Australian Meteorological and Oceanographic Journal 62 (2012) 83–92

Defining a high-quality daily rainfall candidate network for Western Australia

Marco Marinelli[1 2], Karl Braganza[1], Dean Collins[1], David Jones[1],Shoni Maguire[1], Glenn Cook[2].

[1]National Climate Centre, Australian Bureau of Meteorology, Melbourne, 3001

[2]West Australian Regional Office, Australian Bureau of Meteorology, Perth, 6872.

(Manuscript received October 2011; revised April 2012)

A high-quality daily rainfall candidate network has been defined for Western Australia. The stations in this network have very high rainfall data completeness and have passed a number of statistical tests to assess data quality. This network will be used to supersede the Bureau of Meteorology's current high-quality daily rainfall network in Western Australia.

Introduction

Uncertainty exists in the accuracy of all meteorological datasets, and hence extensive research has been aimed at determining to what degree errors occur in temperature and precipitation datasets. Such analysis has resulted in the development of networks of high-quality stations, which are considered sufficient for use in relevant areas of research. This topic is extensively discussed in Easterling and Peterson (1995); Peterson and Easterling (1994); Peterson et al. (1997, 1998); Torok and Nicholls (1996) and Lavery et al. (1992, 1997). As discussed in Peterson et al. (1998), the methods applied range from empirical techniques, such as comparing plots of annual rainfall totals, to analytical statistical methods. The statistical techniques range from the traditional methods of comparison, such as covariance and correlation, to complex statistics, which determine when errors or discontinuities might be present in a dataset using point detection techniques. Most recent work has been centred on temperature datasets using inter-station comparisons (e.g. Easterling and Peterson (1995); Peterson et al. (1998); Durre et al. (2008); Reeves et al. (2007) and Menne and Williams (2009)).

As discussed by Lavery et al. (1992), 'there are several non-climatic factors which may lead to inhomogeneities in rainfall records' such as changes in observing practices, exposure of the rain gauge, changes in the station's location, and type of gauge used. The methodology applied by Lavery et al. (1992, 1997) in defining the current high-quality daily rainfall network for Australia started with the assumption that all observation stations have errors. The metadata was analysed and stations removed when entries in the metadata suggested a potential problem: for example, when a station changed its gauge type, or if a station was moved. However, it is possible these changes may not have influenced the data to any significant degree, so some rejected stations may actually have been of high quality (see Peterson et al. (1998) for further discussion on metadata). Lavery et al. (1992) did not consider the statistical comparison of rainfall in Australia to be a valid test due to the low density of stations, and thus did little statistical testing on rainfall data; however, more recent research suggests statistical methods are applicable, providing that the reference stations have a sufficiently high climatological consistency with the station of interest. In tests used to compare stations, the normally high variability of the measurements must be minimised. Therefore, in most methods of analysis, the effect of the high variability in daily measurements is minimised by comparison of yearly or seasonal totals (e.g. Reeves et al. (2007), Menne and Williams (2009)).

This paper outlines the methods applied to define a new high-quality daily rainfall candidate network for Western Australia, with the aim of testing the quality of the existing high-quality network and adding more stations to it. Of particular interest is the north of the State where there are only two precipitation stations in the current high-quality daily rainfall network (Lavery et al. 1992). There is a significant difference in the methodology that has been applied in this work from the methods applied by Lavery et al. (1992, 1997). Rather than concentrate on a laborious analysis of metadata entries, as in the work of Lavery et al., and to avoid any bias caused by inadequate metadata entries, a combination of empirical and analytical methods was applied to all daily rainfall observations up to 2008 from all Western Australian sites. As a result a new high-quality daily rainfall candidate network of 157 stations has been defined, a significant increase from the 40 high-quality stations identified by Lavery et al., and providing improved spatial coverage of the State. These stations have a very high completeness of daily

Corresponding author address: Marco Marinelli, National Climate Centre, Australian Bureau of Meteorology, Melbourne, Vic. 3001, Australia
Email: M.Marinelli@bom.gov.au

rainfall data and have a small level of gross error (determined by statistical and empirical tests). Most stations chosen in the high-quality daily precipitation candidate network have continuous records from a single site, and these will be referred to as single stations in the following sections. To supplement these single stations, twelve composite stations were created: (a) in a large geographical area where a station with a sufficiently complete record was not found or (b) the station passed all other tests but was closed earlier than 2000.

Methodologies for quality testing

All of the Bureau of Meteorology's Western Australian precipitation stations were investigated in two steps.

Step 1. Initial selection of candidate high-quality stations by measurement completeness

Initially, analysis of the completeness of the rainfall measurements was performed with the aim of determining which stations had suitable levels of data completeness in all years from 1950 or earlier, until the end of 2008. High quality network stations must have a very high record completeness, so this was the logical first filtering test to apply. It was initially decided that an acceptably complete station is one with at least 95 per cent data completeness over the time period of interest. This equates to a maximum of 1077 missing daily rainfall observations over a minimum period of 59 years, or almost three years of daily data, for example, or a maximum of 1826 missing observations for a site with 100 years of record, or five missing years of daily data. This arbitrary threshold of 95 per cent completeness was chosen as a compromise between an acceptably low amount of missing data, whilst maintaining a useful number of candidate high-quality stations for further investigation. Further analysis, with lower thresholds of data completeness, showed that the 95 per cent value did not exclude stations that might just be of sufficient completeness from further analysis.

Of 3951 Western Australian rainfall stations open at the end of 2008, 341 passed the 95 per cent completeness check.

Secondly, the stations that did not pass this first test were further analysed to see if they were suitable to be part of a composite station comprising a maximum of three stations. To qualify a station must:

1. be within 30 km of one or two closed rainfall sites with minimum overlap periods of six years;
2. have a strong correlation in the overlap period/s with the other member/s of the composite. To test this, the monthly and yearly total precipitation of the station were compared and judged acceptable where correlation was greater than 0.95. Thus, no adjustment of the data was considered necessary and the composites are simply a concatenation of records; and
3. meet the 95 per cent completeness rule described in Step 1 when its data are combined with that of the other members of the composite.

The composite stations are shown in Fig. 1 and detailed in Table 1. The join date was chosen so that the composite had the highest completeness of data during the overlap period, and denotes the transition to the data for the site with the later rainfall record.

None of the composites have a join date earlier than approximately 1930. This is because if a station has a high completeness from that time, its record was considered adequately long (~70–80 years).

To summarise, the stations with very high data completeness fall into two sets: (1) single stations—opened 1950 or earlier and still open at the end of 2008; and (2) composite stations—made up of a maximum of three stations. This analysis step reduced the initial 3951 analysed stations to 363 candidate stations, twenty-two of which are composites.

Step 2. Analysis of the 363 candidate stations

The second step consists of four tests applied to each candidate station: (1) temporal trend analysis and consistency with nearest stations; (2) correlation, regression and covariance with nearest stations; (3) test for the frequency of multi-day periods of missing data; and (4) analysis of rounding errors to identify potential biases in observing or data entry practices. These tests are further detailed in the following sections. These tests sequentially removed stations from the candidate station list.

Fig. 1 Location of composite stations. Locations of each of the composite components are plotted, however due to the scale of the map and the closeness of the sites, for some composite stations only one site may be visible. Colour shows topography.

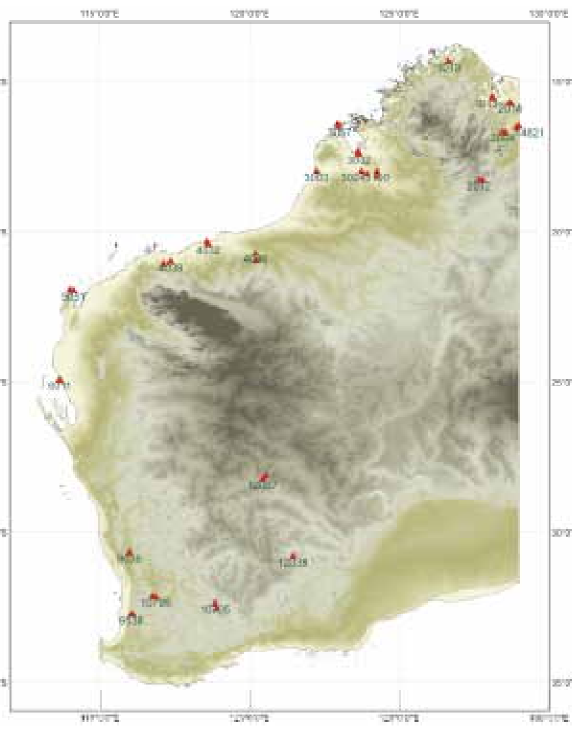


Table 1. Composite stations. This table contains the name and number given to each composite station, the stations in these composites, dates when the records from these stations are joined and relative position to the final station in the composite.

Composite station number	*Composite station name*	*Stations in composite*	*Join date*	*Approximate Distance to current site (km)*	*Approximate direction to current site*	*Station height above mean sea level (m)*
1013	Wyndham	1005				20
		1006	1 Nov 1967	7.4	SE	3.8
		1013	1 April 1968	3.7	SE	11
1019	Kalumburu	1021				23
		1019	1 Jan 2001	0.2	N	23
2012	Halls Creek Airport	2011				360
		2012	1 Jan 1944	12.7	NW	422
2014	Kimberley Research Station	2013				40
		2014	1 Jan 1946	4.8	NE	31
		2013	1 Jan 1959	4.8	NE	40
		2014	1 Jan 1969	4.8	NE	31
2064	Argyle Aerodrome	2016				120
		2064	1 Jan 1985	11.6	WNW	164
3003	Broome Airport	3002				19
		3003	1 Sep 1940	1.6	SW	7.4
3024	Udialla	3016				40
		3024	1 March 1984	22.5	N/W	20
3032	Derby Aerodrome	3007				8
		3032	1 March 1951	8.1	SS/E	6.2
3057	Cygnet Bay	3004				25
		3057	1 Jan 1964	10.6	SE	7
3100	Liveringa Station	3018				7
		3100	1 Jan 2001	19.2	N	55
4032	Port Hedland Airport	4002				8
		4032	1 Aug 1942	8.9	NE	6.4
4039	Warambie	4010				200
		4039	1 Jan 1937	27	NNE	30
4046	Yarrie	4004				250
		4046	1 Jan 1938	28.1	N	100
5051	Exmouth Town	5031				6
		5051	1 Jan 1968	12.9	ESE	13
6011	Carnarvon Airport	6062				5
		6011	1 Feb 1945	2.4	SW	4
8038	Moora West	8091				203
		8038	1 Jan 2003	1.8	W	205
9538	Dwellingup	9702				329
		9538	1 Feb 1934	4.0	W	267
10705	Hyden North	10569				317
		10705	1 Jan 1947	19.0	N	340
10795	Avondale Farm	10642				226
		10795	1 Jan 1965	11.3	SW	200
12038	Kalgoorlie-Boulder Airport	12039				361
		12038	1 Jan 1940	4.1	SSW	365
12067	Pinnacles	12158				499
		12067	1 Jan 1926	16.8	SW	400
14821	Rosewood	2001				120
		14821	1 Jan 1990	10.7	SW	137

For tests one and two, yearly total and seasonal total rainfall were compared. This minimised the effects caused by the naturally high variability present in daily and monthly rainfall measurements. These totals were only calculated for years where every month of the year had more than 25 days of data (the '25-day test'). This reduced the biasing of results that would be caused by large intra-annual periods of missing data.

A high-quality dataset must have a high continuity of measurements. Therefore, if more than six consecutive years failed the 25-day test, the station was excluded from the final candidate station list. 32 stations failed that test. However, if their measurements were mostly complete across the period of this station's life (70–80 per cent of years passed the 25-day test and the temporal distribution of these years are relatively even), the excluded station was still used as a 'reference station' in the assessment of the quality of nearby high-quality candidate stations. An even distribution of the missing data across the life of a reference station was considered necessary to allow it to be used without correcting for bias.

If there were five or six consecutive 'failed' years in a station, it would be included for further analysis: (a) if it was the only station in a data sparse area of the State and hence could not be substituted; and (b) it passed all other tests.

The information from this part of the analysis was also used to determine the first year that could be classified as 'trusted' in a station. For example, if a station was opened in 1900 and passed all the tests, but has seven consecutive failed years prior to 1907, then the first trusted year is 1907. At most stations, it was found that this year was the same, or close to, the year where daily rainfall recordings commence. For example, this frequently occurs for Western Australia stations as many daily rainfall measurements before 1907 are yet to be entered into the Bureau of Meteorology's database.

Test 2.1. Temporal Trend Analysis

The one year total and sliding fifteen-year mean was calculated for each station and the multiple linear regression calculated for time verses both of these rainfall values. This test is statistically simple, but can be prone to biasing caused by missing data, or the length of record of the station. For example, in two nearby stations, one may have been operational from 1900–2008 and the other from 1950–2008. If there is a significant change in the annual rainfall pre- and post-1950, then the regression coefficients may significantly disagree. So, when comparing the change in trend over time between stations, the time when the station was in operation was noted as well as when sufficiently incomplete (excluded) years occurred.

These regression coefficients, for annual and seasonal totals (all months in a season in a year), closely agreed with the temporal climate change grids available from the Bureau of Meteorology (http://www.bom.gov.au/cgi-bin/climate/change/trendmaps.cgi). The comparison of each candidate station's regression was also compared to the other nearest candidate stations. From both comparisons, it was clear that the candidate stations did not have significantly different regressions, which suggests no significant errors in the measurements. In order to further refine, as well as double check this analysis, further statistical analysis was performed.

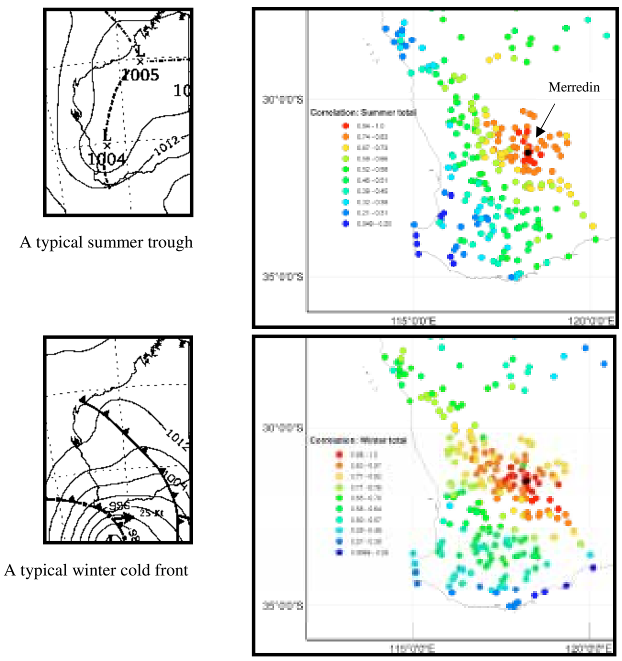
Test 2.2. Statistical Comparison with Nearest Stations

Correlation analysis, covariance, regression and intercept tests were applied to the candidate stations, with each station compared to all other candidate stations within a 7.5 degree grid. Annual and seasonal totals were analysed and results mapped to allow an easy comparison with nearest stations as well as to see how the spatial statistical patterns reflected the climatology of a region. This latter point, when comparing stations across a large area, was considered important due to the difference in rainfall patterns. For example, there should be a high correlation between stations which receive most of their rainfall from winter fronts crossing from the Indian Ocean (such as Cape Leeuwin and Cape Naturaliste). At these, most of the rainfall comes from southwest to northwest depending on the northerly extent of the front. However, for stations east of Albany, such as Cape Riche, most of the rainfall comes from the southeast (from the onshore flow in the wake of the cold front).

Correlation Analysis

Figure 2 shows the correlation of Merredin summer and winter total rainfall at Merredin with all other candidate stations in the southwest corner of Western Australia.

Fig. 2 Correlation between Merredin with southwestern stations: (a) summer and (b) winter periods. Summer rainfall in this inland area is mostly due to the summer trough. In winter it is due to the winter cold fronts.



Merredin was chosen as a good example to describe due to its different seasonal rainfall patterns. In winter, the spatial distribution reflects the passage of the winter fronts and so the stations in higher agreement have a clear northwest/southeast relationship. In summer, the pattern appears more circular, which reflects the summer trough. In the southern half of the State, a station which is in less agreement (and hence lower correlation) was easily identified, as the correlation of a reference station with nearby stations less then ~0.9 occurred infrequently. The quality of such stations was the most questionable and so were the easiest to flag as possibly being of low quality.

Covariance, regression and intercept: The regression and intercept results were also mapped and used in the inter-station agreement analysis. As expected, precipitation at stations in the southwest of the State are in closer agreement (due to the closeness of stations and the lower variability in rainfall during winter). The linear regression and intercept of the nearest stations reflected these values. For example, as shown in Fig. 3, the winter total rainfall of Merredin and Merredin Research station are in close agreement, with a correlation of 0.92.

For every station this analysis was repeated with its nearest stations. For example, both Merredin stations were compared with Doodlakine which is 35 km from Merredin. The correlation is high in both cases (Doodlakine–Merredin 0.87, Doodlakine – Merredin Research Station 0.91, see Fig. 3).

In the north of the State, the stations are expected to be in lower agreement due to: (a) the stations being more distant; and (b) the variability in the rainfall being higher. However, the correlation between summer totals ranged from ~0.8 to ~0.9 for stations less than 50 km apart, and ~0.7 to ~0.8 for stations between 50 and 100 km apart. Given the high variability and distance factors mentioned, this level of correlation was considered a good indicator of agreement.

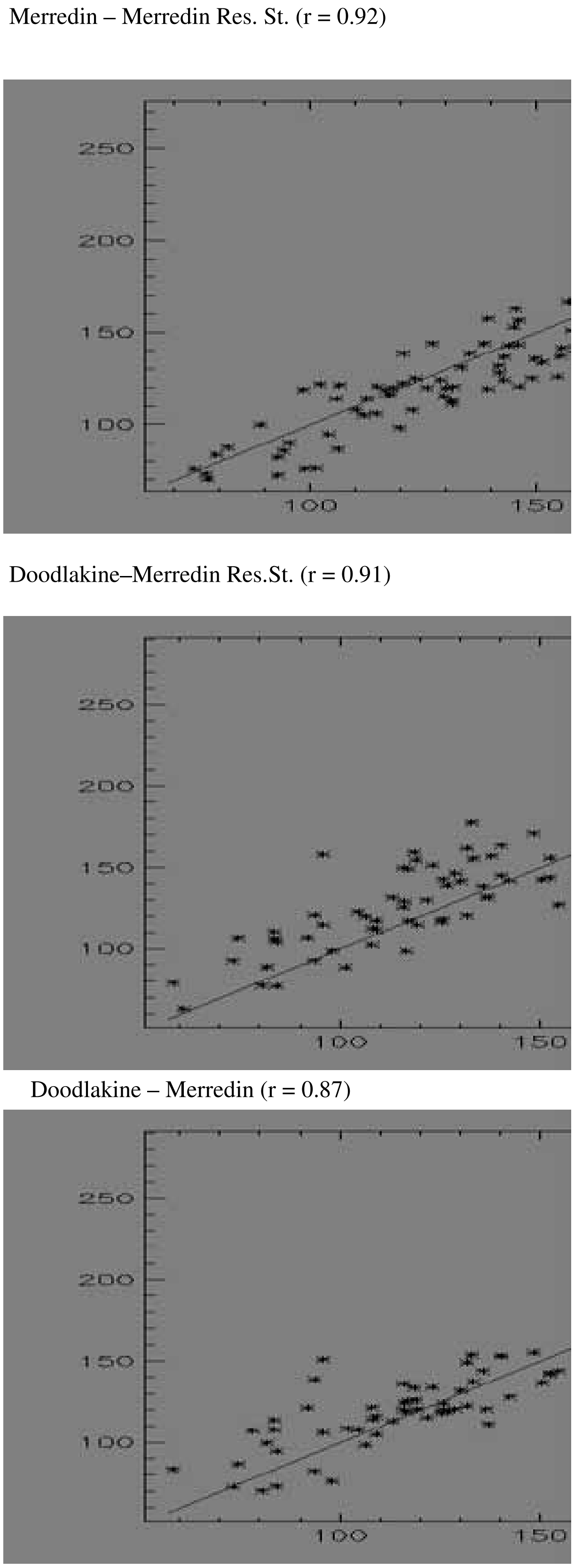
Based only on the statistical comparison results of steps (1) and (2), fourteen stations were rejected.

Test 2.3. Test stations for frequent multi-day periods of missing data

As discussed earlier, the first data completeness test removed stations with more than six consecutive years that failed the 25-day test. This section discusses the results of the second data completeness test, applied to the 317 remaining stations.

Firstly, the frequency of occurrence of small periods of missing measurements, in the remaining candidate stations (missing entries of less than seven days), was investigated. These small periods might be due to when many days of rainfall had been entered as one entry (an aggregation) or when no entry was made. In both cases this occurs to varying degrees at almost all stations. The exceptions to this, in the stations that passed all tests discussed in this paper, are Broome and Carnarvon Airport composites, and Giles and Perth Airport. This is expected for Bureau of

Fig. 3 Comparison of winter total rainfall (mm). Doodlakine, Merredin and Merredin Research Station.



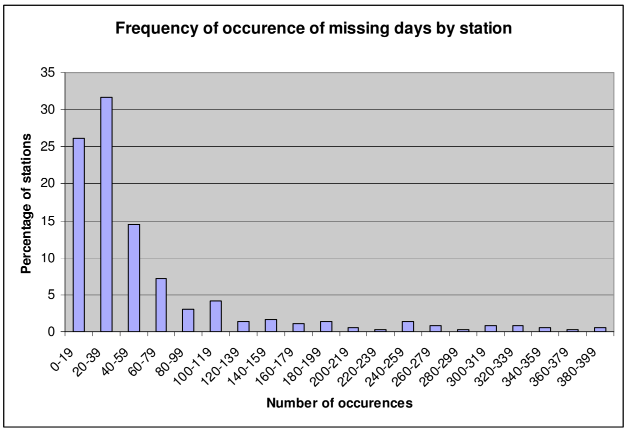
Meteorology airport stations as they have generally been staffed with professional weather observers and thus high-quality weather records are expected.

The frequency of occurrence of missing measurement periods, by percentage of stations, is shown in Fig. 4. 296 stations (81.5 per cent) had up to 100 occurrences of missing data. Note that the percentage of stations per group becomes relatively constant and small when the number of occurrences is ~160 or greater. Using this information, it was subjectively decided that stations with more than 160 occurrences of missing data would be excluded, totalling 41 stations.

Finally, the distribution of any failed years and the period of consecutive failed years in the 276 remaining stations were manually investigated. Of particular interest were stations with: (a) four–six concurrent years that failed the 25-day test but otherwise had very complete records; and/or (b) failed more than one year in approximately the last decade up to 2008. For example, four stations were found to have, from 2000 to 2008, many failed years in an otherwise very complete and long record. This information is of value as such stations are excluded at this point, but may be included in future high-quality datasets if the reason for the missing entries can be fixed; for example, by replacement with a recently opened station that passes the necessary criteria. As expected, the earlier tests had removed most stations which have significant missing entries, so the number which failed this step in the analysis is low (ten stations).

These completeness tests reduced the number of possible candidate stations across most regions of the State. With regard to one of the aims of this work—to increase the number of high-quality stations—this is not significant in the southwest of the State (where there are many candidate stations). However, in the more station-sparse inland and northern areas it is significant (especially in the Pilbara region). Therefore, future developments in statistical rainfall disaggregation and infilling techniques may be of greatest use in the station-sparse areas.

Fig. 4 Frequency of occurrence of missing days by station.



Test 2.4. Rounded measurements (Lavery et al. test)

The tests to this point reduced the number of possible candidate stations to 266. In this final step these stations were further analysed for rounding errors. This is important as a rounding error can reduce rainfall annual totals by four to five per cent (Lavery et al. 1992). Using the method described in Lavery et al. (1992), rounded measurements were identified by generating frequency histograms of the daily rainfall measurements. As expected, the majority of measurements should be low and not rounded to less than 0.2 mm or 1 point (the current decimal and pre-1973 imperial minimum scales of the Bureau of Meteorology rain gauges). If there is a bias toward a certain measurement, this should be clear in the histogram. For example, Fig. 5 shows the histograms of stations with unbiased (Merredin) and biased (Millston) measurements for the post-1973 period. In Millston there are significantly more whole number entries, clearly suggesting a rounding bias. 188 stations remained following this test.

Other excluded stations

In areas where there is a high density of stations the procedure was modified to reduce the number of stations on which a complete analysis was performed due to time constraints. As a result of this 31 stations passed step 1 but were not included as candidate stations, as their data completeness was either low pre-1920 or they opened after this period. Typically these stations are in the southwest and have a high statistical agreement with their nearest stations. Therefore, they are potentially high-quality stations, which future analysis work may determine.

To summarise, these tests reduced the high-quality candidate stations to a final list totalling 157 stations, twelve of which are composites (See Table 3 for details, mapped in Fig. 6).

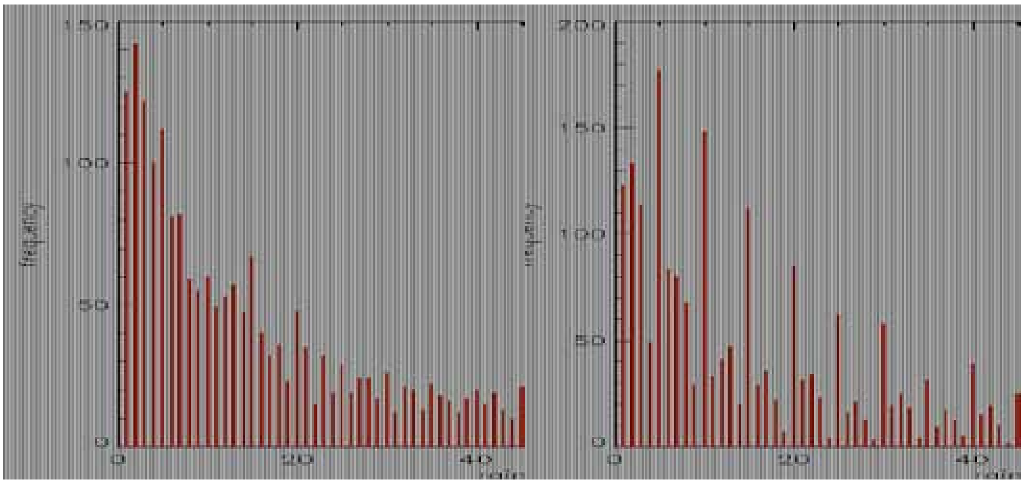
Lavery et al. stations excluded from the high-quality candidate network

There are 40 stations in the Lavery et al. High-Quality Daily Network in Western Australia. Thirteen of these have not been included in the high-quality candidate network. The most common reason for this is the incompleteness of their measurement entries, especially at stations in the northern half of Western Australia. The reason for the exclusion of these stations is written in Table 2.

Summary and discussion on future work

This work has defined the high-quality candidate precipitation network of 157 stations, twelve of which are newly created composites and 29 also in the Lavery et al. high-quality precipitation network. Recent work has increased the statistical tests which can be used to assess the quality of monthly and daily precipitation measurements, as

Fig. 5 Frequency histogram of daily rainfall (mm), 1974–2008. (a) Merredin (b) Millston.



(a) (b)

Fig. 6 High-quality candidate network.

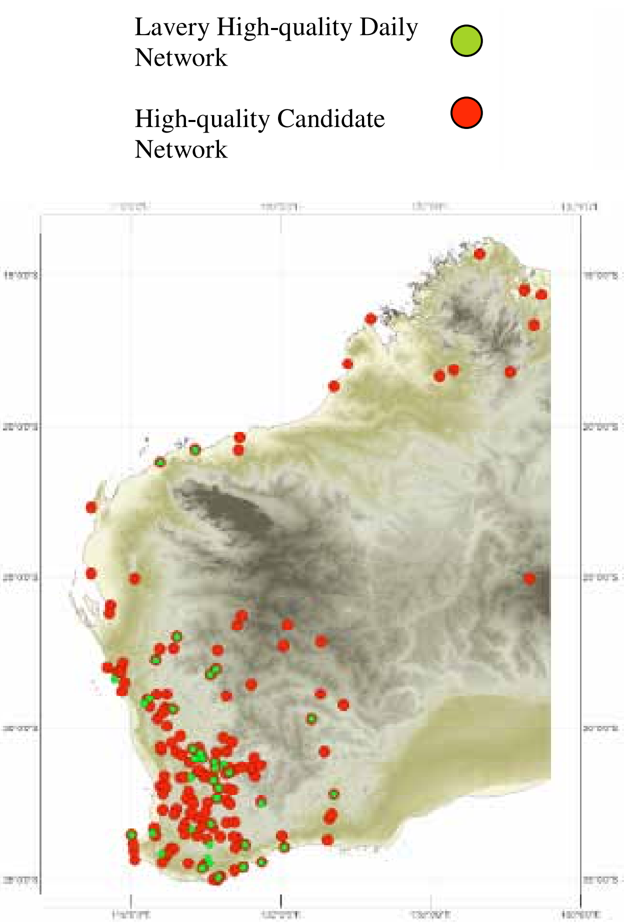


Table 2. Lavery et al. Daily High-quality Network stations that have been excluded from the High Quality Candidate precipitation network. The final column details the test failed, which for most stations was the 160 multi-day periods of missing data test (Test 2.3).

Station Number	*Station Name*	*Reason for exclusion*
8079	Manarra	Missing data in last decade
8088	Mingenew	> 160 missing entry periods.
8141	Willigulli North	Too many consecutive years failed the 25 day test
9503	Boyanup	> 160 missing entry periods.
9561	Kendenup	> 160 missing entry periods.
10045	Ejanding	Poor agreement with nearest stations post 1990
10091	Meckering	> 160 missing entry periods.
10112	Nungarin	> 160 missing entry periods and poor agreement with nearest stations.
10582	Kojonup	> 160 missing entry periods.
10505	Arthur River	> 160 missing entry periods.
10525	Broomehill	> 160 missing entry periods.
10537	Cranbrook	> 160 missing entry periods.
10611	Mount Madden	Failed Lavery et al. test.

well as to make statistically valid corrections, examples of which include work done by Wang et al. (2010) and Costa and Soares (2009). In these papers, tests were used to detect and correct discontinuity in temporal trends (change points). There is also growing interest in, and application of, applying infilling and disaggregation methods, as was done in the temperature record, to the more difficult-to-test (due to high variability) daily rainfall records. As these methods are developed, application of them will determine which of this network's stations will define the updated daily high-quality rainfall dataset network in Western Australia.

Table 3. The High-quality Candidate Network. Composite stations are labelled (C). Stations which are in the Lavery et al. high-quality network are italicised. The first year in which measurements were taken, and the first year with a high measurement of completeness, are tabled in the third and fourth column.

Number	*Name*	*First year where entry > 0*	*Chosen first high completeness (FHC) year*
1013	Wyndham (C)	1898	1900
1019	Kalumburu (C)	1941	1942
2012	Halls Creek Airport (C)	1894	1900
2014	Kimberley Res.Station (C)	1907	1907
2064	Argyle Aerodrome (C)	1907	1907
3003	Broome Airport (C)	1897	1900
3009	Jubilee Downs	1938	1938
3027	Fossil Downs	1909	1921
3030	Bidyadanga	1907	1907
3057	Cygnet Bay (C)	1918	1918
4016	Indee	1911	1911
4032	Port Hedland Airport (C)	1907	1911
4035	*Roebourne*	1907	1908
5008	*Mardie*	1907	1908
5020	Ningaloo	1903	1907
6011	Carnarvon Airport (C)	1907	1907
6027	Jimba Jimba	1916	1916
6048	Twin Peaks	1907	1923
6055	*Woolgorong*	1904	1907
6059	Yaringa Station	1936	1947
6060	Carbla Station	1913	1914
7007	*Boolardy*	1907	1907
7017	Cue	1907	1907
7045	Meekathara Airport	1944	1945
7057	*Mount Magnet*	1907	1907
7064	Murgoo	1907	1908
7067	Narndee	1921	1923
7095	*Yoweragabbie*	1908	1909
7101	Munarra	1921	1929
8004	Balline	1930	1931
8005	Barberton	1911	1912
8010	Binnu	1929	1930
8013	Bowgada	1907	1907
8025	Carnamah	1907	1927
8028	Nabawa	1907	1907
8038	Moora West (C)	1907	1907
8039	Dalwallinu Comparison	1912	1913
8051	Geraldton Airport	1941	1942
8060	South Holmwood	1921	1921
8066	*Kokardine*	1907	1908
8067	Koobabbie	1911	1911
8070	Lake Hinds	1924	1925
8085	Miling	1924	1931
8096	Mumby	1907	1907
8104	Ogilvie	1910	1911
8106	*Perangery*	1910	1913
8113	Riverside	1929	1930
8128	Tenindewa Bindu	1915	1916
8137	Wongan Hills	1907	1907
8143	Yandanooka	1907	1907
8157	Canna	1915	1916
8233	Five Gums	1945	1945
9009	Lower Chittering	1915	1916
9021	Perth Airport	1944	1945
9031	Mundaring Weir	1907	1907
9044	Wungong Dam	1913	1948
9500	Albany	1907	1907
9501	Arundel	1912	1913
9506	Bangalup	1920	1920
9507	Bannister	1907	1907
9518	Cape Leeuwin	1907	1907
9519	*Cape Naturaliste*	1907	1907
9520	*Cape Riche*	1907	1907
9527	Dardanup East	1935	1936
9534	Donnybrook	1907	1907
9538	Dwellingup (C)	1915	1916
9547	Forest Grove	1925	1926
9551	Grassmere	1907	1907
9556	Aldervale	1907	1925
9557	*Hopetoun*	1907	1907
9559	Kalgan River	1911	1913
9564	*King River*	1907	1907
9573	Manjimup	1916	1916
9575	Marradong	1907	1908
9587	Newbicup	1918	1919
9591	*Pardelup*	1907	1907
9592	Pemberton	1941	1942
9594	*Peppermint Grove*	1907	1907
9615	Warriup	1919	1920

Number	*Name*	*First year where entry > 0*	*Chosen first high completeness (FHC) year*
9619	*Wilgarrup*	1907	1907
9626	Peasant Valley	1948	1949
9636	Cowaramup	1926	1941
9657	Roelands	1942	1943
10007	Bencubbin	1925	1925
10011	Bonnie Rock	1936	1936
10012	Boodjerakine	1913	1914
10019	Burracoppin	1907	1907
10030	Wattoning	1907	1928
10032	*Cowcowing*	1907	1907
10037	*Cuttening*	1907	1907
10039	*Doodarding Well*	1907	1914
10041	Doongin Peak	1907	1908
10052	Aleppo	1910	1910
10061	Happy Valley	1937	1937
10073	Kellerberrin	1907	1907
10092	*Merredin*	1907	1914
10093	Merredin Research Station	1912	1912
10097	Moningarin	1910	1911
10108	Nokaning	1914	1915
10111	Northam	1907	1907
10123	*The Granites*	1910	1911
10124	Nangeenan	1907	1907
10126	*Trayning*	1910	1911
10133	*Walk Walkin*	1909	1909
10137	Wialki North	1929	1930
10143	Ygnattering	1928	1929
10149	*Codg Codgen*	1907	1909
10150	Grass Valley	1907	1908
10510	Barooga	1911	1911
10515	Beverley	1907	1907
10524	Brookton	1907	1909
10534	Colorado	1913	1914
10536	*Corrigin*	1910	1910
10541	*Nyerilup*	1910	1912
10542	Darkan	1907	1913
10546	Dumbleyung	1910	1910
10547	Duranillin	1910	1911
10561	Rushy Pool	1911	1911
10564	Hillcroft	1915	1915
10566	Horseshoe	1923	1924
10579	Katanning Comparison	1907	1907
10586	Kumminin	1910	1910
10589	Kwobrup	1915	1917
10592	Lake Grace Comparison	1912	1914
10595	Pingrup South	1935	1936
10600	Nanda Downs	1913	1914
10603	Hartwood	1913	1914
10612	Narembeen	1927	1928
10614	Narrogin	1907	1907
10622	Ongerup	1914	1915
10626	Pingelly	1907	1907
10633	Ravensthorpe	1907	1907
10636	*The Oaks*	1907	1907
10638	St Albans	1912	1914
10644	Bathurst	1913	1914
10647	Wagin	1907	1907
10654	Wickepin	1911	1912
10658	Wonnaminta	1907	1907
10670	*Lake Carmody*	1907	1935
12011	Bullfinch	1911	1912
12038	Kalgoorlie-Boulder Airport (C)	1907	1907
12046	Leonora	1907	1907
12052	*Menzies*	1907	1907
12054	Dobra Scritia	1908	1925
12056	Moree	1929	1930
12064	Noongaar	1926	1928
12065	*Norseman*	1907	1907
12071	Salmon Gums Research Station	1932	1934
12074	Southern Cross	1907	1907
12077	I Dunno	1926	1926
12083	Westonia	1915	1917
12090	Yeelirrie	1928	1929
12092	Yuinmery	1921	1922
12093	Yundamindra	1907	1907
12108	Wonganoo	1928	1929
13012	Wiluna	1907	1907
13017	Giles Meteorological Office	1956	1957

References

Costa, A.C. and Soares, A. 2009. Homogenization of Climate Data: Review and New Perspectives Using Geostatistics. *Mathematical Geosciences, 41*, 291–305.

Durre, I., Menne, M.J. and Vose, R.S. 2008. Strategies for Evaluating Quality Assurance Procedures. *Journal of Applied Meteorology and Climatology, 47*, 1785–1791.

Easterling, D.R. and Peterson, T.C. 1995. A new method of detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.,* 15, 369–337.

Lavery, B., Joung, G. and Nicholls, N. 1997. An Extended High-Quality Historical Rainfall Dataset for Australia, *Aust. Meteorol. Mag., 46,* 27–38.

Lavery, B., Kariko, A. and Nicholls, N. 1992. A Historical Rainfall Dataset for Australia, *Aust. Meteorol. Mag., 40,* 33–39.

Menne M.J. and Williams C.N. 2009. Homogenization of Temperature Series via Pairwise Comparisons, *J. Clim., 22,* 1700–1717.

Peterson, T.C. and Easterling, D.R. 1994. Creation of Homogeneous Composite Climatological Reference Series, *International Journal of Climatology, 94,* 1493–1517.

Peterson, T.C., Easterling, D.R., Kark, T.R., Groisman, P., Nicholls, N., Plummer N., Torok, S., Auer I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland E. J., Hansen-Bauerr, I., Alexanderson, H., Philip, J. and Parker, D. 1998. Homogeneity Adjustments of In Situ Atmospheric Climate Data: A Review, *Int. J. Climatol., 18,* 1493–1517.

Peterson, T.C. and Vose, R.S. 1997. An overview of the Global Historical Climatological Network temperature data set, *Bull. Am. Meteorol. Soc.,* 78, 2837–2849.

Reeves, J., Chen, J., Wang, X., Lund, R. and Lu, Q. (2007). A Review and Comparison of Changepoint Detection Techniques for Climate Data, *Journal of Applied Meteorology and Climatology, 46,* 900–915.

Torok, S.J. and Nicholls, N. 1996. A historical annual temperature dataset for Australia, *Aust. Meteorol. Mag., 45,* 251–260.

Wang, X. L., Chen, H., Wu, Y., Feng, Y. and Pu, Q. 2010. New Techniques for the Detection and Adjustment of Shifts in Daily Precipitation Data Series. *Journal of Applied Meteorology and Climatology,* 49, 2416–2436.