



1 **Using paleoclimate reconstructions to analyse**
2 **hydrological epochs associated with Pacific Decadal**
3 **Variability**

4 LANYING ZHANG, GEORGE KUCZERA

5 *School of Engineering, University of Newcastle, Callaghan, New South Wales, Australia*

6 ANTHONY S. KIEM

7 *Centre for Water, Climate and Land (CWCL), Faculty of Science, University of Newcastle, Callaghan, New*
8 *South Wales, Australia*

9 GARRY WILLGOOSE

10 *School of Engineering, University of Newcastle, Callaghan, New South Wales, Australia*

11 **Abstract**

12 The duration of dry or wet hydrological epochs (run lengths) associated with positive or
13 negative Inter-decadal Pacific Oscillation (IPO) or Pacific Decadal Oscillation (PDO) phases, termed
14 Pacific Decadal Variability (PDV), is an essential statistical property for understanding, assessing and
15 managing hydroclimatic risk. Numerous IPO and PDO paleoclimate reconstructions provide a
16 valuable opportunity to study the statistical signatures of PDV, including run lengths. However,
17 disparities exist between these reconstructions making it problematic to determine which
18 reconstruction(s) to use to investigate pre-instrumental PDV and run length. Variability and
19 persistence on centennial scales are also present in some millennium long reconstructions, making
20 consistent run length extraction difficult. Thus, a robust method to extract meaningful and consistent



1 run lengths from multiple reconstructions is required. In this study, a dynamic threshold framework to
2 account for centennial trends in PDV reconstructions is proposed. The dynamic threshold framework
3 is shown to extract meaningful run length information from multiple reconstructions. Two
4 hydrologically important aspects of the statistical signatures associated with the PDV are explored: (i)
5 whether persistence (i.e. run lengths) during positive epochs is different to persistence during negative
6 epochs and (ii) whether the reconstructed run lengths are stationary during the past millennium.
7 Results suggest that there is no significant difference between run lengths in positive and negative
8 phases of PDV and that it is more likely than not that the PDV run length has been non-stationary in
9 the past millennium. This raises concerns about whether variability seen in the instrumental record
10 (the last ~100 years), or even in the shorter 300–400 year paleoclimate reconstructions, is
11 representative of the full range of variability.

12 Key words: PDV; run length; multiple reconstructions; stationarity

13 1. Introduction

14 A pattern of Pacific ocean-atmosphere climate variability at decadal time scales has been
15 identified and explored at various locations around the world including Africa (Reason and Rouault,
16 2002; Hoell and Funk, 2014), Eastern and Southern Asia (Krishnan and Sugi, 2003; Ma, 2007),
17 America (Mantua and Hare, 2002; Andreoli and Kayano, 2005) and Australia (Kiem et al.,
18 2003; Verdon et al., 2004; Verdon and Franks, 2006; Henley et al., 2011). This pattern, referred to in
19 this paper as Pacific Decadal Variability (PDV), is associated with sea surface temperature
20 fluctuations and sea level pressure changes in the north and south Pacific Ocean.

21 PDV is usually described by the Inter-decadal Pacific Oscillation (IPO) or Pacific Decadal
22 Oscillation (PDO) indices (Mantua et al., 1997; Power et al., 1999). Of particular hydrological
23 relevance are the statistical characteristics of PDV phases where a phase refers to a period during
24 which the PDV index lies above (or below) some thresholds (Dong and Dai, 2015; Verdon et al.,
25 2004; Henley et al., 2011; Vance et al., 2015). The duration of a PDV phase (termed run length) is



1 defined as the time between consecutive crossings of the threshold. The phase may be described as
2 positive (negative) if the PDV index is above (below) the threshold, or dry (wet) if the PDV phases
3 are associated with predominantly dry (wet) hydrological conditions.

4 Of particular relevance to the hydrology and water resources community is the evidence that
5 the PDV phases can be associated with multi-decadal periods of persistently wetter or drier conditions
6 and corresponding increases in flood or drought risk in affected regions, particularly those on the
7 Pacific rim. PDV has been found to be related to a number of hydrological variables including
8 precipitation, streamflow, flood/drought risk (Cook et al., 2013;Kiem et al., 2003;Verdon et al.,
9 2004;Dai, 2013;Goodrich and Walker, 2011;Hu and Huang, 2009;Li et al., 2012;McCabe et al.,
10 2012;Mehta et al., 2011;Wang et al., 2014;Henley et al., 2013). For example, Kiem and Franks (2004)
11 showed in a case study for eastern Australia that the probability of reservoir storages falling below a
12 critical level differs significantly depending on the PDV phases (see Figure 6 in Kiem and Franks
13 (2004)). Kiem and Franks (2003) and Micevski et al. (2006) demonstrated that flood risk in eastern
14 Australia is strongly dependent on PDV phase. Henley et al. (2013) extended this work to show that
15 short-term drought risk is strongly dependent not only on the PDV phase but also on the time spent in
16 a particular phase.

17 Despite the clear relevance of PDV phases to hydrological risk assessment and water resource
18 management, there remain considerable knowledge gaps about the statistical characteristics of PDV,
19 including run lengths. The instrumental record shows that PDV phases have varied irregularly with
20 runs ranging from less than a decade to several decades during the past century. However, the
21 instrumental record is insufficient to characterize the statistical characteristics of PDV run lengths. In
22 response to this, significant advances have been made in reconstructing pre-instrumental PDV
23 behaviour to extend data length. For example, at least 12 IPO or PDO reconstructions have been
24 published (Biondi et al., 2001;Gedalof and Smith, 2001;D'Arrigo et al., 2001;MacDonald and Case,
25 2005;D'Arrigo and Wilson, 2006;Shen et al., 2006;Verdon and Franks, 2006;Linsley et al.,
26 2008;Mann et al., 2009;McGregor et al., 2010;Henley et al., 2011;Vance et al., 2015).



1 In view of these published PDV reconstructions, the fundamental question is whether useful
2 information can be extracted about the statistical characteristics of PDV persistence. In dissecting this
3 question, several unresolved issues are identified:

4 1) Static threshold methods have been used to estimate run lengths of PDV phases. This
5 raises the concern that biased conclusions about the statistical characteristics of decadal climate
6 variability may be drawn when reconstructions exhibit centennial or longer trends. The non-
7 parametric Mann-Whitney test method was used in Verdon and Franks (2006) whereby a crossing
8 was defined when the test detected a statistically significant difference between two halves of data in
9 a 30-year moving window, with zero used as a static threshold to define the sign of PDV phases
10 (Verdon-Kidd, personal communication, 24 April, 2017). Henley et al. (2011) used the instrumental
11 mean of the composite IPO/PDO index, while Vance et al. (2015) used ± 0.5 as thresholds to
12 determine crossings and the sign of PDV phases, and Henley et al. (2017) used the long-term
13 modelled IPO Tripole Index (TPI) mean as the static threshold with run lengths less than 5 years
14 omitted. However, when static threshold methods are used, extraordinary long centennial run lengths
15 may be identified in the reconstructions that exhibit centennial or longer trends (MacDonald and Case,
16 2005; Mann et al., 2009). Analysis based on data with long-term trends can lead to results that are
17 overwhelmed by such trends and hence efforts have been made to detrend data without losing useful
18 signals (Wahl et al., 2006; von Storch et al., 2006; Wu et al., 2007). Given that static threshold methods
19 cannot remove trends, how should the run length extraction method be designed to extract useful and
20 consistent information from all reconstructions, including reconstructions exhibiting centennial trends?
21 To what extent are run length distributions sensitive to the choice of extraction parameters (e.g.
22 threshold, window width etc.)?

23 2) The durations of positive and negative PDV phases have not been treated separately
24 in most previous research (e.g. Verdon and Franks, 2006; Linsley et al., 2008; Henley et al., 2011a)
25 Using a high-resolution millennial-length IPO reconstruction, Vance et al. (2015) observed that
26 positive phases dominate and last longer than negative phases. Owing to nonlinear hydrologic
27 feedback mechanisms, such as elasticity of streamflow to runoff (Chiew, 2006) and shifted rainfall-



1 runoff relationships (Saft et al., 2015), the duration of increased (decreased) precipitation during a wet
2 (or dry) PDV phase may impact on streamflow in a highly nonlinear manner. Therefore, assuming wet
3 and dry runs follow the same distribution may misrepresent the intensity of impacts. Is there sufficient
4 evidence about the dissymmetry between positive and negative phase run lengths from the multiple
5 reconstructions? Will run length samples from each phase lead to significantly different run length
6 simulations?

7 3) Stationarity of climate signatures is a necessary, yet implicit, assumption underlying
8 many analyses based on paleoclimate reconstructions – and hydroclimate stochastic modelling used in
9 water resource assessment (e.g. Henley et al., 2013). However, a number of studies have suggested
10 the possibility of climatic non-stationarity over the past millennium. Phipps et al. (2013) found that
11 the relationships between paleoclimate proxies and climatic variables have been changing at least the
12 last millennium. Razavi et al. (2015) explored the stationarity of the mean, variance and
13 autocorrelation structures of paleo tree-ring proxy data and concluded that the key statistical
14 characteristics of climate had undergone significant change over time. The existence of extraordinary
15 warm and cold periods, known as Medieval Climate Anomaly (~950-1250CE) and Little Ice Age
16 (~1400-1700CE) (Mann et al., 2009;Phipps et al., 2013;Atwood et al., 2015), also challenges the
17 assumption of climatic stationarity. Furthermore, externally forced anthropogenic climate change,
18 arising from elevated atmospheric concentrations of greenhouse gases, may also make recent and
19 future climate different from the long past (Kirtman et al., 2013). Therefore, care should be taken
20 when positing that PDV characteristics are stationary. Statistical signatures of past reconstructed PDV
21 runs should be tested for non-stationarity (and where possible non-stationarity should be attributed to
22 causal mechanisms) to ensure robust and representative estimation of the full range of variability.

23 In view of the above mentioned issues, this study has the following objectives:

24 1) To develop a more robust run length extraction method that is applicable to all
25 reconstructions, including those with centennial trends. Of particular importance is the sensitivity of
26 the inferred run length distributions to the choice of the extraction method parameters.



1 2) To identify hydrologically important statistical properties of PDV that are common to
2 different reconstructions. In particular two fundamental questions are explored: (i) whether
3 persistence (i.e. run lengths) during positive epochs is different to persistence during negative epochs
4 and (ii) whether the reconstructed run lengths are stationary during the past millennium. Both
5 questions address fundamental concerns about using PDV and run length information in stochastic
6 hydroclimatic models used to assess drought and flood risk.

7 The rest of the paper is organized as follows. Section 2 briefly introduces the instrumental and
8 reconstructed PDV records used in this research. The dynamic threshold run length extraction method
9 is introduced in Section 3, along with its comparison with the static threshold method and
10 determination of its parameters. The hydrologically important statistical signatures of PDV are
11 analysed and discussed in Section 4, with conclusions presented in Section 5.

12 **2. Data**

13 **2.1 Instrumental PDV records**

14 A number of instrumental IPO and PDO indices have been published to characterize PDV
15 since ~1900. There is a strong correlation between IPO and PDO indices, suggesting that both indices
16 represent a similar broad pattern of climate variability. The primary difference between the IPO and
17 PDO indices is that the IPO index is based on a broader spatial scale than the PDO index (Folland et
18 al., 2002; Verdon and Franks, 2006; Henley et al., 2011).

19 Three PDV indices are used in this study: (1) The unfiltered instrumental PDO index
20 (<http://research.jisao.washington.edu/pdo/PDO.latest.txt>), denoted as PDO_Mantua, is the monthly
21 standardized value of the leading principal component of monthly SST anomalies in the North Pacific
22 Ocean, poleward of 20N (Mantua et al., 1997; Zhang et al., 1997); (2) The unfiltered monthly IPO
23 index from Parker et al. (2007), denoted as IPO_Parker, is the second covariance empirical orthogonal
24 function of low-pass-filtered SST; (3) The unfiltered monthly Tripole Index (TPI) for the IPO (Henley



1 et al., 2015), denoted as IPO_Henley, is based on the SST anomalies in three large geographic regions
2 of the Pacific. Annual (calendar year) values of these three indices are taken as the average of
3 monthly values.

4 Figure 1 shows time series of the three indices along with probability density plot, quantile-
5 quantile (QQ) plot and autocorrelation plots. The distribution density plot and QQ plot of the annual
6 PDV indices suggest that the instrumental PDV indices are approximately normally distributed with
7 zero mean and unit standard deviation (except for IPO_Parker). The lag-one autocorrelation lies
8 between about 0.3 and 0.5. By lag three, the autocorrelations lie within the 95% confidence bands for
9 zero autocorrelation.

10 The instrumental PDV indices are used for two purposes in this study:

11 (a) to assess the similarity between IPO and PDO indices. If instrumental IPO and PDO
12 indices can be used to represent the same broad pattern of climate variability - and noting that
13 different reconstructions are calibrated against different instrumental indices - there is no need to
14 recalibrate all reconstructions against the same instrumental data.

15 (b) to guide the selection of the parameters of the run length extraction method. If selected
16 parameters can identify credible run lengths in the instrumental PDV record, it is assumed that they
17 can also produce credible run lengths in the reconstructed record.

18 **2.2 PDV reconstructions**

19 Twelve published annual PDV reconstructions from different sites with different proxies are
20 used in this study. Of these, nine reconstructed the last 300-400 years and three reconstructed the last
21 ~1000 years. The reconstructions are based on paleo proxies from the northern and southern, and
22 western and eastern regions of the Pacific Ocean. A summary of these reconstructions is presented in
23 Table 1, while Figure 7 in Section 3.2.3 presents time series plots of the 12 reconstructed indices, and
24 they are denoted throughout by the four-character author name and two-character publishing year. It is



1 worth noting that Mann09 was not directly calibrated to an instrumental PDV index, and also has a
2 negative correlation with other reconstructions. However, because the correlation of this
3 reconstruction with the other ones is relatively high (Henley, 2017), the Mann09 reconstruction is
4 retained with the sign reversed.

5 **3. Run length extraction**

6 **3.1 Static and dynamic threshold run length extraction**

7 **methods**

8 A number of studies have attempted to extract phases (or run lengths) from PDV
9 reconstructions. Verdon and Franks (2006) used the Mann-Whitney test method to determine whether
10 a step change occurred within a 30-year moving window. Each step change point was treated as a
11 potential crossing. They used zero as a static threshold to define the sign of PDV phases (Verdon-
12 Kidd, personal communication, 24 April, 2017). The instrumental mean and ± 0.5 were used as static
13 thresholds in studies by Henley et al. (2011) and Vance et al. (2015) respectively, and later, Henley et
14 al. (2017) used the long-term IPO Tripole Index (TPI) mean as the threshold to determine crossings
15 and the sign of PDV phases.

16 However, for reconstructions Macd05, Mann09 and Vanc15 which exhibit centennial trends
17 (see Table 1 and Figure 2), the above mentioned methods may not identify meaningful decadal phases.
18 Extraordinary long run lengths are identified if the overall mean is used as a threshold with some runs
19 lasting several centuries (see Figure 2). It is not clear whether these centennial trends are due to low
20 frequency climate variability or due to unknown local factors affecting the proxy. With such
21 reconstructions, one can either exclude them from analysis or filter out the centennial trends.
22 Exclusion forgoes any useful information from the reconstructions. On the other hand, filtering out
23 the centennial trends offers the prospect of extracting possibly useful information about run lengths.



1 We adopt the latter approach proposing a dynamic threshold framework to filter out
 2 centennial trends. In this framework, potential crossings are taken to be the change points detected by
 3 the Mann-Whitney test, and the sign of PDV phases is determined by a dynamic threshold that takes
 4 centennial trends into consideration - this relaxes the restriction of a static threshold defining a phase
 5 crossing. The key steps of this framework are:

6 1. Detect step-change points. For a given reconstruction, apply a change point detection
 7 method. This study will use the Mann-Whitney test method (Mauget, 2003) with a given window
 8 width w and confidence level α to identify the step-change points. However, other methods can be
 9 used.

10 2. Merge consecutive step-change points. When two or more step change points occur in
 11 consecutive years, replace them with a single change point. That is, if there are either $2n-1$ or $2n$ years
 12 that are determined as consecutive change points ($n=1, 2, \dots$), replace them with one change point at
 13 year n . This guarantees that the new change point is always closest to the middle of the run of
 14 consecutive step-change points.

15 3. Assign a phase to each run defined by the interval bounded by two step-change points.
 16 Let $i(t)$ denote the PDV index in year t , $i_f(t)$ the filtered PDV index using a first-order Butterworth
 17 filter (Selesnick and Burrus, 1998) with cut-off frequency $1/y \text{ year}^{-1}$, t_{cp}^k the year that the k^{th} change
 18 point occurs (from steps 1 and 2) and $s(t)$ the phase state of year t . The mean PDV index for the k^{th}
 19 run is

20
$$\bar{i}^k = \frac{1}{(t_{cp}^k - t_{cp}^{k-1} + 1)} \sum_{t_{cp}^{k-1}}^{t_{cp}^k} i(t) \quad (1)$$

21 while the corresponding mean of the filtered index is

22
$$\bar{i}_f^k = \frac{1}{(t_{cp}^k - t_{cp}^{k-1} + 1)} \sum_{t_{cp}^{k-1}}^{t_{cp}^k} i_f(t) \quad (2)$$



1 The phase of the k^{th} run length is then defined by

$$2 \qquad s(t) = \begin{cases} 1 & \text{if } \bar{i}^k \geq \bar{i}_f^k \\ 0 & \text{if } \bar{i}^k < \bar{i}_f^k \end{cases}, t \in (t_{cp}^{k-1}, t_{cp}^k - 1) \qquad (3)$$

3 4. Determine run lengths using the time series $s(t)$. The run length of the k^{th} run will be

$$4 \qquad t_{cp}^k - t_{cp}^{k-1} .$$

5 This dynamic threshold method can be seen as a generalization of previous run length
6 extraction methods. If parameter y is set to the total number of years in a given reconstruction, the
7 dynamic threshold method reduces to the method used in Verdon and Franks (2006). If change points
8 are defined by the PDV index crossing the threshold and the cut-off frequency parameter y of the first-
9 order Butterworth filter is set to the total number of years in the reconstruction, the dynamic threshold
10 reduces to the method used in Henley et al. (2011) and Henley et al. (2017).

11 Three parameters need to be specified in the dynamic threshold method: window width w and
12 confidence level α in the Mann-Whitney test, and cut-off frequency y of the Butterworth filter. To
13 filter out centennial trends, y should be set large enough to account for the centennial trend. Hence,
14 $y=100$ is used in this study. It has been shown that reconstructions are sensitive to the choice of the
15 window used in the Mann-Whitney test as well as the choice of threshold (Henley et al., 2017; Henley,
16 2013). Accordingly, this study investigates the sensitivity of dynamic threshold method results to the
17 choice of window width w in the Mann-Whitney test. This sensitivity analysis is used to guide the
18 selection of parameters to be used in the analysis of all 12 PDV reconstructions.



1 **3.2 Results from different run length extraction**

2 **methods**

3 **3.2.1 Comparison of static and dynamic threshold**

4 To further demonstrate the shortcomings of the static threshold method and the need for a
5 dynamic threshold method, run length samples are extracted using the dynamic threshold method
6 proposed in this study (denoted as “Dynamic” in the figures) and the static threshold method used in
7 Verdon and Franks (2006) (denoted as “Static” in the figures). These two methods only differ in step
8 3 of Section 3.1. In the static method, the overall mean defined as $\bar{i}_f = \frac{1}{n} \sum_1^n i_f(t)$ (in which n is
9 number of years in corresponding reconstruction) is used as the threshold instead of the dynamic
10 threshold defined in equation (2). Extracted run length samples using the dynamic and static threshold
11 methods from 3 millennium long reconstructions with centennial trends are plotted in Figure 2. Black
12 lines are filtered standardized PDV indices; green lines represent the Butterworth (1/100) filtered data
13 in the dynamic plots and the overall mean in the static plots; red lines represent the run extracted
14 using either the dynamic or static thresholds method. The run length distributions are plotted in Figure
15 3.

16 It is shown in Figure 2 that extraordinarily long runs are identified when the static threshold
17 method is used - the stronger the centennial trends, the longer the runs. On the other hand, use of the
18 dynamic threshold framework appears to filter centennial trends and produce more meaningful run
19 lengths. The run length density plots show that the dynamic threshold method for the Macd05 and
20 Mann09 reconstructions removes centennial-scale runs. In the case of Vanc15 the centennial trends
21 are muted, possibly because Vanc15 is annually resolved and very accurately dated (Vance et al.,
22 2015) and as such is more likely to exhibit more realistic annual to decadal variability than Macd05
23 and Mann09 (see also Figure 2).



1 **3.2.2 Window width sensitivity analysis**

2 Henley (2013) demonstrated that the mean run length is strongly dependent on the choice of
3 window width in Mann-Whitney method and concluded that a subjective choice of window width is
4 likely to bias the resulting run length distributions. To explore the sensitivity of window width on run
5 length distributions, run length boxplots for all reconstructions are plotted in Figure 4 for window
6 widths from 10 to 60 years with step of 2, with positive and negative samples plotted together and
7 separately.

8 Figure 4 shows that the run length distributions clearly depend on the choice of window width
9 with longer window widths leading to longer run length samples. However, the change is gradual
10 particularly for shorter window widths. As a result, run length distributions are not particularly
11 sensitive to small variations in window width of the order of 10 years. That is, runs extracted by using
12 20-year window width are different from using 50-year window width, but remain similar to using 30-
13 year window width.

14 **3.2.3 Determination of parameters and application of dynamic threshold** 15 **method**

16 To guide the selection of an operationally useful window width, several window widths are
17 applied to the three instrumental PDV time series to identify which produces results that match
18 existing knowledge about run length in instrumental periods. Four window widths, 10, 20, 30 and 40
19 years, are applied to the instrumental time series with the extracted runs plotted in Figure 5.
20 Interpretation of Figure 5 is illustrated for Figure 5d. The green line represents the IPO_Henley index,
21 from which three runs, denoted by green circles, are extracted: ~1900-1938 (positive), ~1938-1976
22 (negative), ~1976-2004 (positive). These are very similar (in number, sign and duration) to the black
23 (PDO_Mantua) and red run lengths (IPO_Parker). It can be seen that a window width of 10 years
24 identifies very short runs which are more likely to represent random fluctuations rather than different



1 PDV phases - we note that Henley et al. (2017) discarded runs with lengths less than 5 years.

2 However, for the longer window widths of 20, 30 and 40 years, the differences become more muted.

3 To better understand these differences, consider window widths of 20 and 40 years. When the
4 40-year window is used, the first 20 years of the window are compared to the last 20 years. As a result,
5 runs of less than 20 years may be overlooked. To demonstrate this, different window widths are
6 applied to all reconstructions to extract run lengths, from which conditional run length distributions
7 are determined. Figure 6 presents run length distributions in six panels. Panels (a1) to (a3) compare
8 distributions for 20 and 40-year windows. Panel (a1) shows the unconditional run distribution, while
9 panel (a2) shows the distribution conditioned on runs less than 20 years and panel (a3) shows the
10 distribution conditioned on runs greater than 20. Panels (b1) to (b3) repeat this for 20 and 60-year
11 windows with conditioning on the 30-year length.

12 Panels (a1) and (b1) show that the unconditional run length distributions are different using
13 different window widths with the difference greater for a greater difference in window width.
14 However, panels (a3) and (b3) show that the conditional distributions for the longer runs are very
15 similar. This shows that both the short and long window widths extract the longer runs in a largely
16 consistent manner. However, the choice of window width strongly affects the distribution of shorter
17 runs as shown in panels (a2) and (b2). Overall, this demonstrates that shorter window widths can “see”
18 higher frequency runs better than longer window widths but both “see” the same distribution of lower
19 frequency runs.

20 These considerations suggest that the window width should be as small as possible, yet
21 sufficiently big to filter out runs that are considered more likely to be random fluctuations rather than
22 PDV phases. Therefore, in this study, a 20-year window is used to ensure decadal or longer runs are
23 identified. Figure 7 plots the extracted runs for the 12 reconstructions using the dynamic threshold
24 method with 90% confidence level and 20-year window in the Mann-Whitney test. The PDV indices
25 plotted in Figure 7 were filtered using a Butterworth filter with a cut-off frequency of $1/11 \text{ years}^{-1}$ to
26 more clearly present the decadal variability in each reconstruction. In the case of Verd06, no PDV



1 indices are available, so the original runs are plotted. Visual inspection of Figure 7 suggests that the
2 identified runs in all reconstructions seem largely consistent with the multi-decadal fluctuations in the
3 filtered PDV indices. In the case of reconstructions with pronounced centennial trends (such as
4 Macd05 and Mann09), positive and negative runs tend to match multi-decadal fluctuations in the
5 filtered indices even when the indices persist above or below the long-term average over centennial
6 scales.

7 **4. Hydrologically important statistical** 8 **signatures associated with PDV**

9 The dynamic threshold method allows run lengths to be extracted from all the PDV
10 reconstructions, even from those with centennial trends. If we had a perfect reconstruction there
11 would be no need for any further reconstructions. However, the fact is that each paleoclimate
12 reconstruction is subject to errors, both random and systematic, that are not fully understood.
13 Therefore, it is pertinent to identify statistical features that are common to the reconstructions and also
14 important to hydroclimatic risk assessments. This addresses the second objective of this study. In
15 particular, the analysis seeks to answer the following questions: Are the distributions of positive and
16 negative PDV run lengths statistically different, and is the variability seen in the instrumental record
17 (the last ~100 years), or the shorter 300-400 year paleoclimate reconstructions, representative of the
18 full range of variability that has occurred in the past or of what is plausible in the future? These
19 questions are of particular importance for predicting near-term PDV phase behaviour and assessing
20 near-term flood and drought risk and therefore are the primary focus of this section.



1 **4.1 Are run lengths different during positive and** 2 **negative PDV phases?**

3 Vance et al. (2015) analysed an IPO reconstruction from AD 1000 to 2003 and concluded that
4 the positive IPO phase ($IPO > 0.5$) has an average duration of 14 years and the negative IPO phase
5 ($IPO < -0.5$) has an average duration of 9 years. This is the first known analysis that considered
6 positive and negative PDV phases separately.

7 To explore the hypothesis that run lengths have different distributions during positive and
8 negative PDV phases, boxplots of run lengths for each reconstruction are shown in Figure 8. Panel (a)
9 shows boxplots for all run lengths for each reconstruction, Panel (b) shows boxplots of pooled
10 positive and negative run lengths, while Panel (c) shows boxplots of positive and negative run lengths
11 for each reconstruction. Panel (b) suggests that the pooled median positive and negative run lengths
12 are similar, but that positive run lengths are more variable. However, Panel (c) shows that there is
13 considerable variability between reconstructions. The absence of a consistent difference between
14 positive and negative run distributions may be a consequence of sampling variability masking any
15 underlying difference. This seems reasonable given that the average number of phases (either positive
16 or negative) per reconstruction is only 13. A more formal statistical analysis is required to explicitly
17 deal with sampling variability arising from the small samples.

18 Henley et al. (2011) adopted the Gamma distribution as the probability model of PDV run
19 lengths. The Gamma distribution has two parameters which are related to the mean μ and standard
20 deviation σ of the run lengths. For each reconstruction, the positive and negative runs are extracted
21 and the posterior distribution of the mean and standard deviation is inferred using MCMC (Gelman
22 and Rubin, 1992). Figure 9 presents the joint posterior distribution of the differences $\mu^+ - \mu^-$ and
23 $\sigma^+ - \sigma^-$ for each reconstruction. If the difference between positive and negative runs is statistically
24 significant, one would expect the posterior distribution to be well removed from the zero-difference
25 origin.



1 Inspection of Figure 9 suggests there is no strong and consistent evidence to reject the
2 assumption that positive and negative phase run length distributions are the same. Although the mean
3 of positive phase run lengths tends to be longer than negative phase run lengths in several
4 reconstructions (e.g. Vanc15, Hen11, Lins08, Darr01), the statistical significance of this difference is
5 weak and the phenomenon is not consistent across all reconstructions. A similar conclusion applies to
6 the standard deviation of positive and negative phase run lengths. This highlights the fundamental
7 limitation of the small PDV run samples derived from the reconstructions.

8 **4.2 Are run lengths stationary over the last millennium?**

9 With paleoclimate reconstructions becoming longer, questions arise whether the PDV run
10 length has been stationary in the past millennium and whether a shorter paleoclimate reconstruction is
11 representative of the full range of variability that has occurred in the past or of what is plausible in the
12 future. The stationarity issue has been explored most PDV reconstructions used in this study. Biondi
13 et al. (2001) identified weakened amplitude of bi-decadal oscillations in the late 1700s to mid-1800s.
14 Based on reconstruction over 1700-1979, D'Arrigo et al. (2001) declared that variations at around 12-
15 17 years are considerably more pronounced from 1700-1849 relative to 1850-1979 with a shift
16 towards decreased amplitude since about 1850. Using visual inspection, Gedalof and Smith (2001)
17 stated that the 30-70 year PDV frequency is confined to the pre-1840 portion of the series over the
18 period of 1599-1983. D'Arrigo and Wilson (2006) reported a broad range of lower (multi-decadal to
19 centennial) frequencies over the 1565-1988 period. Shen et al. (2006) pointed out that the major PDO
20 regime timescale modes of oscillation have not been persistent over 1470-1998 and that 75-115 year
21 and 50–70 year oscillations dominated the periods before and after 1850, respectively. Linsley et al.
22 (2008) argued that decadal to inter-decadal variability in the south Pacific convergence zone region
23 has been relatively constant over 1650-2004. MacDonald and Case (2005) presented evidence for a
24 strong and persistent negative PDO state during the medieval period (AD 900 to 1300), suggesting a
25 cool north eastern Pacific at that time. Mann et al. (2009) observed that the Little Ice Age (LIA, 1400
26 to 1700) and the Medieval Climate Anomaly (MCA, 950 to 1250) showed extra variability and



1 persistence, challenging the assumption of a stationary climate in the past millennium. The
2 mechanism responsible for the extra variability and persistence in studies done by MacDonald and
3 Case (2005) and Mann et al. (2009) is unclear and may be explained by centennial climate variability.
4 It is unknown whether the PDV structure in the past millennium has remained stationary during
5 periods that display extra variability and persistence. This issue can be investigated by filtering out the
6 centennial trends.

7 All the above stationarity studies are based on single reconstructions. A further complication
8 arises for reconstructions that only cover the last ~300-400 years where the underlying non-
9 stationarity may be masked as a consequence of sampling variability or the sampling variability is
10 misinterpreted as non-stationarity. To mitigate this, here we use only millennium-length records,
11 Macd05 (993-1996), Mann09 (500-2006) and Vanc15 (1000-2003), in an attempt to obtain
12 statistically meaningful conclusions on PDV stationarity. The millennium-length records are split into
13 pre-1600 and post-1600 samples to explore whether run length characteristics are statistically similar
14 in these two periods and to shed light on whether shorter records (either the ~100 year instrumental
15 record or the 300-400 year paleoclimate reconstructions) are able to represent the full range of
16 variability. The year 1600 is selected as the change point because most of the shorter PDV
17 reconstructions date back to ~1600 (refer to Table 1 and Figure 2 for details).

18 Figure 10 presents boxplots of pre-1600 and post-1600 run lengths: panel (a) presents
19 boxplots of pooled pre-1600 and post-1600 run lengths, while panel (b) shows boxplots of pre-1600
20 and post-1600 runs for each of the three millennium-length reconstructions. A consistent pattern
21 emerges in which the median run length and interquartile range are greater in the pre-1600 period for
22 both pooled samples (Figure 10a) and each individual reconstruction (Figure 10b).

23 Bearing in mind that the samples are small, a Gamma distribution was inferred for the pre-
24 and post-1600 samples. Figure 11 presents the joint posterior distribution of the differences in the pre-
25 and post-1600 means and standard deviations for each reconstruction using all run lengths. A
26 consistent pattern once again emerges. For all three reconstructions the posteriors of the pre- and post-



- 1 1600 differences lie in the lower left quadrant with the origin minimally intersecting with the posterior.
- 2 Therefore, the evidence that there is a difference is visually strong.

3 **4.3 Discussion**

4 All of the reconstructions in Table 1 have focused on the run length of PDV, with positive
5 and negative PDV run lengths being pooled together in their analysis, except for Vance et al. (2015).
6 Based on a millennium IPO reconstruction, Vance et al. (2015) found that IPO has an average positive
7 phase ($IPO > 0.5$) duration of 14 years and a negative phase ($IPO < -0.5$) duration of 9 years over
8 A.D.1000-2003, and concluded that positive and negative phases durations and frequencies were
9 different in the last millennium. This is the first study that addresses positive and negative PDV runs
10 separately. However, based on comparisons with other reconstructions our results show that there
11 remains uncertainty as to whether or not the run lengths of positive epochs are statistically different to
12 the run length of negative epochs. It is important to keep in mind that the analysis here is
13 fundamentally limited by small PDV run length sample sizes from most of the reconstructions. Two
14 theories may explain the absence of a consistent difference between positive and negative run
15 distributions. One is that no statistically meaningful differences exist and the other is that differences
16 do exist but are masked by sampling variability. The latter seems plausible given the average number
17 of phases (either positive or negative) per reconstruction are only 13.

18 All three millennium-length records lead to higher inferred run length mean and standard
19 deviation in the pre-1600 samples, suggesting longer and more varied PDV runs during the period
20 before 1600 AD. The fact that the differences in the mean and standard deviation of run lengths pre-
21 and post-1600 appear to be statistically significant raises an important question, should the
22 information from pre-1600 reconstructions be used to infer PDV behaviour in the near climate. If one
23 adopts a gradualist view of climate non-stationarity, the immediate past would be considered more
24 representative of the present than the more distant past. Setting aside for the moment the possibility
25 that the climate-proxy relationship may be not be stationary, the gradualist perspective would support



1 discarding the information from the pre-1600 reconstruction on the grounds that it introduces bias
2 when inferring a statistical model of near climate PDV runs. We offer two reasons opposing this
3 perspective.

4 First, there is no assurance that climate non-stationarity evolves in a gradual manner so that
5 PDV behaviour in the near climate is better represented by the post-1600 climate than the pre-1600
6 climate. Indeed there is evidence to the contrary, namely that non-stationarity may be characterized by
7 seemingly shorter-term shifts. Ho et al. (2017) found that American streamflow in the twentieth
8 century featured longer wet and dry spells compared to the preceding 450 years, suggesting the
9 possibility of occurrence of extended dry or wet periods in the future that exceeds variability
10 presented in instrumental and short (i.e. < 500 years) paleoclimate reconstructions. Similar
11 conclusions are also drawn in other studies (Vance et al., 2015; Tozer et al., 2016; Ho et al., 2015b, a).

12 Second, the inclusion of pre-1600 records offers a significantly larger sample of run lengths
13 and therefore is more likely to capture the full range of what is plausible. Extended wet/dry periods
14 considered as rare and extreme (or even impossible) based on recent (i.e. instrumental) history might
15 actually be more likely than thought (or even common) when put into context of climate conditions
16 seen over the last 1000-2000 years. Even if one accepts the gradualist view of non-stationarity, the
17 increased information about PDV more than outweighs the potential bias arising from the use of an
18 apparently statistically inconsistent record. At least until such time as it is proven that the climate has
19 totally shifted and that what occurred in the past is no longer possible – also required is identification
20 of when that shift occurred.

21 It is therefore clear that in addition to short instrumental records inadequately sampling the
22 length and severity of dry/wet epochs, the shorter paleoclimate reconstructions may similarly
23 misrepresent hydroclimatic variability and persistence. For instance, from Ho et al. (2015a), when
24 using the period 1684–1980, both the driest and wettest instrumental decadal rainfall lies closer to the
25 middle of the minimum and maximum decadal reconstructed rainfall range. However, using the same
26 method, Ho et al. (2015a) found that when 2751 years of reconstructions were examined, the wettest



1 instrumental decadal rainfall is near the bottom of the maximum decadal reconstructed rainfall range,
2 while the driest instrumental decadal rainfall is near the top of the minimum decadal reconstructed
3 rainfall range. This suggests that statistics inferred from either instrumental data or short palaeoclimate
4 reconstructions will underestimate the variability contained in the longer palaeoclimate
5 reconstructions. Therefore, research focused on developing and analysing information about
6 hydroclimatic conditions over the last 1000 years or more (e.g. multi-millennium palaeoclimate
7 reconstructions) is critical if we are to better understand, quantify and manage the full range of
8 hydroclimatic variability, and associated risks, that are possible in the future.

9 This study is affected by two fundamental limitations. One is the small PDV run sample size
10 from the reconstructions. For instance, the average number of phases (either positive or negative) per
11 reconstruction are only 13. Another limitation is that all statistical signatures are based on
12 palaeoclimate reconstructions. Although multiple reconstructions are used, this study is constrained by
13 the intrinsic limitations of palaeoclimate reconstruction based interpretations, such as assumptions of
14 stationarity of the proxy-climate relationship (Phipps et al., 2013). The possibility that all these
15 reconstructions are biased in the same direction cannot be ruled out and as such conclusions based on
16 multiple reconstructions may also be biased.

17 5. Conclusions

18 PDV, and associated run lengths of predominantly dry or wet conditions, has profound
19 implications for precipitation/streamflow prediction, flood/drought risk assessment and water resource
20 management. Therefore, a better understanding of the statistical characteristics of PDV is needed.
21 However, because instrumental records are short (~100 years at best in Australia), there is
22 considerable uncertainty about the key statistical signatures of PDV, including run lengths.
23 Palaeoclimate reconstructions, serving as a vehicle to provide longer realizations of PDV, have been
24 widely studied and used. However, for various reasons (e.g. proxy sources, site locations, proxy
25 resolutions, reconstruction methods and local non-PDV effects) temporal coherence between different



1 reconstructions varies significantly (Kipfmüller et al., 2012). Hence, one PDV reconstruction may
2 lead to significantly different conclusions from another reconstruction.

3 The aim of this paper was to explore the characteristics of key statistical signatures of PDV
4 based on multiple PDV reconstructions. We focused on the duration of dry or wet hydrological
5 epochs (i.e. run lengths) as the key statistical signature and developed a robust method for extracting
6 run lengths from multiple reconstructions. Extracting run lengths objectively is challenging, given
7 interactions with sources of variability on a variety of temporal and spatial scales. For instance, when
8 millennium-length reconstructions are used, variability and persistence at the centennial scale can lead
9 to biased characterisation of decadal climate variability. The dynamic threshold framework introduced
10 in this study, which takes centennial trends into consideration, was shown to extract meaningful run
11 length information from multiple reconstructions.

12 No strong evidence was found to support the assumption that run lengths have statistically
13 different distributions during positive and negative PDV phases. Analysis based on three millennium
14 long reconstructions suggests that it is more likely than not that PDV run length has been non-
15 stationarity in the past millennium. This again highlights that the instrumental record (~100 years at
16 best), and even short paleoclimate reconstructions (i.e. less than 400 years into the past), should not be
17 assumed to represent the full range of variability that has occurred in the past or what may occur in
18 the future. Caution should be exercised regarding assumptions that the climate is stationary and the
19 implications of a non-stationary climate should at least be tested. Longer climate reconstructions (i.e.
20 1000 years or more) appear to give more useful information and insights into what has occurred in the
21 past and also tell us more about what is plausible and what needs to be planned for in the future.

22 All of the reconstructions explored in this study have focused on the run length of PDV, yet
23 they have provided quite different run length characterizations. This again highlights that each
24 reconstruction is subject to errors, both random and systematic, that are not fully understood.
25 Therefore, a challenging and important research direction is the analysis of run length characteristics
26 with explicit consideration of the errors. Another important research direction is the continued



- 1 refinement of methods that utilize PDV information in water resource management as we work
2 toward improving how we assess and manage hydrological risks in a variable and changing climate.

3 **Acknowledgments**

- 4 Funding for this research was provided by Australian Research Council Linkage Grant
5 LP120200494 with further funding and/or in-kind support also provided by the NSW Office of
6 Environment and Heritage, Sydney Catchment Authority, Hunter Water Corporation, NSW Office of
7 Water, and NSW Department of Finance and Services.

8 **References**

- 9 Andreoli, R. V., and Kayano, M. T.: ENSO-related rainfall anomalies in South America and associated
10 circulation features during warm and cold Pacific decadal oscillation regimes, *International Journal of*
11 *Climatology*, 25, 2017-2030, [10.1002/joc.1222](https://doi.org/10.1002/joc.1222), 2005.
12 Atwood, A. R., Wu, E., Frierson, D. M. W., Battisti, D. S., and Sachs, J. P.: Quantifying Climate Forcings
13 and Feedbacks over the Last Millennium in the CMIP5–PMIP3 Models, *Journal of Climate*, 29, 1161-
14 1178, [10.1175/JCLI-D-15-0063.1](https://doi.org/10.1175/JCLI-D-15-0063.1), 2015.
15 Biondi, F., Gershunov, A., and Cayan, D. R.: North Pacific decadal climate variability since 1661,
16 *Journal of Climate*, 14, 5-10, [10.1175/1520-0442\(2001\)014<0005:NPDCVS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0005:NPDCVS>2.0.CO;2), 2001.
17 Chiew, F. H. S.: Estimation of rainfall elasticity of streamflow in Australia, *Hydrological Sciences*
18 *Journal*, 51, 613-625, [10.1623/hysj.51.4.613](https://doi.org/10.1623/hysj.51.4.613), 2006.
19 Cook, B. I., Smerdon, J. E., Seager, R., and Cook, E. R.: Pan-Continental Droughts in North America
20 over the Last Millennium, *Journal of Climate*, 27, 383-397, [10.1175/JCLI-D-13-00100.1](https://doi.org/10.1175/JCLI-D-13-00100.1), 2013.
21 D'Arrigo, R., Villalba, R., and Wiles, G.: Tree-ring estimates of Pacific decadal climate variability, *Clim*
22 *Dyn*, 18, 219-224, [10.1007/s003820100177](https://doi.org/10.1007/s003820100177), 2001.
23 D'Arrigo, R., and Wilson, R.: On the Asian expression of the PDO, *International Journal of Climatology*,
24 26, 1607-1617, [10.1002/joc.1326](https://doi.org/10.1002/joc.1326), 2006.
25 Dai, A.: The influence of the inter-decadal Pacific oscillation on US precipitation during 1923–2010,
26 *Clim Dyn*, 41, 633-646, [10.1007/s00382-012-1446-5](https://doi.org/10.1007/s00382-012-1446-5), 2013.
27 Dong, B., and Dai, A.: The influence of the Interdecadal Pacific Oscillation on Temperature and
28 Precipitation over the Globe, *Clim Dyn*, 45, 2667-2681, [10.1007/s00382-015-2500-x](https://doi.org/10.1007/s00382-015-2500-x), 2015.
29 Folland, C. K., Renwick, J. A., Salinger, M. J., and Mullan, A. B.: Relative influences of the interdecadal
30 Pacific oscillation and ENSO on the South Pacific convergence zone, *Geophysical Research Letters*, 29,
31 1643, [10.1029/2001GL014201](https://doi.org/10.1029/2001GL014201), 2002.
32 Gedalof, Z., and Smith, D. J.: Interdecadal climate variability and regime - scale shifts in Pacific North
33 America, *Geophysical Research Letters*, 28, 1515-1518, [10.1029/2000GL011779](https://doi.org/10.1029/2000GL011779), 2001.
34 Gelman, A., and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences,
35 *Statistical Science*, 7, 457-472, [10.1214/ss/1177011136](https://doi.org/10.1214/ss/1177011136), 1992.



- 1 Goodrich, G. B., and Walker, J. M.: The Influence of the PDO on Winter Precipitation During High-
2 and Low-Index ENSO Conditions in the Eastern United States, *Physical Geography*, 32, 295-312,
3 10.2747/0272-3646.32.4.295, 2011.
- 4 Henley, B., Gergis, J., Karoly, D., Power, S., Kennedy, J., and Folland, C.: A Tripole Index for the
5 Interdecadal Pacific Oscillation, *Clim Dyn*, 1-14, 10.1007/s00382-015-2525-1, 2015.
- 6 Henley, B. J., Thyer, M. A., Kuczera, G., and Franks, S. W.: Climate-informed stochastic hydrological
7 modeling: Incorporating decadal-scale variability using paleo data, *Water Resources Research*, 47,
8 W11509, 10.1029/2010WR010034, 2011.
- 9 Henley, B. J., Thyer, M. A., and Kuczera, G.: Climate driver informed short-term drought risk
10 evaluation, *Water Resources Research*, 49, 2317-2326, 10.1002/wrcr.20222, 2013.
- 11 Henley, B. J.: Pacific decadal climate variability: Indices, patterns and tropical-extratropical
12 interactions, *Global and Planetary Change*, 155, 42-55, 10.1016/j.gloplacha.2017.06.004, 2017.
- 13 Henley, B. J., Meehl, G., Power, S., Folland, C., King, A., Brown, J., Karoly, D., Delage, F., Gallant, A.,
14 and Freund, M.: Spatial and temporal agreement in climate model simulations of the interdecadal
15 pacific oscillation, *Environmental Research Letters*, 12, 044011, 10.1088/1748-9326/aa5cc8, 2017.
- 16 Ho, M., Kiem, A. S., and Verdon-Kidd, D. C.: A paleoclimate rainfall reconstruction in the Murray-
17 Darling Basin (MDB), Australia: 2. Assessing hydroclimatic risk using paleoclimate records of wet and
18 dry epochs, *Water Resources Research*, 51, 8380-8396, 10.1002/2015WR017059, 2015a.
- 19 Ho, M., Kiem, A. S., and Verdon-Kidd, D. C.: A paleoclimate rainfall reconstruction in the Murray -
20 Darling Basin (MDB), Australia: 1. Evaluation of different paleoclimate archives, rainfall networks,
21 and reconstruction techniques, *Water Resources Research*, 51, 8362-8379, 10.1002/2015WR017058,
22 2015b.
- 23 Ho, M., Lall, U., Sun, X., and Cook, E. R.: Multiscale temporal variability and regional patterns in 555
24 years of conterminous U.S. streamflow, *Water Resources Research*, 3047-3066,
25 10.1002/2016WR019632, 2017.
- 26 Hoell, A., and Funk, C.: Indo-Pacific sea surface temperature influences on failed consecutive rainy
27 seasons over eastern Africa, *Clim Dyn*, 43, 1645-1660, 10.1007/s00382-013-1991-6, 2014.
- 28 Hu, Z.-Z., and Huang, B.: Interferential Impact of ENSO and PDO on Dry and Wet Conditions in the
29 U.S. Great Plains, *Journal of Climate*, 22, 6047-6065, 10.1175/2009JCLI2798.1, 2009.
- 30 Kiem, A. S., Franks, S. W., and Kuczera, G.: Multi-decadal variability of flood risk, *Geophysical*
31 *Research Letters*, 30, 1035, 10.1029/2002GL015992, 2003.
- 32 Kiem, A. S., and Franks, S. W.: Multi-decadal variability of drought risk, eastern Australia,
33 *Hydrological Processes*, 18, 2039-2050, 10.1002/hyp.1460, 2004.
- 34 Kipfmueller, K. F., Larson, E. R., and St. George, S.: Does proxy uncertainty affect the relations
35 inferred between the Pacific Decadal Oscillation and wildfire activity in the western United States?,
36 *Geophysical Research Letters*, 39, PA2219, 10.1029/2011GL050645, 2012.
- 37 Kirtman, B., Power, S. B., Adedoyin, A. J., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore,
38 A. M., Kimoto, M., and Meehl, G.: Near-term climate change: projections and predictability, in:
39 *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth*
40 *Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press,
41 Cambridge, United Kingdom and New York, NY, USA, 953-1028, 2013.
- 42 Krishnan, R., and Sugi, M.: Pacific decadal oscillation and variability of the Indian summer monsoon
43 rainfall, *Clim Dyn*, 21, 233-242, 10.1007/s00382-003-0330-8, 2003.
- 44 Li, L., Li, W., and Kushnir, Y.: Variation of the North Atlantic subtropical high western ridge and its
45 implication to Southeastern US summer precipitation, *Clim Dyn*, 39, 1401-1412, 10.1007/s00382-
46 011-1214-y, 2012.
- 47 Linsley, B. K., Zhang, P., Kaplan, A., Howe, S. S., and Wellington, G. M.: Interdecadal - decadal
48 climate variability from multicoral oxygen isotope records in the South Pacific Convergence Zone
49 region since 1650 AD, *Paleoceanography*, 23, PA2219, 10.1029/2007PA001539, 2008.



- 1 Ma, Z.: The interdecadal trend and shift of dry/wet over the central part of North China and their
2 relationship to the Pacific Decadal Oscillation (PDO), *Chinese Science Bulletin*, 52, 2130-2139,
3 10.1007/s11434-007-0284-z, 2007.
- 4 MacDonald, G. M., and Case, R. A.: Variations in the Pacific Decadal Oscillation over the past
5 millennium, *Geophysical Research Letters*, 32, L08703, 10.1029/2005GL022478, 2005.
- 6 Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Ammann, C.,
7 Faluvegi, G., and Ni, F.: Global signatures and dynamical origins of the Little Ice Age and Medieval
8 Climate Anomaly, *Science*, 326, 1256-1260, 10.1126/science.1177303, 2009.
- 9 Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C.: A Pacific interdecadal climate
10 oscillation with impacts on salmon production, *Bulletin of the American Meteorological Society*, 78,
11 1069-1079, 10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2, 1997.
- 12 Mantua, N. J., and Hare, S. R.: The Pacific decadal oscillation, *Journal of Oceanography*, 58, 35-44,
13 10.1023/a:1015820616384, 2002.
- 14 Mauget, S. A.: Multidecadal regime shifts in US streamflow, precipitation, and temperature at the
15 end of the twentieth century, *Journal of Climate*, 16, 3905-3916, 10.1175/1520-
16 0442(2003)016<3905:MRSIUS>2.0.CO;2, 2003.
- 17 McCabe, G. J., Ault, T. R., Cook, B. I., Betancourt, J. L., and Schwartz, M. D.: Influences of the El Niño
18 Southern Oscillation and the Pacific Decadal Oscillation on the timing of the North American spring,
19 *International Journal of Climatology*, 32, 2301-2310, 10.1002/joc.3400, 2012.
- 20 McGregor, S., Timmermann, A., and Timm, O.: A unified proxy for ENSO and PDO variability since
21 1650, *Climate of the Past*, 6, 1-17, 10.5194/cp-6-1-2010, 2010.
- 22 Mehta, V. M., Rosenberg, N. J., and Mendoza, K.: Simulated Impacts of Three Decadal Climate
23 Variability Phenomena on Water Yields in the Missouri River Basin1, *JAWRA Journal of the American
24 Water Resources Association*, 47, 126-135, 10.1111/j.1752-1688.2010.00496.x, 2011.
- 25 Micevski, T., Franks, S. W., and Kuczera, G.: Multidecadal variability in coastal eastern Australian
26 flood data, *Journal of Hydrology*, 327, 219-225, <https://doi.org/10.1016/j.jhydrol.2005.11.017>, 2006.
- 27 Parker, D., Folland, C., Scaife, A., Knight, J., Colman, A., Baines, P., and Dong, B.: Decadal to
28 multidecadal variability and the climate change background, *Journal of Geophysical Research:*
29 *Atmospheres*, 112, D18115, 10.1029/2007JD008411, 2007.
- 30 Phipps, S. J., McGregor, H. V., Gergis, J., Gallant, A. J. E., Neukom, R., Stevenson, S., Ackerley, D.,
31 Brown, J. R., Fischer, M. J., and Van Ommen, T. D.: Paleoclimate data–model comparison and the
32 role of climate forcings over the past 1500 years, *Journal of Climate*, 26, 6915-6936, 10.1175/JCLI-D-
33 12-00108.1, 2013.
- 34 Power, S., Casey, T., Folland, C., Colman, A., and Mehta, V.: Inter-decadal modulation of the impact
35 of ENSO on Australia, *Clim Dyn*, 15, 319-324, 10.1007/s003820050284, 1999.
- 36 Razavi, S., Elshorbagy, A., Wheeler, H., and Sauchyn, D.: Toward understanding nonstationarity in
37 climate and hydrology through tree ring proxy records, *Water Resources Research*, 51, 1813-1830,
38 10.1002/2014WR015696, 2015.
- 39 Reason, C. J. C., and Rouault, M.: ENSO-like decadal variability and South African rainfall, *Geophysical
40 Research Letters*, 29, 16-11-16-14, 10.1029/2002GL014663, 2002.
- 41 Saft, M., Western, A. W., Zhang, L., Peel, M. C., and Potter, N. J.: The influence of multiyear drought
42 on the annual rainfall-runoff relationship: An Australian perspective, *Water Resources Research*, 51,
43 2444-2463, 10.1002/2014WR015348, 2015.
- 44 Selesnick, I. W., and Burrus, C. S.: Generalized digital Butterworth filter design, *Trans. Sig. Proc.*, 46,
45 1688-1694, 10.1109/78.678493, 1998.
- 46 Shen, C., Wang, W. C., Gong, W., and Hao, Z.: A Pacific Decadal Oscillation record since 1470 AD
47 reconstructed from proxy data of summer rainfall over eastern China, *Geophysical Research Letters*,
48 33, L03702, 10.1029/2005GL024804, 2006.
- 49 Tozer, C. R., Vance, T. R., Roberts, J. L., Kiem, A. S., Curran, M. A. J., and Moy, A. D.: An ice core
50 derived 1013-year catchment-scale annual rainfall reconstruction in subtropical eastern Australia,
51 *Hydrol. Earth Syst. Sci.*, 20, 1703-1717, 10.5194/hess-20-1703-2016, 2016.



- 1 Vance, T. R., Roberts, J. L., Plummer, C. T., Kiem, A. S., and van Ommen, T. D.: Interdecadal Pacific
2 variability and eastern Australian megadroughts over the last millennium, *Geophysical Research*
3 *Letters*, 42, 129-137, 10.1002/2014GL062447, 2015.
- 4 Verdon, D. C., Wyatt, A. M., Kiem, A. S., and Franks, S. W.: Multidecadal variability of rainfall and
5 streamflow: Eastern Australia, *Water Resources Research*, 40, W10201, 10.1029/2004WR003234,
6 2004.
- 7 Verdon, D. C., and Franks, S. W.: Long-term behaviour of ENSO: Interactions with the PDO over the
8 past 400 years inferred from paleoclimate records, *Geophysical Research Letters*, 33, L06712,
9 10.1029/2005GL025052, 2006.
- 10 von Storch, H., Zorita, E., Jones, J. M., Gonzalez-Rouco, F., and Tett, S. F. B.: Response to Comment
11 on "Reconstructing Past Climate from Noisy Data", *Science*, 312, 529, 10.1126/science.1121571,
12 2006.
- 13 Wahl, E. R., Ritson, D. M., and Ammann, C. M.: Comment on "Reconstructing Past Climate from
14 Noisy Data", *Science*, 312, 529, 10.1126/science.1120866, 2006.
- 15 Wang, S., Huang, J., He, Y., and Guan, Y.: Combined effects of the Pacific Decadal Oscillation and El
16 Niño-Southern Oscillation on Global Land Dry–Wet Changes, *Scientific Reports*, 4, 6651,
17 10.1038/srep06651
- 18 <https://www.nature.com/articles/srep06651#supplementary-information>, 2014.
- 19 Wu, Z., Huang, N. E., Long, S. R., and Peng, C.-K.: On the trend, detrending, and variability of
20 nonlinear and nonstationary time series, *Proceedings of the National Academy of Sciences*, 104,
21 14889-14894, 10.1073/pnas.0701020104, 2007.
- 22 Zhang, Y., Wallace, J. M., and Battisti, D. S.: ENSO-like interdecadal variability: 1900-93, *Journal of*
23 *climate*, 10, 1004-1020, 10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2, 1997.

24

25

1 **Tables**2 **Table 1 Available PDV paleoclimate reconstructions and their descriptions**

Reconstruction	Abbreviation	Period	Location	Proxy
Biondi et al. (2001)	Bion01	1661-1991	North Pacific (North America)	Tree ring based PDO
D'Arrigo et al. (2001)	Darr01	1700-1979	Northeast Pacific (Western North America)	Tree ring based PDO
Gedalof and Smith (2001)	Geda01	1599-1983	North Pacific (Coastal western North America)	Tree ring based PDO
MacDonald and Case (2005)	Macd05	993-1996	North Pacific (South California and Canada)	Tree ring based PDO
D'Arrigo and Wilson (2006)	Darr06	1565-1988	North Pacific (Eastern Asian)	Tree ring based PDO
Shen et al. (2006)	Shen06	1470-1998	Western North Pacific (Eastern China)	Proxy data of Summer Rainfall
Verdon and Franks (2006)	Verd06	1662-1998	Both North and South Pacific	Other reconstructions based PDO
Linsley et al. (2008)	Lins08	1650-2004	South Pacific (Fiji and Tonga)	Oxygen isotope from coral cores based IPO
Mann et al. (2009)	Mann09	500-2006	Global	Global tree ring, ice core, coral, sediment, and other assorted proxy records
McGregor et al. (2010)	Magr10	1650-1977	Global	low frequency variability of the proxies of ENSO
Henley et al. (2011)	Hen11	1471-2000 (filtered)	Both North and South Pacific	Other reconstructions
(Vance et al. (2015))	Vanc15	1000-2003 (filtered)	East Antarctica	Law Dome ice core based IPO

3

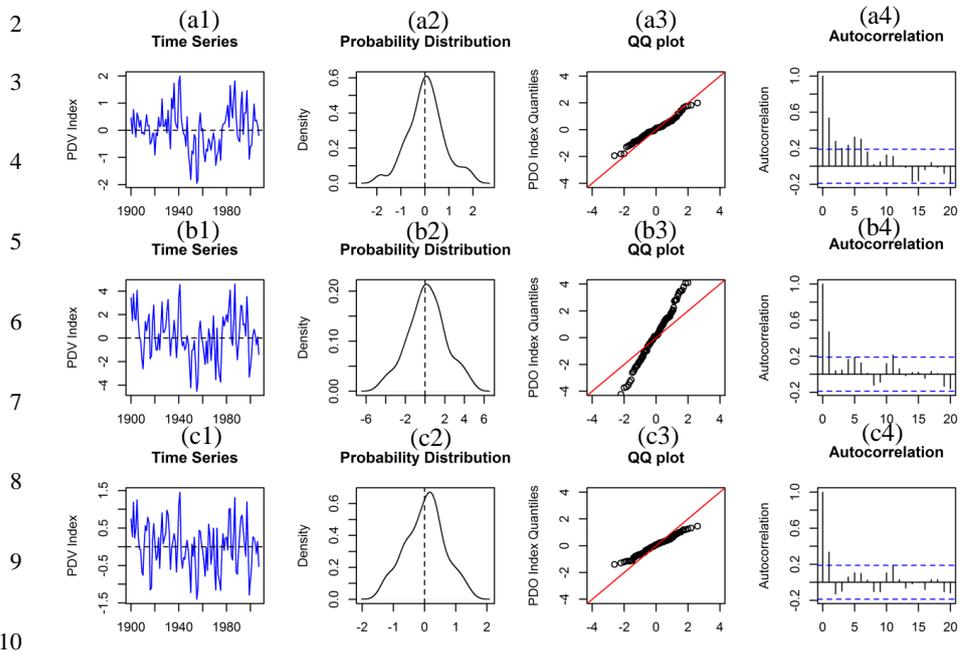


1 A list of figure captions

- 2 Figure 1 Annual instrumental time series, probability density plot, QQ plot and autocorrelation plot:
3 (a): PDO from Mantua (1900-2015); (b): IPO from Parker (1871-2007); (c): IPO from Henley (1870-
4 2007)
- 5 Figure 2 Run length time series extracted using the dynamic and static threshold methods for
6 reconstructions with centennial trends
- 7 Figure 3 Comparison of run length distributions extracted using the dynamic and static threshold
8 methods in reconstructions with centennial trends
- 9 Figure 4 Boxplot of run lengths samples from all reconstructions with different window width in
10 Mann-Whitney method
- 11 Figure 5 Time series plots of different instrumental PDV indices and corresponding run lengths
12 (represented by dots in corresponding colour) extracted using different window: (a) 10; (b) 20; (c) 30;
13 (d) 40 years
- 14 Figure 6 Density plot of run length samples by applying different window widths to all
15 reconstructions, (a1-a3): comparison between window width 20 and 40, in which (a1) is unconditional
16 run length distribution, panel (a2) is the distribution conditioned on run lengths less than 20 years and
17 panel (a3) is the distribution conditioned on run lengths greater than 20 years; (b1-b3): comparison
18 between window width 20 and 60, in which (b1) is unconditional run length distribution, panel (b2) is
19 the distribution conditioned on run lengths less than 30 years and panel (b3) is the distribution
20 conditioned on run lengths greater than 30 years
- 21 Figure 7 Filtered PDV reconstruction time series (black line) with extracted run length (red line)
- 22 Figure 8 Boxplot of run lengths samples from different reconstructions: (a) all run lengths for each
23 reconstruction; (b) pooled positive (grey) and negative (white) run lengths; (c) positive (grey) and
24 negative (white) run lengths for each reconstruction
- 25 Figure 9 Density and scatter plots of the differences between mean and standard deviation of positive
26 and negative run length simulations (positive minus negative)
- 27 Figure 10 Boxplot of run lengths samples from different reconstructions (white box indicates pre-
28 1600 and grey box indicates post-1600): (a) pooled run lengths from all three reconstructions; (b) run
29 lengths for each reconstruction
- 30 Figure 11 Density and scatter plots of the differences between mean and standard deviation of pre-
31 1600 and post-1600 run length simulations (post-1600 minus pre-1600): (a) Macd05; (b) Mann09; (c)
32 Vanc15



1 **Figures**

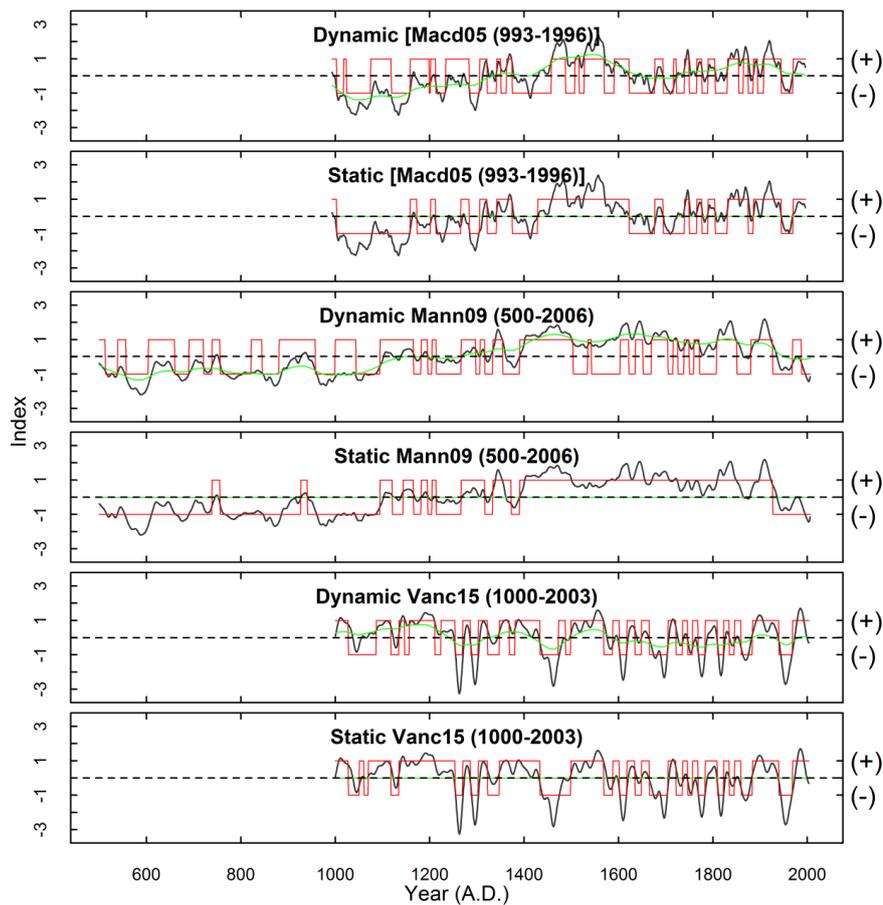


11 **Figure 1 Annual instrumental time series, probability density plot, QQ plot and autocorrelation plot: (a):**
 12 **PDO from Mantua (1900-2015); (b): IPO from Parker (1871-2007); (c): IPO from Henley (1870-2007)**

13



1



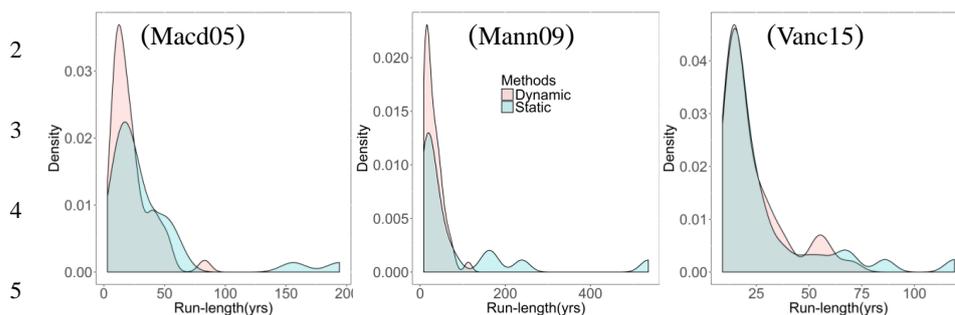
2

3 **Figure 2** Run length time series extracted using the dynamic and static threshold methods for
4 reconstructions with centennial trends

5



1

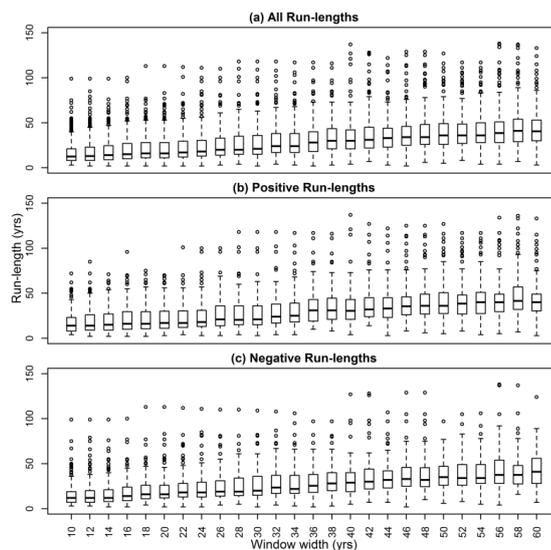


2
3
4
5
6 **Figure 3 Comparison of run length distributions extracted using the dynamic and static threshold**
7 **methods in reconstructions with centennial trends**

8



1



2

3 **Figure 4** Boxplot of run lengths samples from all reconstructions with different window width in Mann-
4 **Whitney method**

5



1

2

3

4

5

6

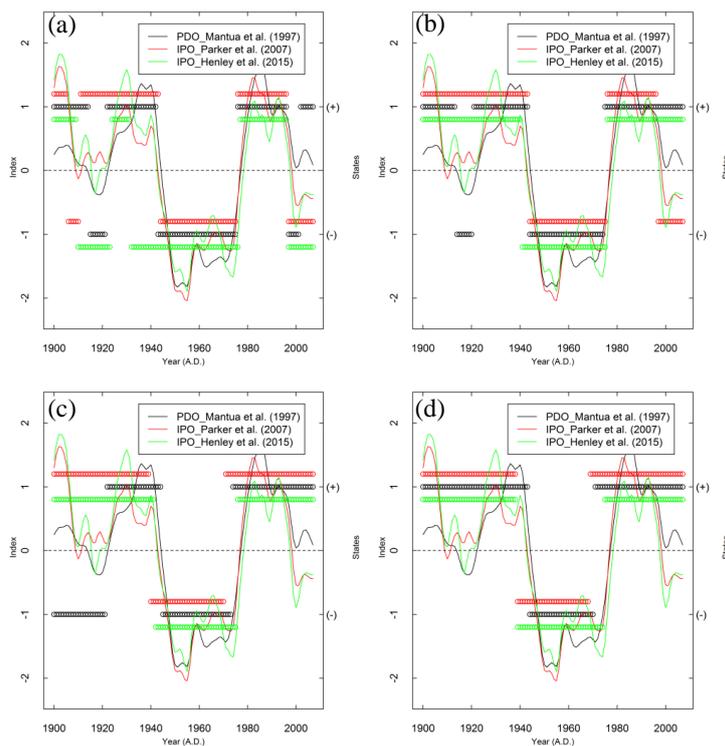
7

8

9

10

11



12

Figure 5 Time series plots of different instrumental PDV indices and corresponding run lengths

13

(represented by dots in corresponding colour) extracted using different window: (a) 10; (b) 20; (c) 30; (d)

14

40 years

15

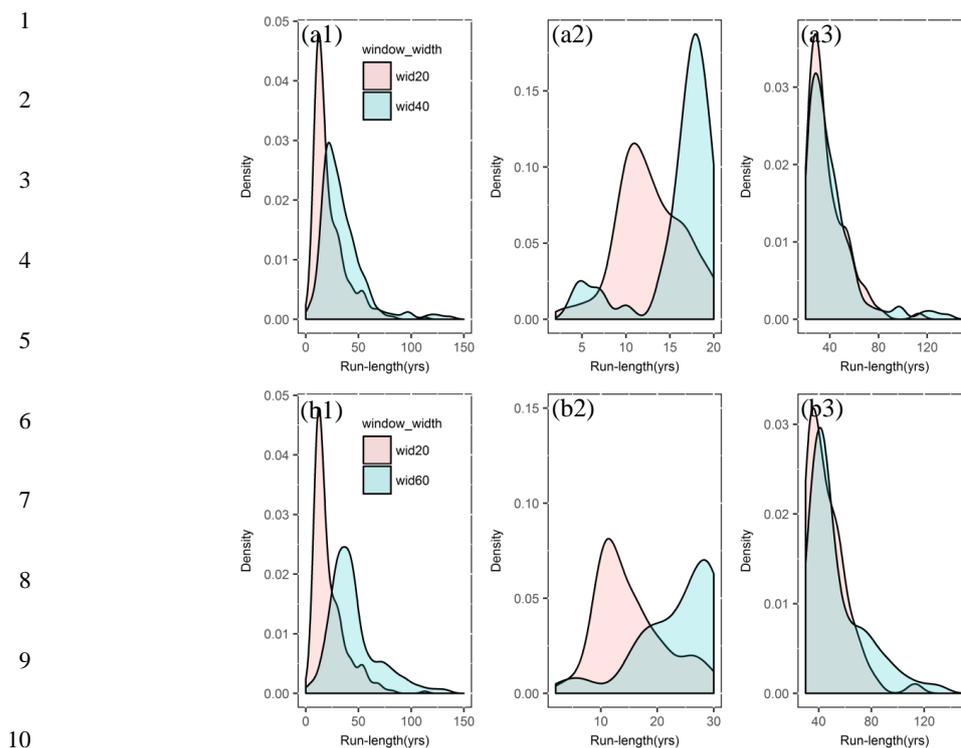
16

17

18

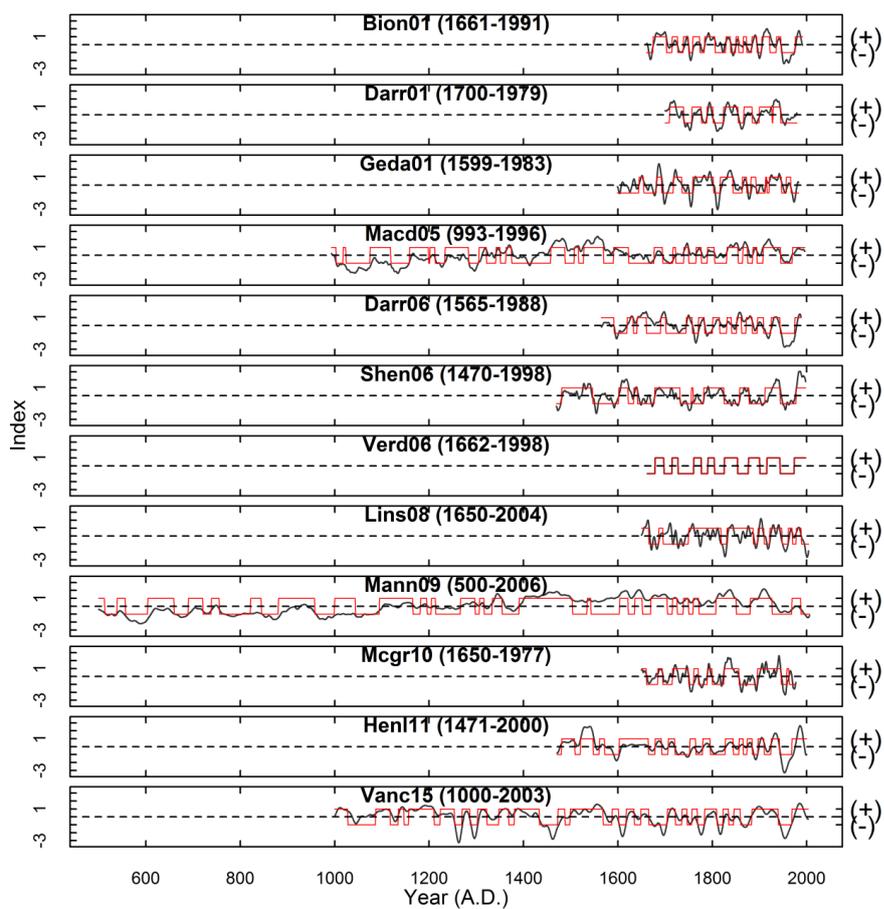
19

20



11 **Figure 6** Density plot of run length samples by applying different window widths to all reconstructions,
12 **(a1-a3):** comparison between window width 20 and 40, in which **(a1)** is unconditional run length
13 **distribution, panel (a2)** is the distribution conditioned on run lengths less than 20 years and panel **(a3)** is
14 **the distribution conditioned on run lengths greater than 20 years; (b1-b3):** comparison between window
15 **width 20 and 60, in which (b1)** is unconditional run length distribution, panel **(b2)** is the distribution
16 **conditioned on run lengths less than 30 years and panel (b3)** is the distribution conditioned on run lengths
17 **greater than 30 years**

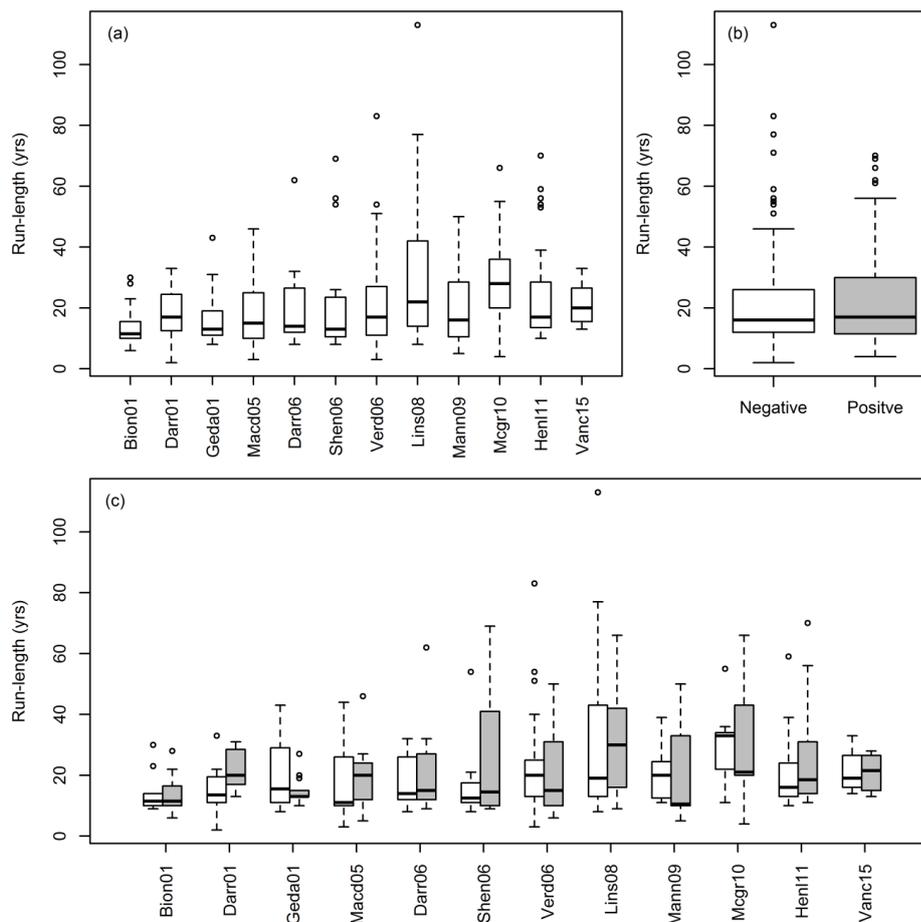
18



1

2 **Figure 7 Filtered PDV reconstruction time series (black line) with extracted run length (red line)**

3



1

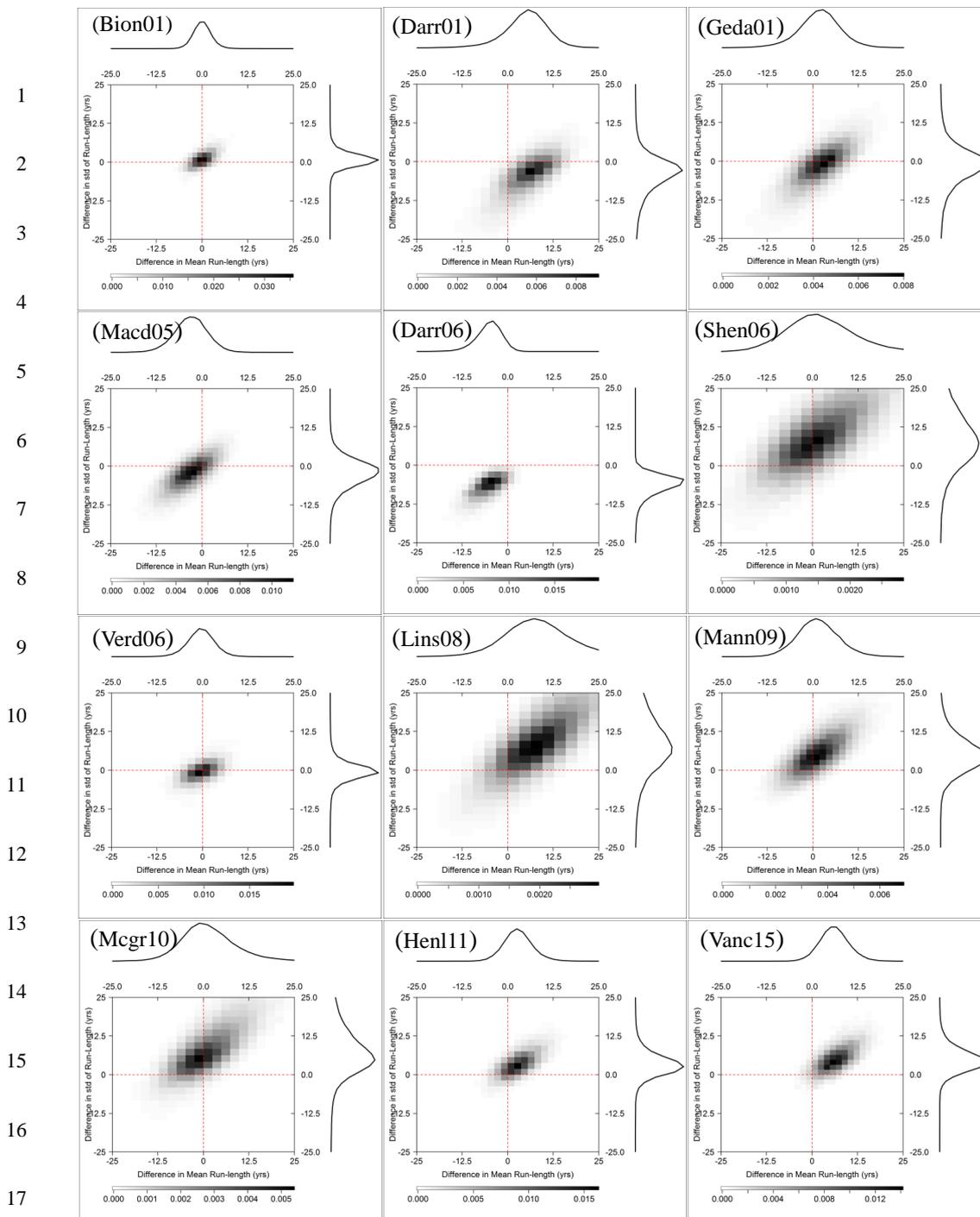
2 **Figure 8** Boxplot of run lengths samples from different reconstructions: (a) all run lengths for each
3 reconstruction; (b) pooled positive (grey) and negative (white) run lengths; (c) positive (grey) and
4 negative (white) run lengths for each reconstruction

5

6

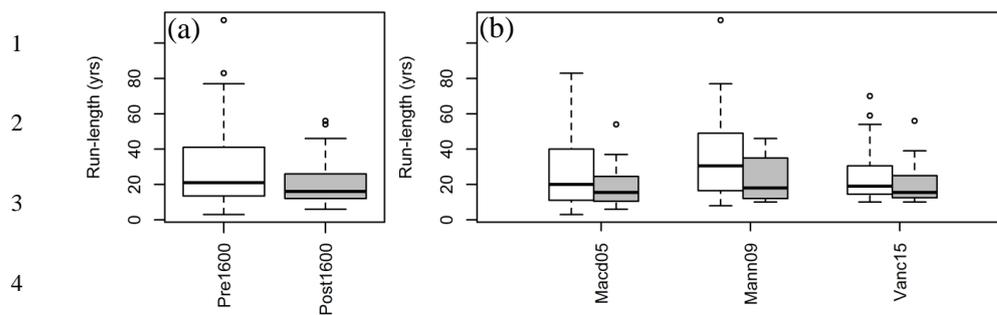
7

8



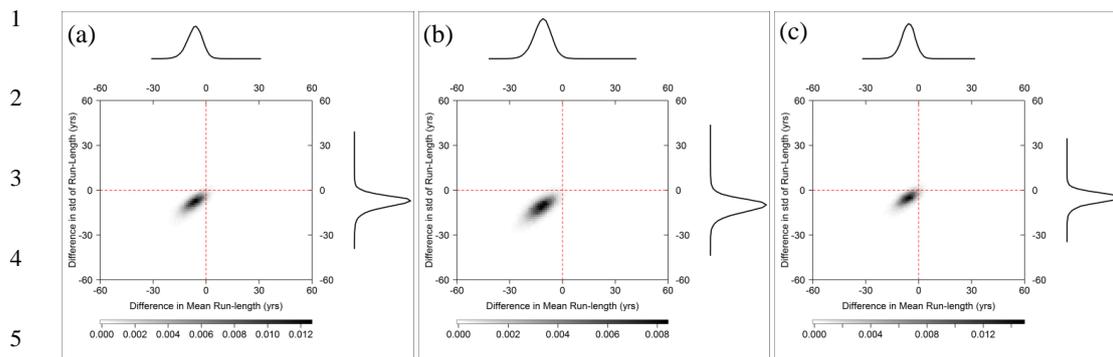
18 **Figure 9** Density and scatter plots of the differences between mean and standard deviation of positive and

19 **negative run length simulations (positive minus negative)**



5 **Figure 10** Boxplot of run lengths samples from different reconstructions (white box indicates pre-1600
6 and grey box indicates post-1600): (a) pooled run lengths from all three reconstructions; (b) run lengths
7 for each reconstruction

8



6 **Figure 11** Density and scatter plots of the differences between mean and standard deviation of pre-1600
7 and post-1600 run length simulations (post-1600 minus pre-1600): (a) Macd05; (b) Mann09; (c) Vanc15

8