changing atmospheric circulation (Baker et al., 1988; Gu et al., 2017). Merz et al. (2016) assumed that the clustering of floods across Germany is caused by catchments memory effects. However, the effects of memory cannot explain the flood clustering at specific times. Therefore, we first examined the relations between the intra-annual time of flood occurrence and the climate indices, then investigated atmospheric circulation associated with the periods of flood clustering. We show that the flood occurrence is not independent and exhibits clustering modulated by climatic indices. More importantly, the clustering periods of floods are in line with atmospheric circulation both at intra-annual and inter-annual scales. In summary, the variation of atmospheric circulation related to climate variability is the main reason for flood clustering, meanwhile, the catchment memory also plays a role in the strength of clustering.

6. Conclusions

Though several studies have been carried out in the Northern Hemisphere, there are few reports in literatures on temporal clustering of floods for large regions in Southern Hemisphere where the causative factors behind the flood clustering remain poorly understood. This study, for the first time, assesses the flood clustering and its causes at intra-annual and inter-annual scales across 413 Australian catchments for the last 36 years (i.e., 1975–2010). The following conclusions can be drawn:

- (1) Cox regression model is used as a framework to investigate the relation between intra-annual time of flood occurrence and four climate indices representing the influence of the Pacific Ocean and India Ocean. The variation of flood occurrence time within the year significantly relates to the climate indices, showing the occurrence of floods is not independent and exhibits clustering modulated by climatic indices.
- (2) There exist significant flood-rich and flood-poor months across Australia, with distinct north-south differences. In northern Australia, January to March are flood-rich months, and longer flood-poor months are from June to October. In contrast, in southern Australia, July to September are flood rich, and January to May are flood poor.
- (3) The strength of inter-annual flood clustering varies across Australia. Southern Australia exhibits stronger clustering of floods than northern Australia. The strength of clustering decreases with the increase of flood severities.
- (4) Significant flood-rich/flood-poor periods are detected across Australia. 1978–1997 is a significant flood-poor period for most drainage divisions across the northern Australia, but a significant flood-rich period for all drainage divisions except for TAS in southern Australia. In contrast, 2001–2006 is a significant flood-poor period for LEB and NEC in northern Australia and all drainage divisions except for TAS in southeastern Australia.
- (5) Coherence between the variations of atmospheric circulation and flood clustering is verified at intra-annual and inter-annual scales. At intra-annual scale, northerly winds and ascending motion lead to active floods across northern Australia in February; prevailing westerlies and ascending motion result in frequent floods across southwest and southeast of Australia in August. During 1987–1992, anomalies in descending motion in February bring scarce flood over northern Australia; enhanced anomalies easterly winds in

August from the Pacific Ocean with moist water vapour generate rich flood over southeastern Australia. During 2001–2006, anomalies in southerly winds from inland in February control most regions across northern Australia, leading to scarce flood over these regions; the large-scale anomalous anticyclone in August over southeastern Australia leads to the scarce flood in flood season. For the period of 1976–2010, the significant trend of ascending motion and westerly winds results in increasing flood occurrence rates over northwest of Australia; anomalous anticyclone over southeastern Australia brings decreasing flood occurrence rates across the drainage divisions in this region.

Using daily streamflow data obtained from 413 unregulated catchments across Australian Continent, this study comprehensively investigates the multi-temporal clustering of continental floods at different thresholds. It sheds new light on physical mechanisms behind the flood clustering by analysing the spatiotemporal variation of atmospheric circulation.

Acknowledgements

This work was carried out in CSIRO Land and Water, with support from International Program for Ph.D. Candidates, Sun Yat-Sen University. We thank hard editorial work from the Editor Tim R. McVicar and the Associate Editor Sergio M. Vicente-Serrano, and constructive comments and suggestions from Manuela Nied, Jose Luis Salinas and an anonymous reviewer. We thank Ruth Palmer for proofreading this paper. We also appreciate Qiang Zhang, Xihui Gu and Mingzhong Xiao for their kind assistance.

References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19 (6), 716–723.
- Baker, V., Kochel, R.C., Patton, P.C., 1988. *Flood geomorphology*. Wiley-Interscience.
- Barnston, A.G., Livezey, R.E., 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Weather Rev.* 115 (6), 1083–1126.
- Boschat, G., Pezza, A., Simmonds, I., Perkins, S., Cowan, T., Purich, A., 2015. Large scale and sub-regional connections in the lead up to summer heat wave and extreme rainfall events in eastern Australia. *Clim. Dyn.* 44 (7–8), 1823–1840.
- Cai, W., Van Rensh, P., Cowan, T., Hendon, H.H., 2012. An asymmetry in the IOD and ENSO teleconnection pathway and its impact on Australian climate. *J. Clim.* 25 (18), 6318–6329.
- Cunderlik, J.M., Ouarda, T.B., Bobée, B., 2004. Determination of flood seasonality from hydrological records/Détermination de la saisonnalité des crues à partir de séries hydrologiques. *Hydrol. Sci. J.* 49 (3).
- De Michele, C., Salvadori, G., 2002. On the derived flood frequency distribution: analytical formulation and the influence of antecedent soil moisture condition. *J. Hydrol.* 262 (1), 245–258.
- Diggle, P., 1985. A kernel method for smoothing point process data. *Appl. Stat.*, 138–147.
- Fasullo, J.T., Boening, C., Landerer, F.W., Nerem, R.S., 2013. Australia's unique influence on global sea level in 2010–2011. *Geophys. Res. Lett.* 40 (16), 4368–4373.
- Greene, W.H., 2003. *Econometric Analysis*. Pearson Education India.
- Gu, X., Zhang, Q., Singh, V.P., Chen, X., Liu, L., 2016a. Nonstationarity in the occurrence rate of floods in the Tarim River basin, China, and related impacts of climate indices. *Global Planet. Change* 142, 1–13.
- Gu, X., Zhang, Q., Singh, V.P., Chen, Y.D., Shi, P., 2016b. Temporal clustering of floods and impacts of climate indices in the Tarim River basin, China. *Global Planet. Change* 147, 12–24. <https://doi.org/10.1016/j.gloplacha.2016.10.011>.
- Gu, X., Zhang, Q., Singh, V.P., Shi, P., 2017. Changes in magnitude and frequency of heavy precipitation across China and its potential links to summer temperature. *J. Hydrol.* 547, 718–731.
- Hall, J. et al., 2014. Understanding flood regime changes in Europe: a state of the art assessment. *Hydrol. Earth Syst. Sci.* 18 (7), 2735–2772.
- Hendon, H.H., Thompson, D.W., Wheeler, M.C., 2007. Australian rainfall and surface temperature variations associated with the Southern Hemisphere annular mode. *J. Clim.* 20 (11), 2452–2467.
- Hirschboeck, K.K., 1988. *Flood Hydroclimatology*. In: Baker, V.R., Kochel, R.C., Patton, P.C. (Eds.), *Flood Geomorphology*. Wiley, NY.
- Ishak, E., Rahman, A., Westra, S., Sharma, A., Kuczera, G., 2013. Evaluating the non-stationarity of Australian annual maximum flood. *J. Hydrol.* 494, 134–145.
- Jain, S., Lall, U., 2001. Floods in a changing climate: does the past represent the future? *Water Resour. Res.* 37 (12), 3193–3205.
- Johnson, F., White, C.J., van Dijk, A., Ekstrom, M., Evans, J.P., Jakob, D., Kiern, A.S., Leonard, M., Rouillard, A., Westra, S., 2016. Natural hazards in Australia: floods. *Clim. Change* 139 (1), 21–35.
- Kalnay, E. et al., 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77 (3), 437–471.
- Koutsoyiannis, D., 2005. Uncertainty, entropy, scaling and hydrological stochastics. 1. Marginal distributional properties of hydrological processes and state scaling/Incertitude, entropie, effet d'échelle et propriétés stochastiques hydrologiques. 1. Propriétés distributionnelles marginales des processus hydrologiques et échelle d'état. *Hydrol. Sci. J.* 50 (3).
- Lang, M., Ouarda, T., Bobée, B., 1999. Towards operational guidelines for over-threshold modeling. *J. Hydrol.* 225 (3), 103–117.
- Lavender, S.L., Abbs, D.J., 2013. Trends in Australian rainfall: contribution of tropical cyclones and closed lows. *Clim. Dyn.* 40 (1–2), 317–326.
- Li, H.X., Zhang, Y.Q., 2017. Regionalising rainfall-runoff modelling for predicting daily runoff: comparing gridded spatial proximity and gridded integrated similarity approaches against their lumped counterparts. *J. Hydrol.* 550, 279–293. <https://doi.org/10.1016/j.jhydrol.2017.05.015>.
- Liu, J.Y., Zhang, Q., Singh, V.P., Gu, X.H., Shi, P.J., 2017. Nonstationarity and clustering of flood characteristics and relations with the climate indices in the Poyang Lake basin, China. *Hydrol. Sci. J.* <https://doi.org/10.1080/02626667.2017.1349909>.
- Macdonald, N., Phillips, I.D., Mayle, C., 2010. Spatial and temporal variability of flood seasonality in Wales. *Hydrol. Process.* 24 (13), 1806–1820.
- Mallakpour, I., Villarini, G., 2016. Investigating the relationship between the frequency of flooding over the central United States and large-scale climate. *Adv. Water Resour.* 92, 159–171. <https://doi.org/10.1016/j.advwatres.2016.04.008>.
- Marshall, G.J., 2003. Trends in the Southern Annular Mode from observations and reanalyses. *J. Clim.* 16 (24), 4134–4143.
- Mediero, L. et al., 2015. Identification of coherent flood regions across Europe by using the longest streamflow records. *J. Hydrol.* 528, 341–360. <https://doi.org/10.1016/j.jhydrol.2015.06.016>.
- Mediero, L., Santillán, D., Garrote, L., Granados, A., 2014. Detection and attribution of trends in magnitude, frequency and timing of floods in Spain. *J. Hydrol.* 517, 1072–1088.
- Merz, B. et al., 2014. Floods and climate: emerging perspectives for flood risk assessment and management. *Nat. Hazards Earth Syst. Sci.* 14 (7), 1921.
- Merz, B., Nguyen, V.D., Vorogushyn, S., 2016. Temporal clustering of floods in Germany: do flood-rich and flood-poor periods exist? *J. Hydrol.* 541, 824–838. <https://doi.org/10.1016/j.jhydrol.2016.07.041>.
- Merz, B., Plate, E., 1997. An analysis of the effects of spatial variability of soil and soil moisture on runoff. *Water Resour. Res.* 33 (12), 2909–2922.
- Min, S.K., Cai, W., Whetton, P., 2013. Influence of climate variability on seasonal extremes over Australia. *J. Geophys. Res. Atmos.* 118 (2), 643–654.
- Mudelsee, M., 2010. Climate time series analysis: classical statistical and bootstrap methods. *Atmospheric and Oceanographic Sciences Library*, vol. 42. Springer, p. 474. Price.
- Mudelsee, M., Börngen, M., Tetzlaff, G., Grünewald, U., 2004. Extreme floods in central Europe over the past 500 years: Role of cyclone pathway "Zugstrasse Vb". *J. Geophys. Res. Atmos.* 109 (D23). <https://doi.org/10.1029/2004JD005034>.
- Mumby, P.J., Vitolo, R., Stephenson, D.B., 2011. Temporal clustering of tropical cyclones and its ecosystem impacts. *Proc. Natl. Acad. Sci.* 108 (43), 17626–17630.
- Murphy, B.F., Timbal, B., 2008. A review of recent climate variability and climate change in southeastern Australia. *Int. J. Climatol.* 28 (7), 859–879.
- Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., Anquetin, S., 2008. Flash flood warning based on rainfall thresholds and soil moisture conditions: an assessment for gauged and ungauged basins. *J. Hydrol.* 362 (3), 274–290.
- Pepler, A., Timbal, B., Rakich, C., Coutts-Smith, A., 2014. Indian Ocean Dipole overrides ENSO's influence on cool season rainfall across the Eastern Seaboard of Australia. *J. Clim.* 27 (10), 3816–3826.
- Pui, A., Lal, A., Sharma, A., 2011. How does the Interdecadal Pacific Oscillation affect design floods in Australia? *Water Resour. Res.* 47 (5).
- Pui, A., Sharma, A., Santoso, A., Westra, S., 2012. Impact of the El Niño–Southern oscillation, Indian Ocean dipole, and Southern annular mode on daily to subdaily rainfall characteristics in East Australia. *Mon. Weather Rev.* 140 (5), 1665–1682.
- Pinto, J.G., Bellenbaum, N., Karremann, M.K., Della-Marta, P.M., 2013. Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions. *J. Geophys. Res. Atmos.* 118 (22), 12476–12485. <https://doi.org/10.1002/2013jd020564>.
- Risbey, J.S., Pook, M.J., McIntosh, P.C., Wheeler, M.C., Hendon, H.H., 2009. On the remote drivers of rainfall variability in Australia. *Mon. Weather Rev.* 137 (10), 3233–3253.
- Silva, A.T., Portela, M.M., Naghettini, M., 2012. Nonstationarities in the occurrence rates of flood events in Portuguese watersheds. *Hydrol. Earth Syst. Sci.* 16 (1), 241–254. <https://doi.org/10.5194/hess-16-241-2012>.
- Smith, J.A., Karr, A.F., 1986. Flood frequency analysis using the Cox regression model. *Water Resour. Res.* 22 (6), 890–896.
- Stedinger, J.R., 1993. Frequency analysis of extreme events. In: *Handbook of Hydrology*, 18.
- Therneau, T., Lumley, T., 2011. *Survival: Survival analysis, including penalised likelihood*. R package version 2.36-5. *Survival: Survival analysis, including penalised likelihood*. R package version: 2.36-2.2010.

- Thomas, L., Reyes, E.M., 2014. Tutorial: survival estimation for Cox regression models with time-varying coefficients using SAS and R. *J. Stat. Software* 61 (1), 1–23.
- Thompson, D.W., Wallace, J.M., 2000. Annular modes in the extratropical circulation. Part I: month-to-month variability. *J. Clim.* 13 (5), 1000–1016.
- Ummenhofer, C.C. et al., 2015. How did ocean warming affect Australian rainfall extremes during the 2010/2011 La Niña event? *Geophys. Res. Lett.* 42 (22), 9942–9951.
- Villarini, G., 2016. On the seasonality of flooding across the continental United States. *Adv. Water Resour.* 87, 80–91. <https://doi.org/10.1016/j.advwatres.2015.11.009>.
- Villarini, G. et al., 2011. On the frequency of heavy rainfall for the Midwest of the United States. *J. Hydrol.* 400 (1–2), 103–120. <https://doi.org/10.1016/j.jhydrol.2011.01.027>.
- Villarini, G., Smith, J.A., Vecchi, G.A., 2013a. Changing frequency of heavy rainfall over the central United States. *J. Clim.* 26 (1), 351–357.
- Villarini, G., Smith, J.A., Vitolo, R., Stephenson, D.B., 2013b. On the temporal clustering of US floods and its relationship to climate teleconnection patterns. *Int. J. Climatol.* 33 (3), 629–640. <https://doi.org/10.1002/joc.3458>.
- Vitolo, R., Stephenson, D.B., Cook, I.M., Mitchell-Wallace, K., 2009. Serial clustering of intense European storms. *Meteorol. Z.* 18 (4), 411–424.
- Wadey, M., Haigh, I., Brown, J., 2014. A century of sea level data and the UK's 2013/14 storm surges: an assessment of extremes and clustering using the Newlyn tide gauge record. *Ocean Sci.* 10 (6), 1031.
- Wolff, N.H. et al., 2016. Temporal clustering of tropical cyclones on the Great Barrier Reef and its ecological importance. *Coral Reefs* 35 (2), 613–623.
- Xiao, M., Zhang, Q., Singh, V.P., Chen, X., 2013. Regionalization-based spatiotemporal variations of precipitation regimes across China. *Theor. Appl. Climatol.* 114 (1–2), 203–212.
- Yang, Y., McVicar, T.R., Donohue, R.J., Zhang, Y., Roderick, M.L., Chiew, F.H., Zhang, J., 2017. Lags in hydrologic recovery following an extreme drought: assessing the roles of climate and catchment characteristics. *Water Resour. Res.* 53, 4821–4837. <https://doi.org/10.1002/2017WR020683>.
- Yilmaz, A.G., Imteaz, M.A., Perera, B.J.C., 2017. Investigation of non-stationarity of extreme rainfalls and spatial variability of rainfall intensity-frequency-duration relationships: a case study of Victoria, Australia. *Int. J. Climatol.* 37 (1), 430–442.
- Yilmaz, A.G., Perera, B.J.C., 2015. Spatiotemporal trend analysis of extreme rainfall events in Victoria, Australia. *Water Resour. Manage.* 29 (12), 4465–4480.
- Zhao, M., Running, S.W., 2010. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* 329 (5994), 940–943.
- Zhang, Y., Chiew, F.H., 2009. Relative merits of different methods for runoff predictions in ungauged catchments. *Water Resour. Res.* 45, W07412. <https://doi.org/10.1029/2008WR007504>.
- Zhang, Y., Viney, N., Frost, A., Oke, A., Brooks, M., Chen, Y., Campbell, N., 2013. Collation of Australian modeller's streamflow dataset for 780 unregulated Australian catchments. *Water for a Healthy Country National Research Flagship*, 115pp.
- Zhang, Y., Zheng, H., Chiew, F.H.S., Arancibia, J.P., Zhou, X., 2016. Evaluating regional and global hydrological models against streamflow and evapotranspiration measurements. *J. Hydrometerol.* 17 (3), 995–1010.
- Zhou, Y., Zhang, Y., Vaze, J., Lane, P., Xu, S., 2013. Improving runoff estimates using remote sensing vegetation data for bushfire impacted catchments. *Agric. For. Meteorol.* 182, 332–341.