Supplementary Information

1. Statistical methods

Confidence interval for the mean pentad rainfall

A confidence interval for mean rainfall per pentad is calculated in order to define a period considered as the lower and upper limit of the wet season. A 95% confidence interval for the mean pentad rainfall amount ($\bar{x}_p$) is given as:

$$\bar{x}_p \pm \frac{z_{0.95} s_p}{\sqrt{n_p}} \quad (1)$$

where s_p represents the standard deviation of pentad rainfall amounts, n_p is the number of observations for that pentad, and $z_{0.95} = 1.96$ is the standard normal z statistic that corresponds with a 95% confidence level. The confidence interval is calculated for each of the 73 non-overlapping pentads across the calendar year.

Fitting statistical distributions

Wet/dry spell lengths may be characterised through statistical distributions. The geometric, compound geometric, generalised Pareto, logarithmic and Polya distributions have been used for such purposes (Chowdhury and Beecham, 2013). To estimate the distribution that best characterises wet/dry spells, distributional parameters must be estimated. Parameter estimates are obtained by either employing the 'method of moments' or the 'maximum likelihood method'. Probability mass functions and parameters considered in our study are presented in Table S2. Geometric and Compound geometric parameter estimates are obtained using the method of moments. Such moments, including the mean, variance, skewness and kurtosis, are estimated from identified dry/wet spells using the formulae in Table S2; where n_l = the number of dry/wet spells and l_t = the length of dry/wet spells. The goodness of fit of the fitted distributions is confirmed using the Chi-square goodness of fit test.

Pearson's Goodness of fit test

The Pearson's goodness of fit test is used to ascertain whether a set of observations has an assumed probability distribution (null hypothesis), as opposed to the data not having been drawn from an assumed distribution. The test statistic is given by:

$$\chi^2 = \sum_{l=1}^{L} \frac{(O_l - E_l)^2}{E_l^2} \quad (2)$$

where O_l are the observed frequencies of wet/dry spells of length l, and E_l are the expected frequencies of spells of length l, given the assumption of a specified distribution. The distributions tested in this study are provided in Table S2.

2. Supplementary Tables

Table S1: Statistical distributions and parameters fitted on wet and dry spell lengths. The letters a, b, d, m and p are distribution parameters. Formulae for parameter estimates are given below using sample data spell lengths (l_t) combined for each year (t).

Distribution	**Probability mass function**	**Parameters**
Geometric	$f(x) = p(1-p)^{x-1}$	$p = 1/\mu$
Compound geometric	$f(0) = \frac{b}{a-1}$ $f(x) = \frac{(a-x-2)f(x-1)}{a+b+x-1}$	$a = (\mu-1)(b-1)$ $b = \frac{2s^2}{\mu(1-\mu)+s^2}$

Parameter	Sample Estimate
Mean	$\mu = \frac{1}{n_l}\sum_{t=1}^{n_l} l_t$
Variance	$s^2 = \frac{1}{n_l - 1}\sum_{t=1}^{n_l}(l_t - \mu)^2$

Pentad Summary

Table S2(a): Pentad profiles per decade at Maitland. Pentad long-term mean = 8.44mm and dry pentads have < 5mm of rain. Pentad 19-24 (1-30 April); Pentad 25-30 (1-30 May); Pentad 51-55 (8 September -2 October); and Pentad 56-61(3 October –1 November).

Decade	Lowest pentad above mean	Highest pentad above mean	N° of Pentads above mean	N° of dry pentads	Peak Pentad	Peak Pentad rainfall
1906 - 1909	19	60	28	9	44	35.0
1910 - 1919	19	58	32	8	28	33.0
1920 - 1929	24	56	23	13	32	24.1
1930 - 1939	26	56	21	13	44	15.4
1940 - 1949	19	52	26	10	30	20.4
1950 - 1959	20	60	33	5	22	23.9
1960 - 1969	21	58	26	10	32	27.1
1970 - 1979	24	59	29	10	33	26.5
1980 - 1989	20	56	28	11	40	23.9
1990 - 1999	19	59	34	5	38	38.3
2000 - 2009	20	59	31	7	38	23.8
2010 - 2018	19	61	28	13	42	25.5

Table S2(b): Pentad profiles per decade at Kirstenbosch. Pentad long-term mean = mm and dry pentads have < 5mm of rain. Pentad 19-24 (1-30 April); Pentad 25-30 (1-30 May); Pentad 51-55 (8 September -2 October); and Pentad 56-61(3 October –1 November).

Decade	Lowest pentad above mean	Highest pentad above mean	N° of Pentads above mean	N° of dry pentads	Peak Pentad	Peak Pentad rainfall
1915 - 1919	22	61	26	3	41	84.5
1920 - 1929	23	56	27	3	32	62.2
1930 - 1939	26	58	23	1	47	46.0
1940 - 1949	19	61	31	2	40	64.4
1950 - 1959	20	56	34	0	40	72.8
1960 - 1969	21	58	28	1	32	70.3
1970 - 1979	20	59	30	1	32	65.3
1980 - 1989	23	56	25	2	27	73.0
1990 - 1999	22	59	30	2	40	79.2
2000 - 2009	20	54	27	0	49	49.7
2010 - 2018	22	54	29	3	30	58.2

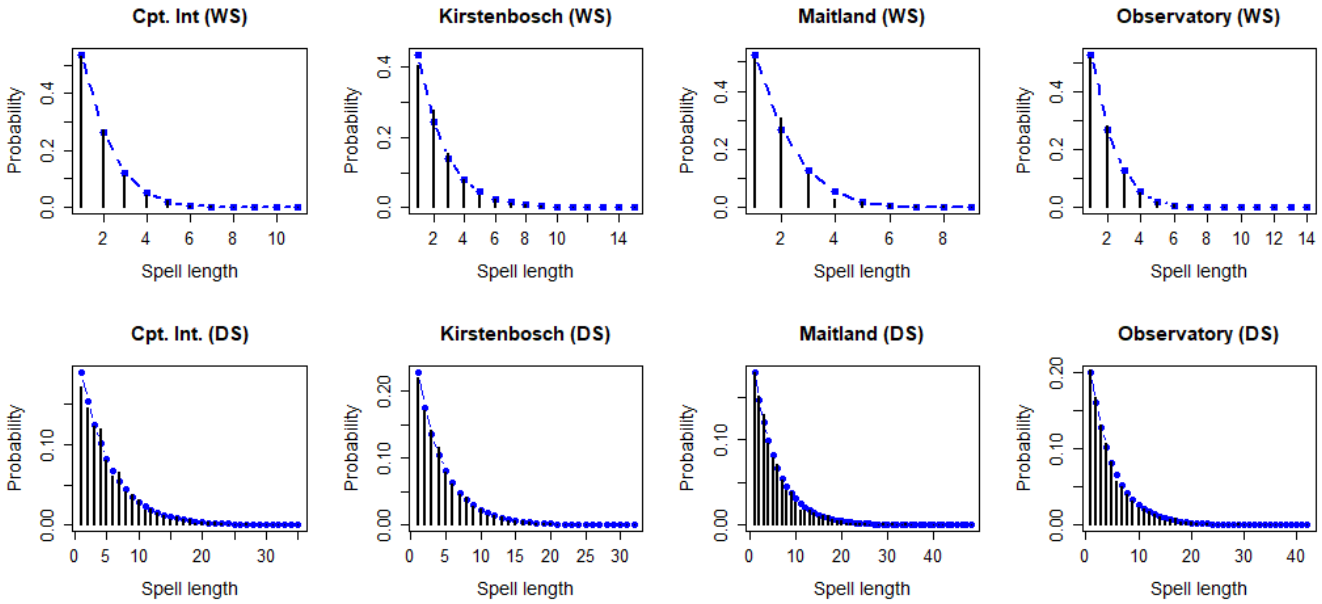
Table S2(c): Pentad profiles per decade at Cape Town Int. Pentad long-term mean = mm and dry pentads have < 5mm of rain. Pentad 19-24 (1-30 April); Pentad 25-30 (1-30 May); Pentad 51-55 (8 September -2 October); and Pentad 56-61(3 October –1 November).

Decade	Lowest pentad above mean	Highest pentad above mean	N° of Pentads above mean	N° of dry pentads	Peak Pentad	Peak Pentad rainfall
1950 - 1959	20	60	34	6	36	29.4
1960 - 1969	21	58	28	7	28	30.9
1970 - 1979	24	51	25	5	43	23.1
1980 - 1989	20	59	29	12	36	22.1
1990 - 1999	19	60	33	5	35	25.7
2000 - 2009	20	54	28	7	44	21.7
2010 - 2018	25	54	24	12	31	23.0

3. Supplementary Figures

Figure S1: Statistical probability distributions of wet and dry spells (during the wet season). Blue solid line represents the observed probabilities and the red dotted line represents the fitted distribution (using geometric and compound geometric distributions). The cumulative distribution functions for wet and dry spells with blue dashed lines indicating the 95th percentile.

(a) Probability distributions



(b) Cumulative distribution functions

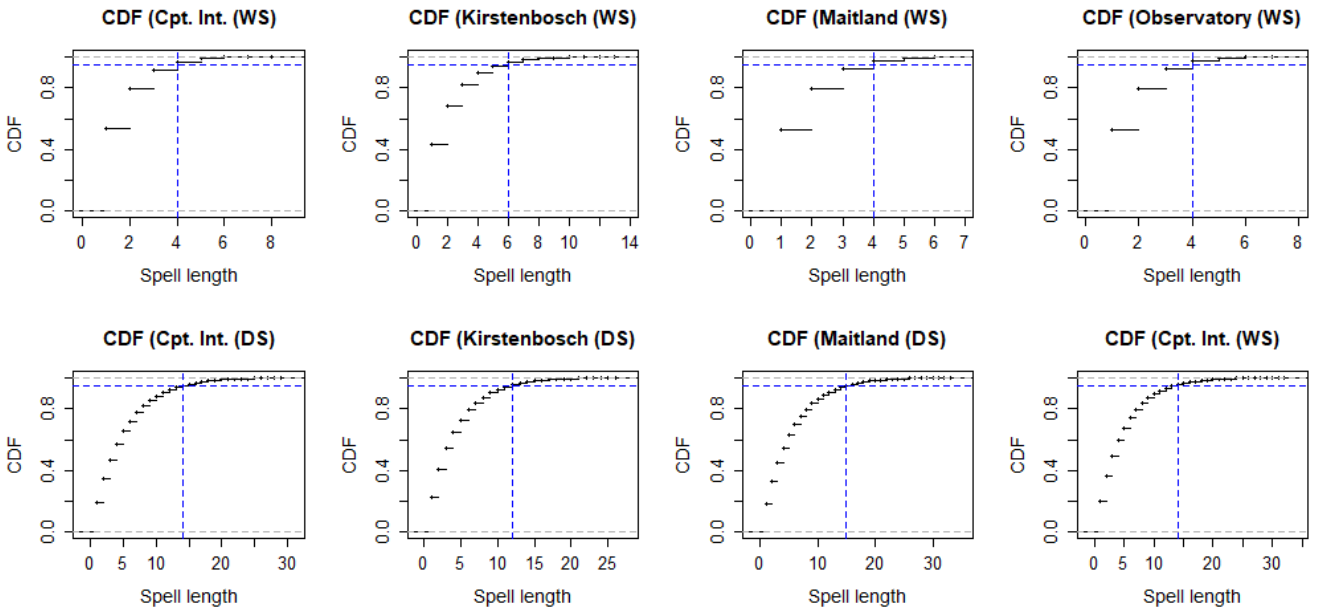
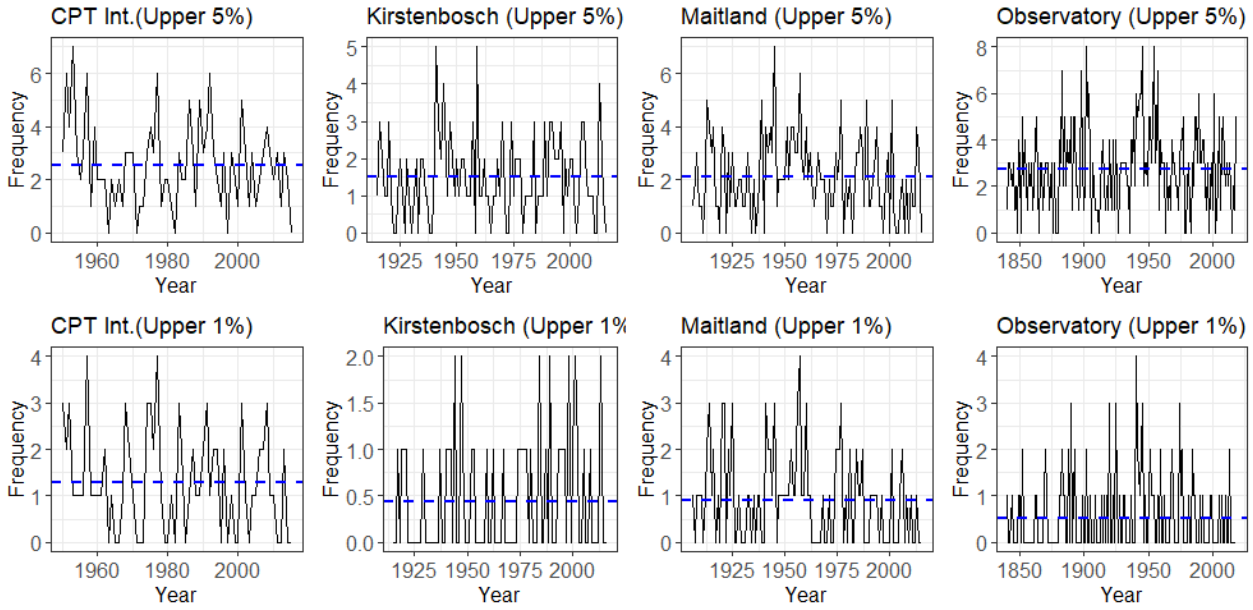


Figure S2: Frequencies of extreme (95th and 99th percentile) for wet and dry spells (during the wet season). Blue dashed lines represent the mean number (frequency) of extreme wet spells over the full record at that station.

(a) Upper 5% (95th percentile) and Upper 1% (99th percentile) for wet spells at each station.



(b) Upper 5% (95th percentile) and Upper 1% (99th percentile) for dry spells at each station.

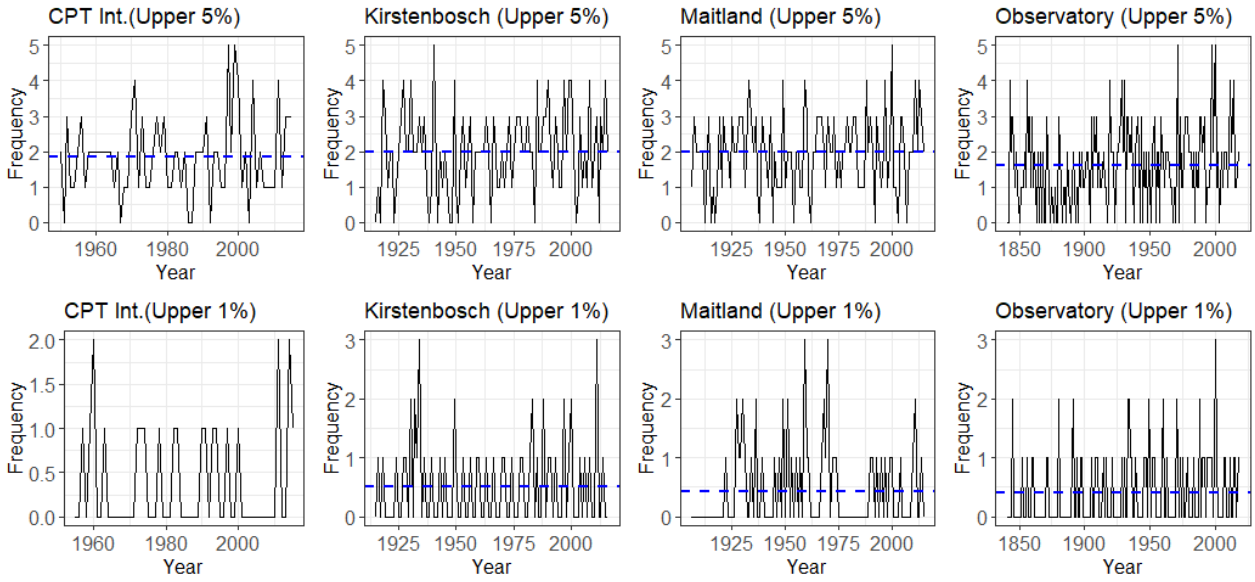
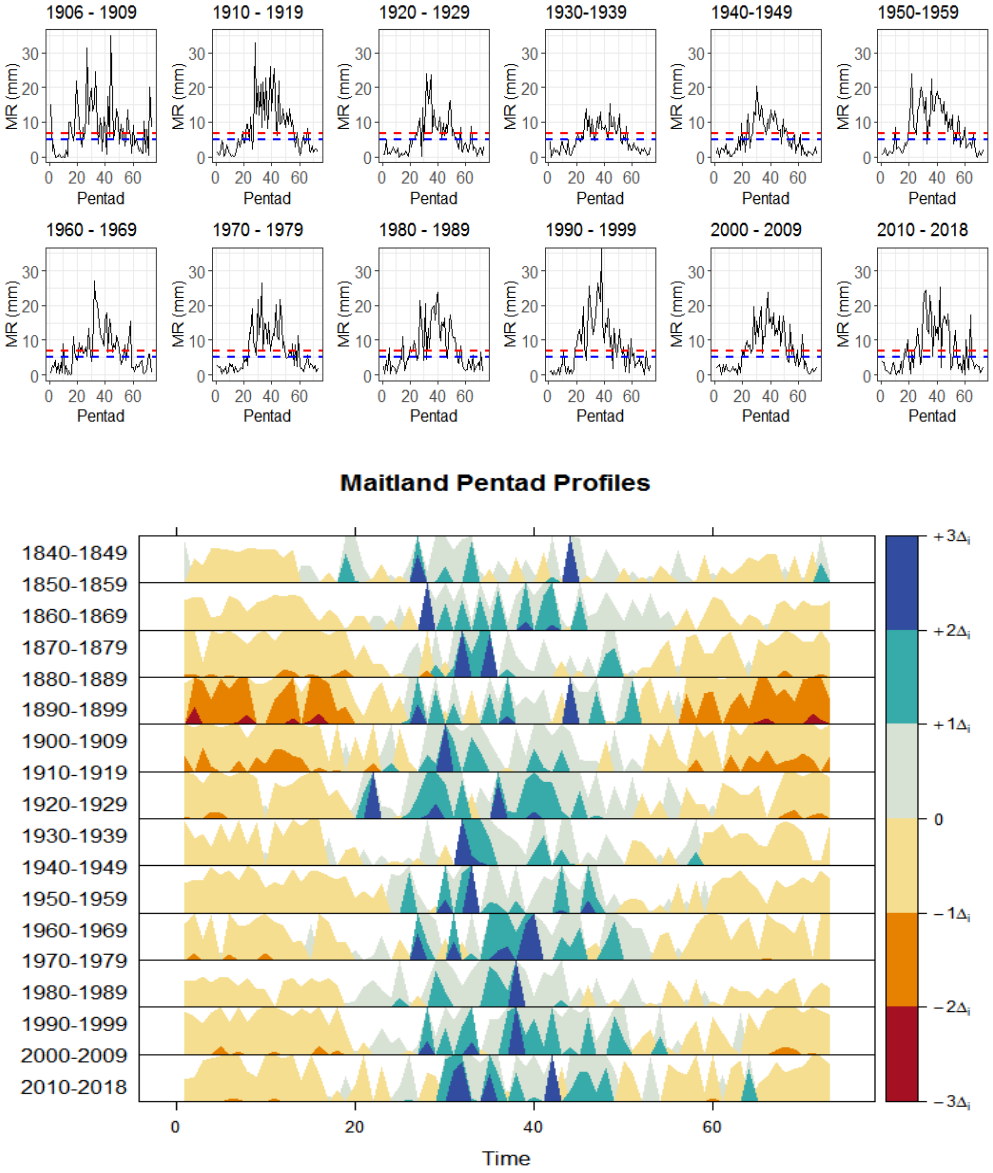
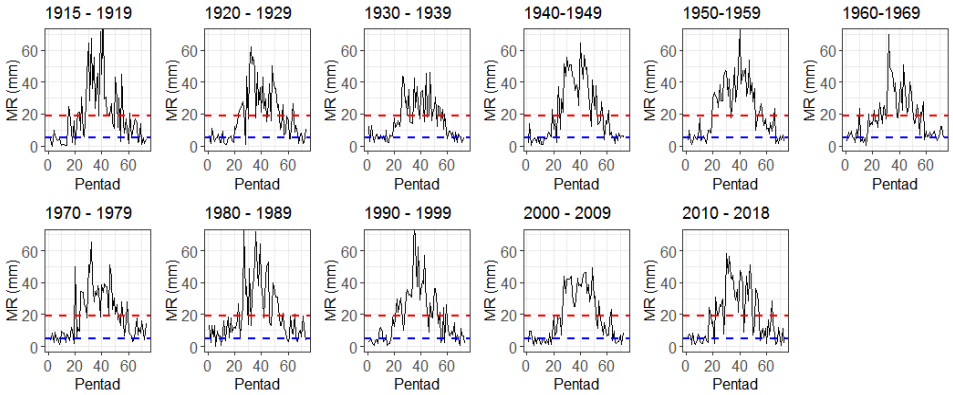


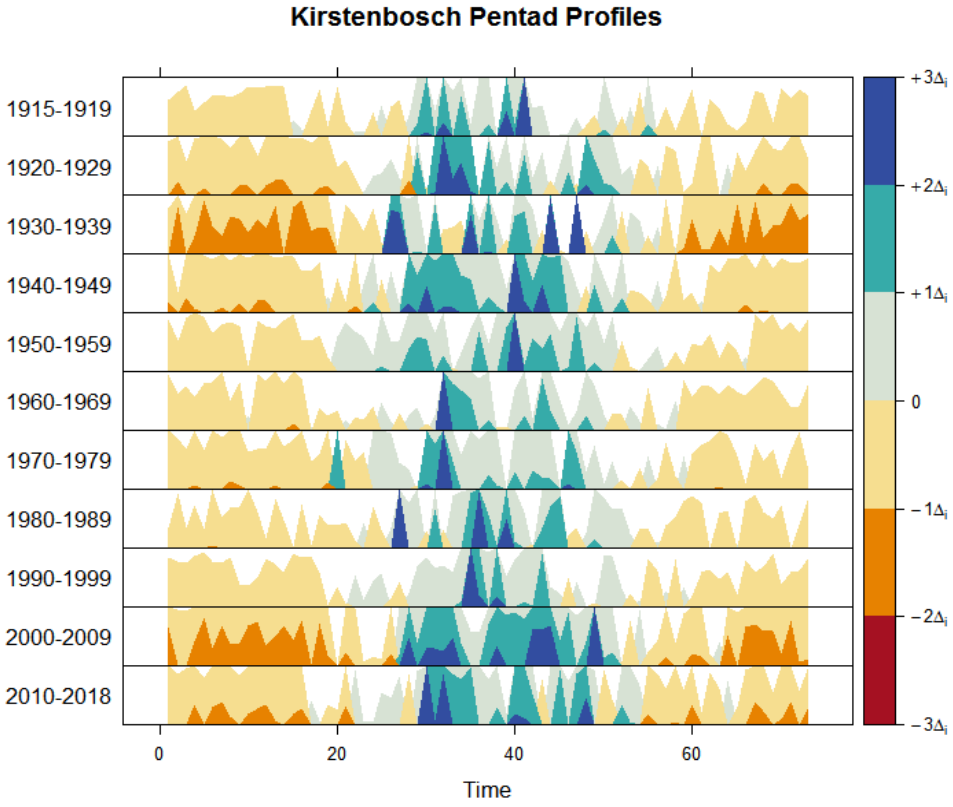
Figure S3: Pentad rainfall profiles for each station: (a) Maitland (b) Kirstenbosch and (c) Cape Town International, with mean rainfall (MR) pentad totals for each decade. Red dashed line = long-term pentad average ($\overline{D}$) using all pentads of the year and all years for the length of each station record. Blue dashed line = 5mm, which separates wet and dry pentads. The horizon plots are plotted for difference in the pentad mean relative to the overall long-term pentad mean for the series. Positive values (blue) represent pentad mean values above the overall long-term pentad mean.

(a) Maitland Pentad profile

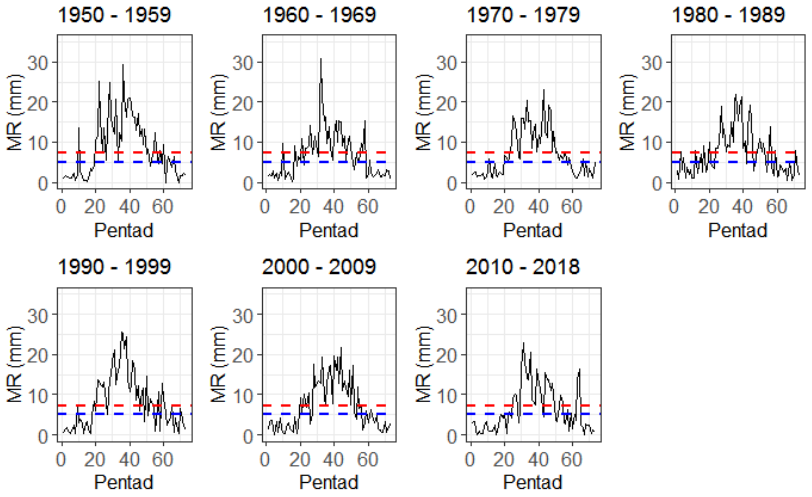


(b) Kirstenbosch Pentad profiles





(c) Cape Town International



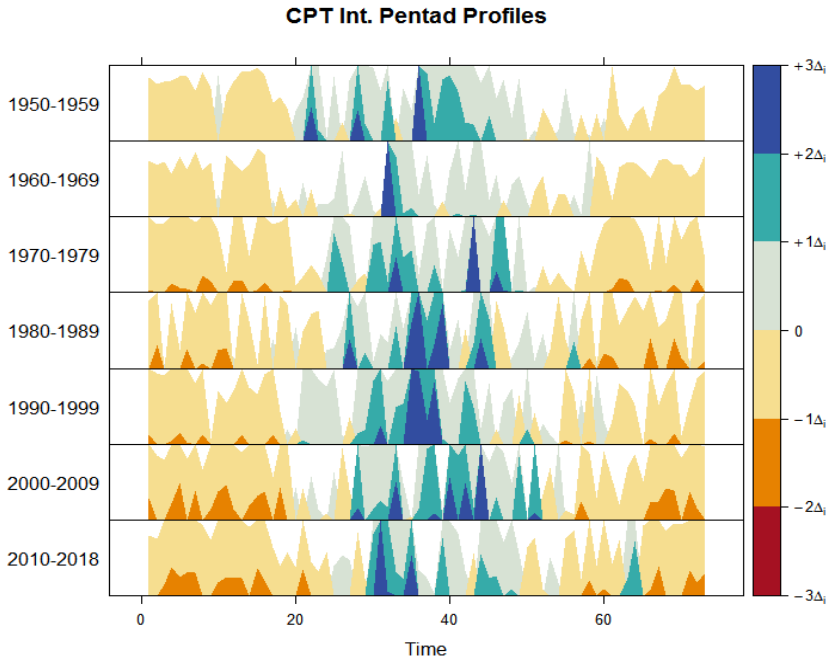


Figure S4: Correlation coefficient matrix for wet season onset, termination and length for the SAAO (O), Maitland (M) and Kirstenbosch (K) (1915-2018)