Is homogenisation of Australian temperature data any good?

Part 2. Marble Bar, the warmest place in Australia

Dr Bill Johnston[1]

scientist@bomwatch.com.au

Maximum temperature (Tmax) at Marble Bar shows no trend or change attributable to the climate. While instruments and observers changed and the site moved to locations that were cooler/warmer, the climate remained the same. Metadata (data about the data) failed to mention supply of a Stevenson screen in 1911, moving to the new post office on the corner of Francis and Contest streets in 1948, construction of the telephone exchange in 1961, the move to the LPO then across town to the Travellers Rest … and more. Metadata is demonstrably useless as a basis for adjusting for station moves and changes

Homogenising data to support the green agenda is unconscionable. The Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) project is unscientific and should be abandoned.

Read on …

Summary

Homogenisation of Australian temperature data by the Bureau of Meteorology (BoM) creates trends in homogenised data that are unrelated to the climate. While metadata used to identify changepoints is unreliable, datasets used to construct reference series selected on the basis of inter-site correlations likely embed parallel faults. Use of co-correlated data to adjust faults in target-site data has no scientific or statistical merit. Robust, replicable methods require changepoints to be identified statistically and for adjustments to be made objectively without bias.

Daily maximum temperature data (Tmax) for weather stations at Marble Bar, Western Australia were merged and analysed using www.bomwatch.com.au protocols outlined previously using Parafield as the case study. Inhomogeneities in 1913, 1945, 1988, 1999 and 2013 were cross-referenced to BoM metadata, a 1996 PhD thesis, maps, plans, photographs and files held by the National Archives of Australia, and reports in newspapers from the National Library of Australia's *Trove* database.

Metadata ignored that a Stevenson screen supplied in 1911 was probably not installed until 1912. Also, that the site moved to the new post office in Francis Street, which opened on 5 August 1948, and was extended in 1961/62 to house a telephone line exchange. Step-changes in 1913 (Stevenson screen), 1945 (probable staff change) and 1988 (move) could be explained but those in 1999 and 2013 could not. (While the site was inspected by the BoM in August 1999, a 60-litre screen was reportedly not installed until September 2000.) Additional analysis showed numbers of observations >95th day-of-year percentiles were less from 1966 to 1985, probably due to shading or watering, and markedly higher from 2015.

Historic data were not recorded to be used decades later to detect trend and change. Due to inaccurate metadata, poor site control (deterioration of the screen, watering, shade), and prior to metrication on 1 September 1992, lack of precision, Marble Bar data could not be regarded as high-quality. As the Bureau's homogenisation methods are deeply flawed, the ACORN-SAT project should be abandoned.

[1] Former NSW Department of Natural Resources research scientist and weather observer.

Introduction

> Maximum temperature data consists of an ambient background signal, which is assumed to be constant; variation due to rainfall, which is the physically causal or deterministic covariable, and unexplained residual variation reflecting the wider climate. Disaggregating the total signal into component parts is the key to determining unbiased climatic trend and change.
>
> Read on …

Measurement of climatic trend and change requires that long-term observations are not affected by extraneous factors such as station moves, instrument changes and changes in the local environment. Otherwise, site-related effects may be confused as being due to the climate. As temperature timeseries consist of at least two overlaid signals: one reflecting changes in the ambience of a site, the other, the weather, it is unlikely that trends in raw data are due to the climate alone.

Consequently, naïve trend analysis should not be used for data-shopping or to prove a point. For instance, Excel analysis of the combined dataset for Bourke, New South Wales, gives the impression that maximum temperature (Tmax) has cooled 0.08°C/decade since 1879. Also, that Tmax at Bridgetown, Western Australia, has warmed by 0.03°C/decade since 1907, and at Rutherglen, Victoria by 0.08°C/decade since 1913. However, as raw data embed multiple signals, overall trends are misleading.

Time-series analysis cannot discriminate between parallel signals. Exemplified by the Parafield study, protocols developed here at www.BomWatch.com.au partitions data into components: the part related to causal or physically deterministic covariables, and that related to site changes. The procedure is a form of covariate analysis with checks and balances built-in to assure the outcome is replicable and robust. A warming climate would be detected as an overarching systematic trend that was independent of covariable and site change effects.

This study focusses on the climate of Marble Bar, in the Pilbara region of W.A., allegedly the warmest place in Australia. Marble Bar is an Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) site used to track Australia’s warming; however, the quality of the Tmax data has not been rigorously assessed.

In addition to examining raw data, BomWatch protocols are also used to evaluate successive rounds of maximum temperature homogenisation by Bureau of Meteorology (BoM) scientists led successively by Neville Nicholls, Paul Della-Marta et al. and most recently, Blair Trewin. **The overarching question is whether the Bureau’s homogenisation methods are any good, or whether they create biased trends that are unrelated to the true climate**.

1. Background

> As all Australian weather stations have moved or changed, adjusting data for non-climate effects is essential for determining unbiased trend. However, homogenisation methods must be rigorous, objective and replicable.
>
> Read on …

Homogenisation of temperature data aims to remove or adjust the effects of site and instrument changes so homogenised data reflect the climate alone[1] and provided methods are objective, robust and replicable, it is an essential precursor to determining unbiased trend and change. While BoM scientists had been homogenising temperature data since before 1990[2], prior to a review by Peterson

---

[1] See: Trewin, Blair (2010). Exposure, instrumentation, and observing practice effects on land temperature measurements. WIREs Climate Change 1, 490-506 (DOI: 10.1002/wcc.46)

[2] Torok, S.J. and Nicholls, N. (1996). A historical annual temperature dataset for Australia. *Aust. Met. Mag.*, 45, 251-260.

and 21 co-authors, including BoM scientists Neville Nicholls, Neil Plummer and Simon Torok in 1998[1] the Bureau's methods had not been explained.

The approach they outlined in Peterson et al. (1998) involved:

- Identification of breakpoints or discontinuities in the target series to be homogenised using a combination of station metadata, statistical tests, visual (graphical) analysis and comparisons of interannual temperature differences between the target and a reference series derived from highly correlated first-differenced neighbouring data; and,
- Adjustment of target site data relative to pairs of highly correlated neighbours, or a composite derived from up to 40 cross-correlated datasets (within 6^{o} of Latitude and Longitude; Pearson's liner correlation coefficient for first-differenced series >0.7 and a significance level of $P < 0.05$).

<u>However, that is not what they did.</u> While the approach they outlined in Section 4.1.1 involved tortuous combinations of metadata, graphical and statistical tests, visual analysis of diurnal temperature ranges and correlation of first-differences between target and neighbouring site data, they say on page 1504: "*Finally, based on all this information, a* ***subjective decision was made on which adjustments were necessary and the magnitude of each***". So, under the guise of implementing complex methods, subjective criteria were used to potentially create pre-determined trends. Updating the previous Torok and Nicholls (1996) HQ_1 dataset, Della-Marta et al. (2004)[2] noted that "*The decision of whether or not to correct for a potential inhomogeneity is often a subjective one*", and that "*it was often impossible to reproduce the exact homogeneity adjustments of Torok and Nicholls (1996)*". While the need to account for data inhomogeneities is unarguable, the methods they advocated were problematic.

The two most obvious problems are: (i), metadata for most BoM weather stations was either scant or non-existent, and (ii), most stations had undergone concurrent up-grades (or site changes) since they were first established.

For example, former Aeradio meteorological observers and forecasters transferred from the Royal Australian Air Force to the Bureau in June 1946 around the time that many post offices were up-graded and telephone exchanges were being installed in post office yards. Former commonwealth-owned aerodromes were also upgraded before they were handed to civilian control in the 1950s. Temperature observations were metricated across Australia from 1 September 1972. Australia's telecommunications network was also substantially upgraded in 1970s often involving microwave towers and additional works adjacent to local post offices, and by 1975 over 90% of Australian subscribers had access to subscriber trunk dialling. Satellite communications (SATCOMS) were installed at airports sometimes in close proximity to meteorological (met) enclosures in 1986, and small (90-litre) Stevenson screens and automatic weather stations (AWS) were progressively introduced from the 1990s.

Long-term observations of rainfall and temperature were also not collected to detect climatic trend and change but rather to describe and monitor the weather and make short term predictions of inclement conditions likely to endanger aircraft, coastal shipping and local communities. Using imperfect long-term data to detect trends in the climate became a 'bolt-on' experiment that had its origins in activities undertaken by the World Meteorological Organisation (WMO) in the 1980s to support the Intergovernmental Panel on Climate Change (IPCC).

A chronology of Australia's involvement in the WMO and so-called *greenhouse policy* generally was summarised by the Commonwealth Parliamentary Library Service (https://www.aph.gov.au/

---

[1] Peterson TC, Easterling DR, Karl TR, 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, Forland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology* 18: 1493–1517.

[2] Della-Marta, P., Collins, D. and Braganza, K. (2004). Updating Australia's high quality annual temperature dataset. *Aust. Met. Mag*. 53, 15-19 https://www.cmar.csiro.au/e-print/internal/braganza_x2004a.pdf

About_Parliament/Parliamentary_departments/Parliamentary_Library/Publications_Archive/Background_Papers/bp9798/98bp04) and also Taplin (2013): Greenhouse: An Overview of Australian Policy and Practice (http://dx.doi.org/10.1080/14486563.1994.10648305).

According to the chronology, the first world climate conference was sponsored by WMO in Geneva in February 1979. Under the auspices of the United Nations Environment Program (UNEP) and WMO, another from 9 to 15 October 1985 in Villach, Austria "*increased the awareness of climate change*" and urged "*further research into both the causes and effects of climate change*". A third conference held in Toronto, Canada from 27 to 30 June 1988 resulted in the Toronto Agreement, which aimed to cut CO_2 emissions by 20 per cent by the year 2005 relative to 1988 as the base. The WMO and UNEP jointly established the IPCC in October 1988 and the Australian government adopted the Toronto target in 1990. At that time, anthropogenic greenhouse gas emissions were forecast to increase global temperatures by a 3°C by 2030.

The Australian Greenhouse Office (AGO) was formed in 1998[1] and in March 1989 at The Hague, along with 23 other countries Australia pledged to "*work through the United Nations to take measures to control global climate change*". So driven by the AGO, with CSIRO overseeing the BoM, and the Division of Atmospheric Research the lead agency, detection of trend in Australia's long-term temperature and sea level records consistent with model predictions became a high priority.

The AGO eventually morphed into the Department of Climate Change, Energy, the Environment and Water, which with 37 sub-agencies or points of contact (Appendix 1) is arguably the most powerful, politically dominant bureaucracy since Federation.

Despite its lack of objectivity and the potential for homogenisation to create bogus trends, BoM scientists were apparently keenly supported from 1978 to 2003 by then Bureau Director, Dr John Zillman. Zillman was also Australia's permanent representative with the WMO from 1978 to 2004, Vice-President from 1987 to 1995 and President from 1995 to 2003. Neville Nicholls was involved with the WMO Working Group on Long Range Weather Forecasting Research from 1982 to 1986, and later from 1986 to 1990 with the World Climate Research Program. Nicholls was also Lead Author of the Section C, Supplement to the IPCC First Scientific Assessment Report (1991-1992) and Convening Lead Author of Chapter 3, Observed climate variability and change, of the IPCC Second Scientific Assessment (1993-1995) (see: https://www.eoas.info/biogs/P003289b.htm).

Thus, with assistance from Nicholls, Zillman directed the Bureau and its scientists and also activities and projects within WMO that supported the IPCC. He was also a member of the Prime Minister's Science, Engineering and Innovation Council within the office of Australia's Chief Scientist[2] and he was influential within the Australian Academy of Science and CSIRO. With Agenda 21 morphing into government policy[3] following the Rio Earth Summit in June 1992 and Zillman overseeing everything there was little likelihood that BoM's homogenisation methods would be seriously scrutinised. Thus, up to this time the Bureau's methods have not been independently, objectively audited.

2.1 The problem of inter-site correlations and comparisons.

Methods used by BoM scientists to homogenise Australian temperature records are unsound. Correlated comparators used to detect and adjust changepoints in target-site data likely embed parallel faults, which create bogus trends in homogenised data.

Read on …

Leaving aside the apparent complexity of BoM's methods discussed previously, the paucity of reliable metadata and the use of subjective criteria for identifying and adjusting breakpoints, selection of comparator datasets from the population of potential neighbours using linear correlation of *first-*

---

[1] https://en.wikipedia.org/wiki/Australian_Greenhouse_Office
[2] https://en.wikipedia.org/wiki/John_Zillman
[3] https://www.un.org/en/conferences/environment/rio1992

differenced[1] monthly or annual means is neither statistically valid nor scientifically sound. Whether constructed using anomalies, first differences, or multi-site medians of first difference, as correlated data likely embed parallel faults the key assumption that highly intercorrelated reference series would be representative of the background climate[2] is flawed. It is also clear from Torok (1996), Torok and Nicholls (1996), Della-Marta et al. (2004) and subsequent papers by Trewin[3], that ignoring changes that happened, or adjusting for some that made no difference (vis-à-vis Townsville Airport: https://www.bomwatch.com.au/ data-quality/climate-of-the-great-barrier-reef-queensland-climate-change-at-townsville-abstract-and-case-study/) provided considerable scope for Nicholls and protégés, including Trewin, to manipulate changepoints and adjustments and thereby create pre-determined trends. Lack of objectivity and replicability as noted by Della-Marta et al. (2004) are the hallmarks of potentially dodgy methods.

2.2 The crucial importance of station metadata

Verifiable metadata (data about the data) is a touchstone issue. However, as sources of information available for verifying that site conditions have or have not changed is sparse or non-existent, the Bureau's site-summary and ACORN-SAT metadata is demonstrably unreliable.

Read on …

As most original station files cannot be viewed in the public domain, the unreliability and unverifiability of station metadata and/or the use of statistical tests and inter-site comparisons for pinpointing times when conditions changed is of paramount concern. Also, adjusting changepoints using highly correlated neighbours, whose data are not verified as homogeneous, potentially smooths discontinuities in target site data so they appear to be continuous. The method lacks quality assurance and statistical control in detecting changepoints and also in making adjustments. For instance, if an abrupt change in ambience causes average Tmax to increase by 0.3°C, the appropriate adjustment is to reduce average temperature after the change by the same amount, not smooth it over relative to correlated neighbours leaving the effect on trend intact. It is also confusing that BoM adjusts temperatures *into the past*, rather than statistically adjusting near-term and incoming data.

Objective, replicable homogenisation would investigate time-related changes in dataset properties using objective statistical methods and relate those to what is known about the site *post hoc*. For example, by cross-referencing with BoM metadata, and/or documents, maps, plans and aerial photographs held by the National Archives, the National Library of Australia (NAA and NLA), state authorities, museums, historical societies and other sources such as newspaper reports etc. Even if supporting information is not available, statistical detection based on the probability of an abrupt sustained change against the NULL hypothesis of *no change* should be sufficient to evidence that change occurred.

Multiple studies reported by BomWatch have shown that BoM metadata is the least reliable source of information relating to individual weather stations, and that claims by the Bureau, CSIRO, and university professors and scientists that HQ and ACORN-SAT sites had been well researched, and that data had been unbiasedly adjusted **is substantially untrue**. Multiple instances have also been found where site changes were ignored/adjusted to infer or reinforce trend or change in homogenised data, which was erroneously attributed to the climate.

There is a human dimension too. Since the 1990s, homogenised temperature datasets have been widely used by teachers, professors and communicators to stir unwarranted fear in the minds of

---

[1] Differences calculated between successive values (e.g., $x_2 - x_1$). Usually used to remove first-order autocorrelation (dependence of one value on that previous). However, it can be shown that as differences are negative from summer to winter and positive from winter to summer, first-differenced monthly data embed cycles that inflate significance of Pearson linear correlation between independent series. (An underlying trend or cycle common to both series invalidates the test.)

[2] Trewin, B.C. 2012. Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset. CAWCR Technical Report 49, Centre for Australian Weather and Climate Research, Melbourne.

[3] Trewin, B.C., 2018. The Australian Climate Observations Reference Network – Surface Air Temperature (ACORNSAT) version 2. Bureau of Meteorology Research Report BRR 032.

vulnerable school children, mothers, families and undergraduates. They also feed into well-funded, coordinated campaigns by the cohort of mainly WWF-related activist groups such as GetUp!, Australian Youth Climate Coalition, Earth Hour, 1-million women, the Wentworth Group of Concerned Scientists, the Climate Council, Farmers for Climate Action etc[1]. and elitest political *independents* including Zali Steggall and the emerging 'Teals'. Policy-driven science also undermines trust in Australia's scientific and political institutions including CSIRO, the Australian Academy of Science, education Trade Unions, superannuation funds, boardrooms and across the lavishly funded, increasingly woke university sector.

2.3 The Stevenson screen issue

Prior to installation of 230-litre Stevenson screens that protect thermometers from both direct and indirect radiation, historic data were often reported as 'shade temperatures' to indicate thermometers were not directly exposed to the sun. However, as they were still subject to indirect heat sources and sinks, particularly at night, data were nevertheless biased. The *West Australian* newspaper reported on Friday 22 November 1901 that 'true shade' readings at the Marble Bar post office during the previous week were: Thursday, 107°F (41.7°C); Friday 107.5°F (41.9°C); Saturday, 111.2°F (44.0°C); Sunday, 109.2°F (42.9°C); Monday 110.6°F (43.7°C); Tuesday 111.8°F (44.3°C) and Wednesday, 110.8°F (43.8°C) (http://nla.gov.au/nla.news-article24763859). As Celsius values reported for those days by BoM's climate data online database are identical (Station ID 04020), raw values are unadjusted for possible site and instrument biases.

Writing in the *Western Mail* on 25 December 1909, commentator E. Paget Thurston wrote "*What makes the mistake of building the* (Marble Bar) *post office*" (out of stone) "*more unfortunate is that the official meteorological records are taken from a thermometer in the rear. Almost certainly the maxima are affected by radiation from the adjacent mass of heated buildings thus making the records higher than they truly are, quite an unnecessary exaggeration in the case of Marble Bar*." (http://nla.gov.au/nla.news-article37401491)

Nevertheless, temperatures measured in 2022 by the rapid-sampling automatic weather station (AWS) probe housed in a small 60-litre Stevenson screen beside a dusty track southwest of the Marble Bar Roadhouse and Travellers Rest Motel are unlikely to be comparable with "true shade" temperatures observed in 1901 at the rear of the original Post Office (Figure 1). Therefore, naïvely 'joining the dots' without considering the effect of ambient site factors on observations is unlikely to provide unbiased estimates of long-term trend and change.

Figure 1. Left: Built of local stone, the original government office in Marble Bar housed the police station, mining registrar, courthouse and post office (closest to the camera). Water tanks were supplied from a well in the bed of Sandy Creek, which is below the ridge and the building in the distance was the leprosy sanatorium/hospital. While shaded in this 1938 National Archives of Australia photograph, the rear of the building where temperature was measured faces east. Right: The current automatic weather station site, and behind the vehicle, the 60-litre wooden screen beside a dusty track southwest of the Roadhouse and Travellers Rest Motel (BoM photograph, 16 January 2012).

[1] See p. 39 in: https://uploads.prod01.sydney.platformos.com/instances/403/assets/Annual-Reports/PEF%20Annual%20Review%202015%20-%20final.pdf?updated=1619513905 b

Furthermore, Excel or other naïve analysis tools can only show superficial trajectories inclusive of site change effects (Figure 2). Believing that Bureau scientists have diligently searched station records for station and instrument changes is also misplaced, especially for sites whose data commenced near the beginning of the 20th Century. For instance, BoM metadata failed to mention that a Stevenson screen was not installed at Marble Bar before 1911. First correspondence between the Bureau and the post office was in March 1930[1] when it was noted that the screen was knocked over by a willy-willy (local tornado) and they (presumably the post office people) sat it back up and set the legs in concrete. They also proposed to move the screen and instruments in November 1944, but there was no record of the move taking place. Just a month later in December 1944, the screen was “brittle and splitting” but there is no record of how long it had been in that state or if or when it was repaired or replaced.

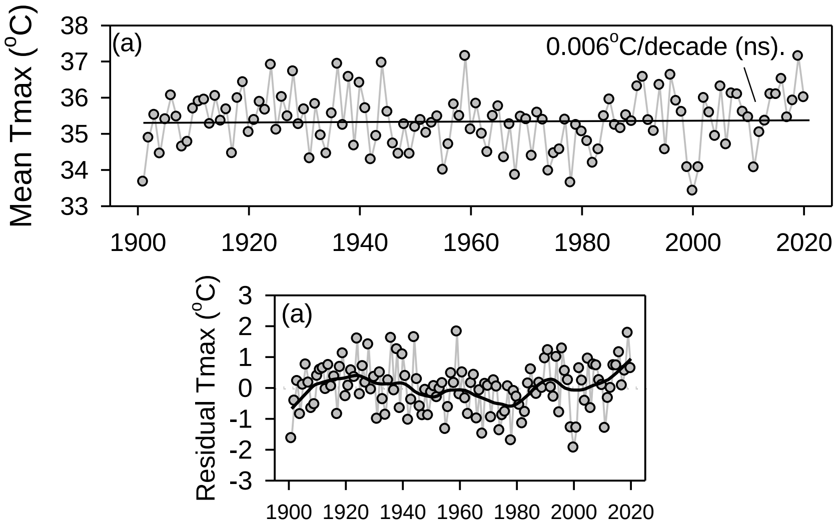


Figure 2. Annual mean maximum temperature at Marble Bar (a) shows a non-significant (ns) trend of 0.006°C/decade since 1901. However, as zero-centred residuals in (b) embed patterns and irregularities typical of missing step-change variables (smoothed by the LOWESS curve), trend in (a) is likely to be spurious.

Simon Torok further noted that because it was reading low, the Tmin thermometer was replaced 21 years later in January 1965, then in 1967 the screen was moved a small distance towards the Post Office possibly to be out of the way of the telephone exchange built earlier in 1961. Some 17-years later in August 1984, the screen reportedly moved from the Post Office to the general store 500m east and in July 1988, due to the site becoming overgrown by trees, it moved again 1 km to a residence. Subsequent research has found that the relocations listed by Torok (1996), and subsequently by ACORN-SAT and site-summary metadata were incorrectly reported.

Metadata failed to mention that the post office was unroofed and severely damaged and that the raingauge was blown away by a cyclone on 2 March 1941 (https://trove.nla.gov.au/newspaper/article/78813277). Also, that it was again unroofed and damaged on 30 December 1947 (https://trove.nla.gov.au/newspaper/article/79679196) possibly forcing the site to relocate.

While data were probably telegraphed daily with hard copies posted to the Bureau each month, it is hardly credible that a site deemed by BoM scientists to be ‘high quality’ operated for decades between inspections and reportage. Marble Bar post office station files from 1930 to 1969 are located in the National Archives of Australia office in Perth WA but due to cost constraints have not been accessed.

[1] Torok SJ (1996). Appendix A1, in: “The development of a high-quality historical temperature data base for Australia”. PhD Thesis, School of Earth Sciences, Faculty of Science, The University of Melbourne.

2.4 A framework for auditing the Bureau's homogenisation methods

Methods used to homogenise Australian temperature records are fundamentally unsound. Strongly correlated comparator datasets used to detect and adjust changepoints in target site data likely embed parallel faults, which create bogus trends in homogenised data.

Read on …

The key problem in the BoM's homogenisation methods is the reliance for both the detection of target-site inhomogeneities and their adjustment on comparisons with *strongly correlated* first-differenced series that likely embed parallel faults. While it is not possible to replicate the Bureau's various homogenisation iterations using their data, metadata or Python code, objective statistical methods are highly suitable for comparing data for the same site with breakpoint scenarios used for homogenising the same data. It would be expected for example, that provided site change-effects were faithfully reproduced, homogenisation scenario 'A' would result in better fitting data or a measurably superior outcome compared to say, a scenario based on guesswork.

This study investigates the overall quality of maximum temperature data for Marble Bar and outlines an objective statistical method for auditing the Bureau's homogenisation methods.

2. Case study Marble Bar

3.1 The general climate

As rainfall is low and strongly summer-dominant and potential evaporation is high and the landscape highly weathered, the Pilbara region of north Western Australia is warm throughout the year.

Read on …

Situated in northwest Western Australia on the edge of the Great Sandy Desert, the climate of Marble Bar is classified as 'hot, winter-dry'[1]. Median rainfall is 363 mm; 50% of years from 1901 experience between 253 (1st quartile or Q_1) and 451 mm (Q_3). As 18 of the highest annual rainfalls (which were received in 15% of years) accounted for >25% of the 120-year grand-total, rainfall is highly skewed or long-tailed (i.e., dominated by infrequent extreme events). Average rainfall may thus poorly indicate rainfall likelihood.

Monthly rainfall distributions for the SILO[2] grid cell that encompasses Marble Bar (Latitude -21.15, Longitude 119.85) shows rainfall is highly seasonal with practically zero rain from April to November (Figure 3). During the 'wet' from December to March, monthly evaporation grossly exceeds rainfall even in the wettest month of February.

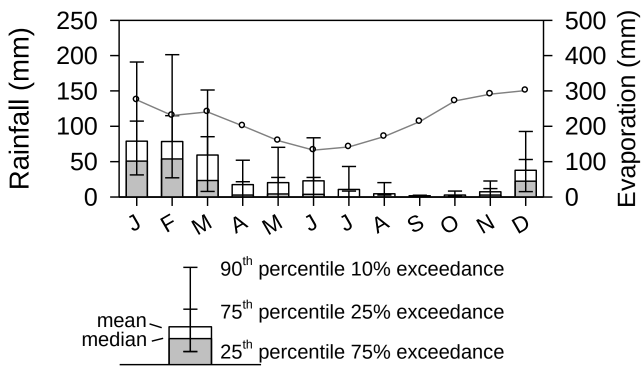


Figure 3. Monthly rainfall distributions for the SILO grid-cell surrounding Marble Bar (centred on Latitude -21.15, Longitude 119.85). Rainfall is strongly summer dominant and as the mean exceeds the median in all months rainfall is highly positively skewed. Even during the wet season from December to March, the likelihood that rainfall would exceed monthly A-pan evaporation (right-hand scale) is less than in 25% in the wettest month of February. Due to the northward progression of the subtropical ridge in spring and the anti-clockwise rotation of associated high-pressure cells across the continent, the low likelihood of rain in November and December has implications for the occurrence, intensity and duration of summer heatwaves across southern Australia.

A notable feature of the long-term record is that average rainfall from 1900 to 1974 was less than from 1975 to 2020 (326 vs. 425 mm) (Figure 4). Rainfall was also notably higher from November 1993.

[1] Stern, H. and de Hoedt, G. (2000). Objective classification of Australian climates. *Aust. Met. Mag*. 49 (2000) 87-96.

[2] https://www.longpaddock.qld.gov.au/silo/

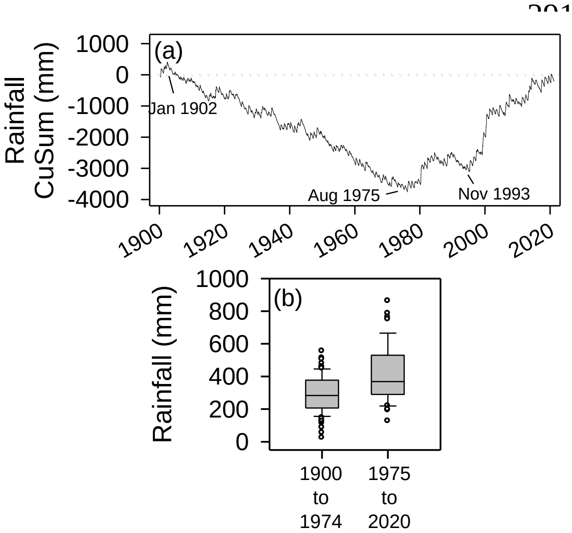


Figure 4. The rainfall CuSum curve (a) shows average *monthly* rainfall was cumulatively less from January 1902 to August 1975 (the curve declines), near-average for the period to about 1993 (it showed no consistent trend) then above average for the remainder of the record to December 2020 (the curve ascends).

Box-plots (b) contrast the distribution of annual rainfall pre- and post-1974. The box represents the upper and lower quartiles, the line within indicates the median, while whiskers represent the 10th and 90th percentiles. Values outside those limits are indicated as outliers.

Post-1975 median rainfall was higher and the distribution of values was skewed to a greater extent by upper-range values than from 1900 to 1974.

Assuming a plant-available water holding capacity of 50 mm, residual soil water plus rainfall minus potential evaporation calculated using a monthly water balance found surpluses occurred only 17 times since 1903 and all after an upward 'bump' in 1968. It could not be argued that droughts are more frequent and longer-lasting than before the rainfall reversal in 1975.

3.2 Methods

> Daily maximum temperature data downloaded from the BoM were abutted as they were for ACORN-SAT, summarised and investigated using objective statistical methods.
>
> Read on …

Daily temperature data for Marble Bar (BoM ID 04020) from 9 February 1901 to 22 August 2006 and for the AWS (04106) from 26 September 2000 to 31 December 2020 were downloaded from http://www.bom.gov.au/climate/data/ abutted on 24 March 2003 (which aligned with the ACORN-SAT merge) and summarised into annual averages and other attributes using the statistical program R v.4.1.2 (summary data are provided separately). Average annual Tmax was aligned with annual rainfall, which was also sourced from the BoM, and analysed using mainly the *Rcmdr* and *lsmeans/emmeans* packages[1] using protocols previously outlined in the Parafield Case Study.

3.3 Results

> The BomWatch approach is an application of covariance analysis, where sources of variation are explained and accounted for objectively prior to determining trend. As rainfall and site change effects only explained 50.3% of Tmax variation (R^2_{adj} = 0.503) data quality was poor. Further, as no trend was evident in multiple linear regression residuals, it could not be claimed that the climate has warmed due to a third variable.
>
> Read on …

3.3.1 Raw data and step-change analysis

Rainfall naïvely explains 26.1% of variation in Tmax (R^2_{adj} = 0.261) and slightly more (29.3%) if outliers are ignored (Figure 5(a)). Residuals (the non-rainfall portion of the signal rescaled by adding grand-mean Tmax) embed significant step-changes in 1913, 1945, 1988, 1999 and 2013, which are indicative of underlying inhomogeneities (Figure 5(b)). As the deterministic rainfall portion of the Tmax signal was removed by linear regression, residual inhomogeneities reflect underlying structural changes in the data. A non-significant up-step from 1935 to 1945 suggested the screen was in a deteriorated condition for a decade before December 1944 when according to Torok it was noted on the file.

---

[1] https://www.rcommander.com/; https://cran.r-project.org/web/packages/lsmeans/lsmeans.pdf

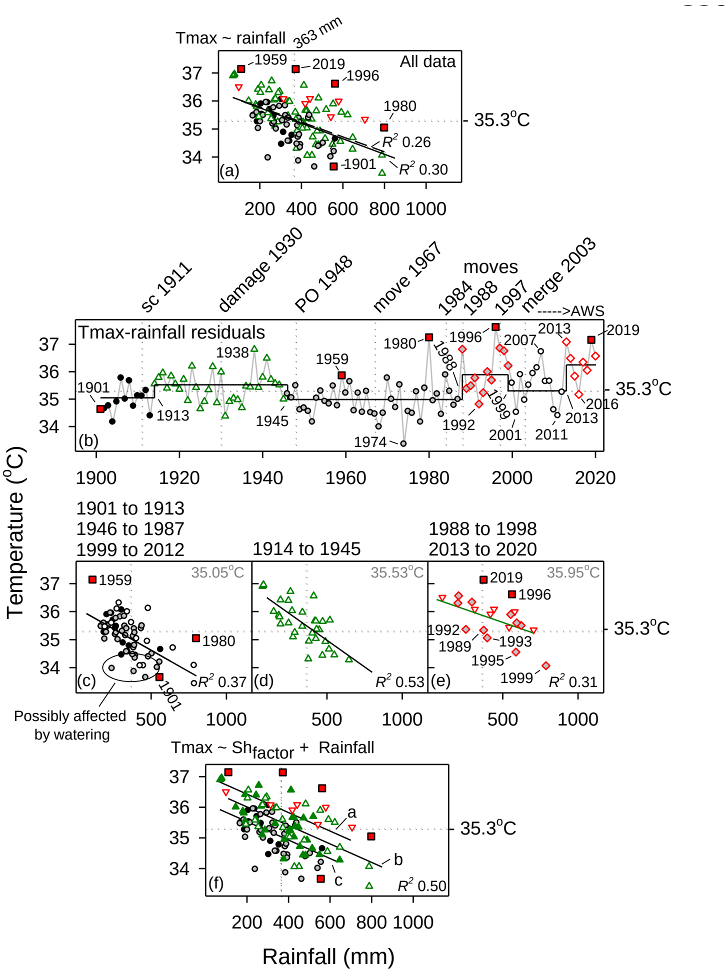


Figure 5. Composite analysis of Marble bar Tmax.

Overall average Tmax (35.4°C) comprises three groups who's within-group means are the same but who's between-group means are different (according to 1-way analysis of variance and Turkey pair-wise comparisons, Figure 5(c) to (e)). Note that regression lines in Figures 5(c) to 5(e) were free-fit, whereas for the pooled dataset, those in Figure 5(f) were confirmed as parallel by multiple linear regression.

Multiple linear regression of the pooled dataset of the form: Tmax ~ Groups + rainfall (Figure 5(f)) verified that group means were different (lines were not coincident) and that grouped responses to rainfall were the same (interaction was not significant thus, regressions by group were statistically parallel). However, groups and rainfall simultaneously explained only 50.3% of Tmax variation, which is low.

Of the data segments shown in Figure 5(c) to (e) data from 1914 to 1945 was the soundest (R^2_{adj} >0.50). The pooled difference in Tmax between Figure 5(c) *vs*. (d) was 0.48°C (sem = 0.114), that between Figure 5(d) *vs*. (e) was 0.42°C (0.161). Thus, apparent warming attributable to site changes (Figure 5(c) *vs*. (e)) was 0.9°C (0.145). No trend was evident in MLR residuals, which were acceptably normally distributed, independent, with equal variance across categories.

3.3.2 Precision of daily observations

Marble Bar was not a Bureau-manned site but was operated by post office staff or paid or unpaid volunteers. As the Stevenson screen fell into disrepair and the site moved to various locations whose coordinates were unknown or not stated in metadata, there is reason to question whether daily observations were accurate and reliable.

The precision with which observations are made (i.e., whether daily observations were rounded to the nearest decimal place, or to the nearest whole or ½ degree) provides insights relating to operator training, diligence and many assist in diagnosing bias or other data problems.

Irrespective of scale (°F or °C), if daily data were diligently observed to one decimal place over a year, each of 10 decimal fractions (x.0, x.1 … x.9 (where x is an integer)) would proportionally equal 10% of the total number of observations, while their average [(0.0 + 0.2 + .. + 0.9)/10] is expected to be 0.45. However, as the Fahrenheit scale is not linear to Celsius but only Celsius values are available from the BoM, data before 1 September 1972 required back-transformation from °C to °F prior to calculating fractions and ratios.

To aid interpretation of Figure 6, Monto Carlo simulations involving of >2000 random observations within the range of Fahrenheit Tmax, found that:

- As stated above, Fahrenheit temperature observed <u>consistently to the nearest decimal fraction</u> back-transform from Celsius in the ratio of about 10% per fraction. Likewise, after 1 September

1972, Celsius temperatures observed to 1-decimal place also sum to a frequency of about 10%/fraction/yr.

- However, <u>temperature observed in whole and ½°F</u> back-transform from °C in the ratio of about 30% x.0, 10% x.9 and x.1; and, 30% x.5, 10% x.4 and x.6, with no back-transformed x.2 or x.3, or x.7 or x.8 decimal fractions. After 1 September 1972, over representation of x.0°C and x.5°C occurs at the expense of other fractions and indicates rounding-preference by observers.
- <u>Temperature observed in whole °F only</u>, back-transform from °C in the ratio of about 55% x.0 and 45% shared between x.1 and x.9. Temperature observed in whole °C after 1 September 1972 are indicated by that fraction (x.0°C) approaching 100% of all fractions.

Before the Stevenson screen was supplied in 1911, most observations were reported as whole°F, about 20% were reported x.5°F with other decimals rarely reported. Whole degrees were again predominant after about 1924 and were reported almost without exception from 1945 to 1960 following which, until metrication, around 60% of observations were reported as whole and ½°F. Coinciding with metrication and the introduction new thermometers and revised protocols, reporting of whole°C stepped-down to around 10%, nevertheless, observers continued to preferentially record whole degrees and x.2, x.5 and x.8°C at the expense of x.1 and x.9, while the frequency of other fractions was randomly high or low.

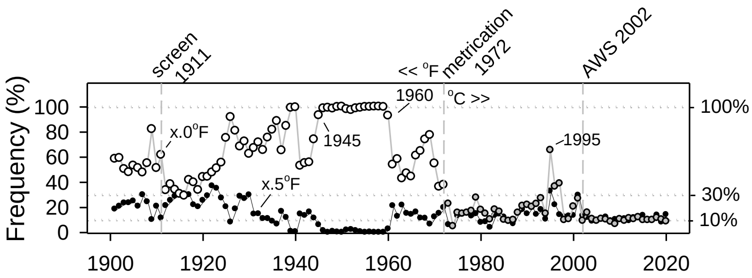


Figure 6. Frequency/yr of reporting whole and ½°F before November 1972 and Celsius thereafter (open and solid circles). From 1995 to 1997 whole and ½°C were over-reported at the expense of other fractions, particularly x.3, x.4 and x.9°C.

While differences probably reflect training, diligence and observer preference, Figure 6 shows that despite data for Marble Bar being considered 'high-quality', historic data were not observed with high precision particularly for the period from 1945 to 1960. It is not certain whether observers consistently rounded-up (or down) i.e., whether temperatures observed within a 1°F index were routinely rounded-up (or down) to the next, which could result in upward (downward) bias, or only those that exceeded x.5°F with the remainder being rounded-down. As precision is not homogeneous, comparing highly precise AWS data with historic data measured at a range of locations under variable conditions observed in whole and ½°F results in uncertainty, which may invalidate such comparisons.

3.3.2 Temperature extremes

Extreme temperatures are those within the tails of data distributions, typically observations less than the 5th or greater than the 95th day-of-year percentiles calculated across the entire daily record (Figure 7(a)). For an unchanging site, counts of low and high extremes/yr are expected reflect the weather and vary randomly over time and within the bounds of statistical probability their ratio (Hi_N/Lo_N) would be constant.

Applying the logarithmic transformation so data are symmetrical, Hi_N/Lo_N ratios were not homogeneous (Figure 7(b)). Numbers of high extremes declined abruptly from 1966 to 1985 and increased markedly from 2015. Changes over a period of 20 years (1966 to 1985) and from 2015 going forward are likely to reflect site changes rather than persistent changes in the weather.

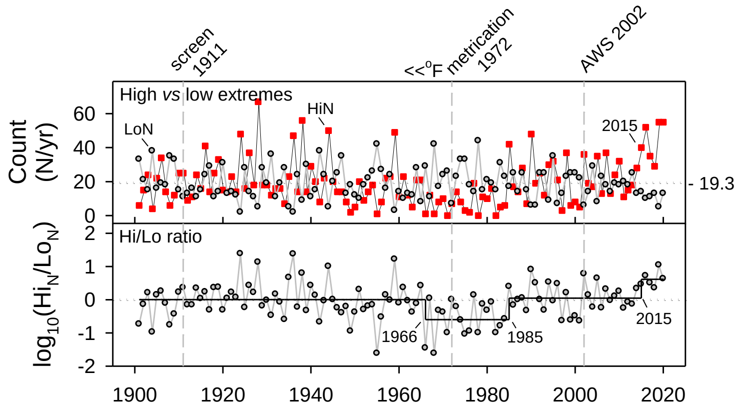


Figure 7. Frequency (N) of daily Tmax/year <5th and >95th day-of-year percentiles (top panel), and their log$_{10}$ transformed ratio (lower panel). The step-down in 1966 indicates the Stevenson screen relocated to a cooler environment than beforehand, while up-steps in 1985 and 2015 indicate conditions ambient conditions were generally warmer. The average across counts of extremes was N = 19.3/year.

According to Simon Torok[1] the minimum temperature thermometer was replaced in 1965. It may have previously been reading high, or a faulty Tmax thermometer may have been replaced as well. Also, that in May 1967 the Stevenson screen moved "up sloping ground toward the (then) PO". Further, the post office may have closed after August 1984 when site moved "from the PO to the general store 500 metres east". The abrupt decline in the Hi_N/Lo_N ratio indicates the move up-slope in 1967 was to an area that was possibly watered or shaded, while conversely, the site at the general store was more exposed and warmer. The predominance of upper range extremes after 2015 is consistent with a site change rather than a persistent change in the weather.

The 2015 step-change is further investigated in Figure 8(a), which compares daily Tmax probability density functions (pdf plots) for equivalent time periods before and after the step-change (from 1 January 2008 to 31 December 2014, N = 2441 days, vs 1 January 2015 to 31 December 2021, N = 2557 days).

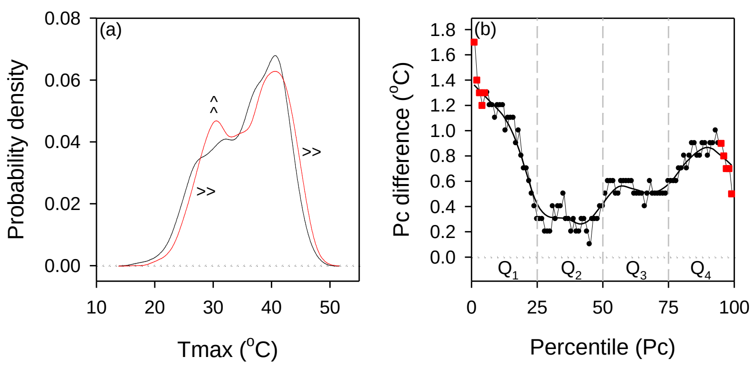


Figure 8(a) post-2015 probability density functions (red) have shifted warmer relative to pre-2015, with much of the increase attributable to daily Tmax values <30°C, or within the average of the dry season (may-September) range. Differences between cumulative distribution functions in (b) show unexpected systematic trends and clustering suggestive of artificial influences in tails of data distributions (red squares).

Probability density functions can be likened to histograms of the probability of daily values falling within an infinite number of bins. Calculated over the same range the area subtended by each curve is unity so plots are directly comparable. (The bi-modal distribution in Figure 8(a) reflects the underlying sinusoidal seasonal cycle; i.e., the single peak in summer (December) and the predominance of lower peaks in autumn and spring.)

Figure 8(b) shows the difference in percentile distributions (post-2015 minus pre-2015) for the same data overlaid by a LOWESS curve (**Lo**cally **We**ighted **S**catterplot **S**moothing), which is a non-parametric line of best fit. Consistent with the pdf-plot, post-January 2015 data are >1°C warmer within the 1st quartile (Q_1) compared with pre-January 2015 data with the lower (coolest) 5% of values being 1.2°C to 1.8°C warmer than those for pre-2015. Daily values within the 2nd Quartile are uniformly about 0.3°C warmer but corresponding with the median, differences step-up to 0.6°C warmer. A quasi-linear increase to 0.9°C occurs over the course of the 3rd Quartile to an apparent maximum at the 95th percentile after which differences decline. The shape and rate of decline suggests daily values greater than about 45°C were censored (i.e., that due to the AWS exceeding the range of its calibration, higher values were adjusted-down.)

---

[1] Torok, S.J. (1996). Appendix A1, in: "The development of a high-quality historical temperature data base for Australia". PhD Thesis, School of Earth Sciences, Faculty of Science, The University of Melbourne.

So, while across the range of daily values, pdf plots show the relative *shape* of daily data distributions, differences between respective cumulative distribution functions show the *nature* of the disturbance (i.e., where differences across the percentile range are random or systematic).

All things being equal, it was expected that differences in percentile distributions would be random around the rainfall-adjusted 2015 up-step of about 0.4ºC. However, as tails of distributions were impacted by abrupt changes within the 1st and 3rd quartiles this was not the case. Except that on 21 March 2016 metadata noted data were "processed by ASOS (PBA)", whether changes in distributions reflected the weather, the site, the Stevenson screen, the instrument or data processing could not be determined from data alone.

3.4 Interim discussion

> Tmax data for Marble Bar were affected by watering and shade, deterioration in the condition of the Stevenson screen and instruments, lackadaisical practices including reporting in whole degrees from 1945 to 1960, infrequent reporting of site conditions, site moves after 1967, and possible daily data in-filling.
>
> Read on …

Lack of precision (low R^2_{adj}) in the relationship between Tmax, site changes and rainfall shows the quality of Marble Bar data is relatively poor. However, as residuals were random there is no evidence of an unaccounted-for underlying trend or change. Given that knowledge about the site is sparse and inaccurate and that data are fairly rough, it could not be claimed that the climate of Marble Bar has changed or warmed independently of site factors.

Torok's and ACORN-SAT metadata do not cross-reference to coordinates provided in site-summary metadata; distances between locations do not agree with those calculated by Google Earth Pro; changepoints listed by Torok and ACORN-SAT v1 and v2 do not agree, and without removing the effect of rainfall, changepoints detected statistically by Torok and ACORN-SAT were for a confounded Tmax dataset, which was faulty.

While some documented changes made no difference (screen deterioration then refurbishment in 1930, moves in 1967 and 1984 and merging with the AWS in 2003), others may reflect faulty or missing metadata. For instance, although supplied in 1911 the Stevenson screen may not have been installed until 1912/13. Also, although the step-change in 1945 is consistent with a site move in 1944, the new Post Office, residence and fuel store was not completed until 1948 (Figure 9). Metadata also ignored the impact of cyclones in March 1941 and December 1947 which may have damaged the Stevenson screen or even rendered the post office uninhabitable. NAA files also show the telephone line exchange was annexed to the post office in 1961/62.

Following the move to a residence (for which no coordinates are provided), then 1 km east in July 1988, then 300 m southwest in October 1997 the quality of the data, the sites, the screen or the instruments deteriorated markedly. Also, data after 1998 was subject to unknown effects that resulted in systematic warm-cool sequences: Tmax warmed from 1992 to 1996, successively cooled to 2001, warmed again to 2007, cooled to 2011, warmed to 2013, cooled to 2016 and warmed again to 2019. Although possibly reflecting haphazard practices, poor site control or irregular servicing/painting of the screen, patterns were irregular, non-random, not due to rainfall and not evident in data before the site allegedly moved from the post office in 1988.

Metadata is a pivotal issue. The new post office which opened in August 1948 was not mentioned in ACORN-SAT or site-summary metadata. The claimed move 500m *east* on 1 August 1984 probably marked when the post office closed and the licenced post office opened at the general store opposite the Ironclad Hotel, 200m *west*. The subsequent move on 19 July 1988 was probably to the Roadhouse and Travellers Rest Motel 500m east (not 1 km, which would have placed the new site outside the village). It was probably from there that the screen moved 100 m (not 300m) southwest to its present position in October 1997.

Figure 9. The left-hand panel shows (L to R) the police station, mining registrar's office, court house, and post office; and (right) the new post office that was opened on 1 August 1949. (Due to asbestos contamination the building at 23 Francis Street is no longer used.) Plans indicate the postmaster's residence, with generator shed and fuel bay behind, was built adjacent to the new post office facing Contest Street. (Photographs courtesy of the National Archives of Australia). According to files in the National Archives the telephone line exchange was probably built in 1961. That building is currently occupied by Telstra.

From a climatic trend perspective, Marble Bar weather station was not well documented, historic data was never very good and it is clear from site-summary and ACORN-SAT metadata that Bureau scientists had not researched the site as they claimed they did. The many scientists using ACORN-SAT perpetuated the problem by not undertaking due diligence on data they used.

3. Homogenisation

Choice of comparator datasets is limited by their sparseness, with coastal and near-coastal sites predominant. No in-depth checks are made to assure comparator datasets are homogeneous, which combined with the use of linear correlation in their selection results in bias.

Read on …

The choice of comparator datasets used to homogenise Marble Bar HQ and ACORN-SAT is limited by the sparseness of their distribution across the Pilbara and Kimberley regions of WA (Figure 10). The majority of long-term weather stations occur along the NW-coast with many of those having been originally set up as Aerado stations used to monitor the air route between Perth and the former Dutch East-Indies capital of Batavia (now Jakarta) and on to London.

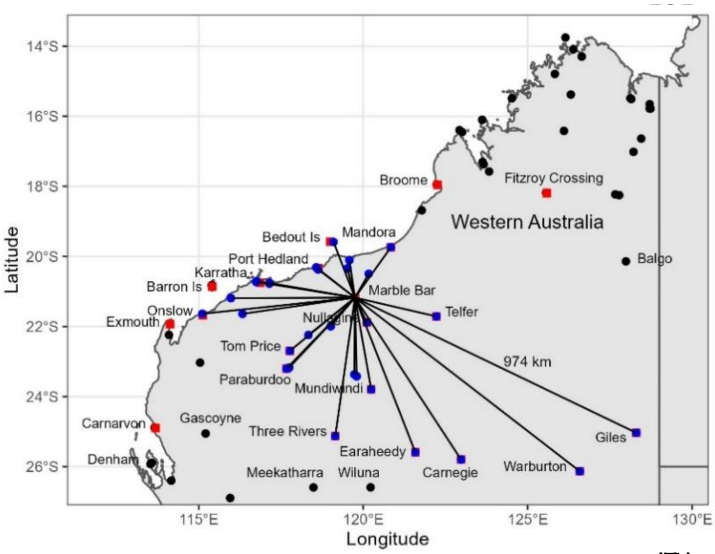


Figure 10. The distribution of weather stations having more than 15 years of data across northern WA (black circles, some named), major towns also having data (red squares) and sites used to homogenise Marble Bar HQ and ACORN-SAT (blue circles, some not named). Note the lack of comparators E and NE of Marble Bar. A list of ACORN-SAT adjustments and comparator stations used is given in the ACORN-SAT catalogue (www.bom.gov.au/climate/acorn.sat/stations/#/4106). It is noted that most comparator sites including other ACORN-SAT sites (Wittenoom, Giles, Port Hedland, Onslow, Exmouth and Meekatharra) are also impacted by station moves and changes, so like Marble Bar their data is not homogeneous.

As all those sites (and others used to homogenise Marble Bar) have changed or moved or been affected by developments and are not themselves homogeneous, their use as comparators is scientifically and statistically flawed. For instance, in their early days many sites including Mandora, Port Hedland and Onslow were watered during dry seasons either for comfort or to reduce sand-blasting and instrument damage. In contrast, Telfer is a fly-in-fly-out workcamp run by Newcrest

Mining (ASX: NCM) and with Giles, is one of the most isolated weather stations in Australia. Poor metadata and lack of site control are a consistent problem across the Bureau's network.

4.1 Homogenisation breakpoint analysis

Should homogenisation breakpoints align with changes in the data, segmented relationships between Tmax and rainfall would show an equal or better fit to that determined for raw data in Figure 5. However, this was not the case…

Read on …

In this section, breakpoints specified by Della-Marta *et al* (2004), and Blair Trewin's ACORN-SAT v.1 and v2 shown in Figure 11, were evaluated as scenarios. Scenarios were compared with results discussed previously in relation to Figure 5 using the same BomWatch framework.

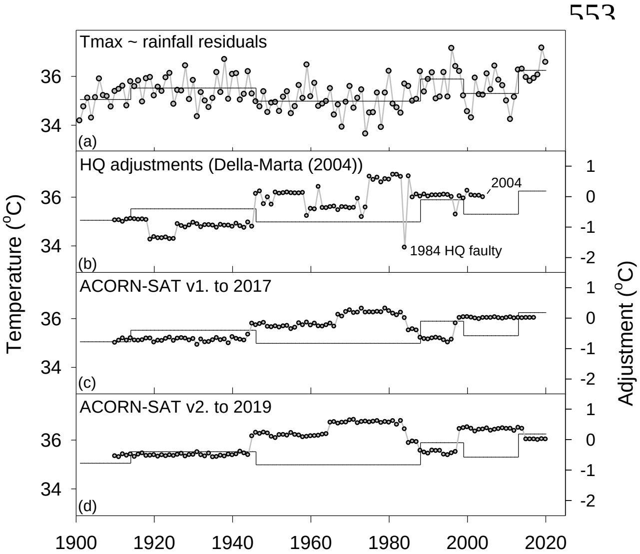


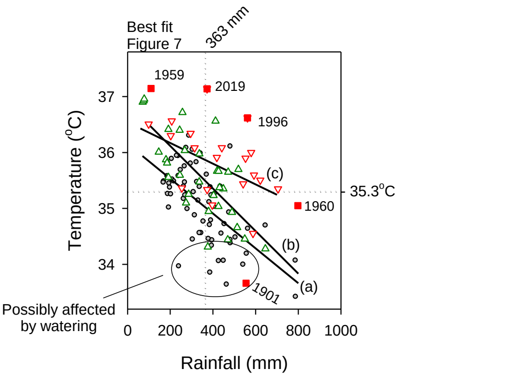
Figure 11. Step-changes (horizontal lines) in Tmax rainfall residuals ((a) from Figure 5(b)) compared with adjustments (grey circles, right axis) applied by Della-Marta *et al* (2004) to derive the homogenised High Quality (HQ) v.2 dataset (b), and Blair Trewin's ACORN-SAT v.1 (to 2017, (c)) and ACORN-SAT v.2 (to 2019, (d)). Adjustments were calculated as differences between raw data and annually summarised homogenised data (homogenised minus raw). Note that while Marble Bar data commenced in 1901, homogenised data commenced in 1910.

Should homogenisation breakpoints fit the data more accurately, it is expected that segmented relationships between Tmax and rainfall would be significant (P <0.05) and offset (rainfall adjusted means would be different); that responses to rainfall would be the same (lines would be parallel - interaction between segments and rainfall would not be significant); that R^2_{adj} would be higher and that rainfall-domain residuals would be independent, normally distributed with equal variance across categories. Homogenisation changepoints are evaluated graphically and statistically in Table 1.

Each analysis in Table 1 examines: (i), previously archived homogenised HQ, Acv.1 and Acv.2 data and Marble Bar rainfall, and, (ii), breakpoint scenarios used to create each series. Thus, in the case of the Della-Marta *et al.* (2004) HQ dataset, rainfall alone explains 15.7% of variation in HQ Tmax which is a poorer fit than for $Tmax_{raw}$ (top panel) where rainfall explains 29.7% of variation. HQ_{shift} scenarios ($Tmax_{raw}$ ~ Sh_{HQ} + rainfall) are also a poorer fit compared to the Sh_{group} scenarios in the top panel (R^2_{adj} = 0.439 vs 0.503). As superscripts (^{a,b}) show no clear differences between segment means, although lines are mostly parallel, they are also coincident, which indicates the breakpoints do not reflect underlying site change effects on Tmax.

Finally, with the effect of rainfall removed, re-scaled HQ and ACORN-SAT residuals were not homogeneous as they should be if underlying changes were correctly specified (Figure 12). Homogenised data therefore still embed confounded signals – that due to rainfall and underlying step-changes resulting from maladjustments or biases imparted by the homogenisation process.

Table 1. Multiple linear regression and interaction analysis of breakpoints or shifts (Sh) identified in Figure 11 (a) verses those applied by High Quality (ShHQ), ACORN-SAT v.1 (ShAcv.1) and ACORN-SAT v.2 (ShAcv.2). Tmax refers to average annual Tmax (i.e., raw data). Graphs show free-fit regressions for respective data-segments, with segment means tabulated on the right. Superscripts identify means that are not significantly different (*P* <0.05) applying Turkey *P* value adjustments for pair-wise comparisons. Sem refers to the standard error of means. Each analysis presents goodness of fit (R^2_{adj}) of homogenised data (HQ, Acv.1 and Acv.2) with rainfall, and respective breakpoint scenarios, which were evaluated separately.

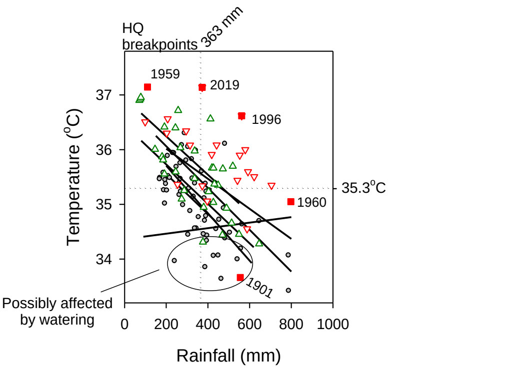


Tmax ~ sh_{group} + rain	Mean (°C)	Sem
Cool	35.03[a]	0.065
Intermediate	35.53[b]	0.093
Warm	35.95[c]	0.129

$Tmax_{raw}$ ~ rainfall R^2_{adj} = 0.297 (outliers ignored)

$Tmax_{raw}$ ~ Sh_{group} + rainfall R^2_{adj} = 0.503 (outliers ignored)

(For consistency, individual regressions were free-fit vis-à-vis Figure 7(c) to (e). Interaction was not significant therefore lines are statistically parallel (refer Figure 7(f)).

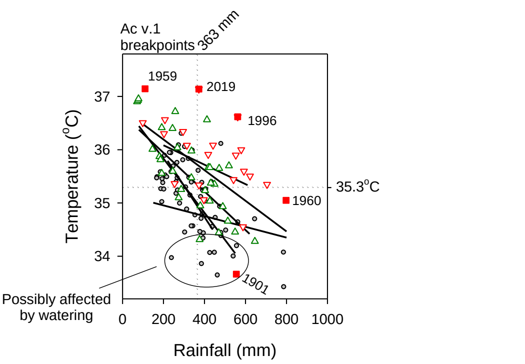


$Tmax_{raw}$ ~ sh_{HQ} + rain		
1	35.2[a,b]	0.143
2	35.7[a]	0.223
3	35.5[a]	0.132
4	35.0[a,b]	0.163
5	35.1[a,b]	0.148
6	34.8[b]	0.179
7	35.6[a]	0.138

$Tmax_{HQ}$ ~ rainfall R^2_{adj} = 0.157 (All data)

$Tmax_{raw}$ ~ Sh_{HQ} + rainfall = 0.439 (All data)

Interaction possibly significant (P = 0.10); note no clear differences between segment means.

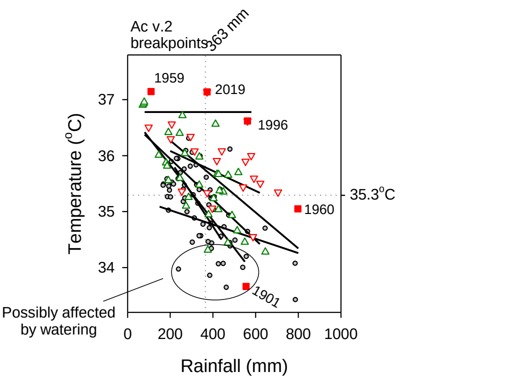


$Tmax_{raw}$ ~ sh_{Acv1} + rain		
1	35.4[a,b]	0.0942
2	35.1[a,c]	0.1294
3	34.8[c]	0.1489
4	35.1[a,b,c]	0.3072
5	35.8[b]	0.2046
6	35.7[b]	0.1295

$Tmax_{AcV.1}$ ~ rainfall R^2_{adj} = 0.226 (All data)

$Tmax_{raw}$ ~ ShAcv.1 + rainfall = 0.395 (All data)

Interaction possibly significant (P = 0.08); note the lack of discrimination between segments.



$Tmax_{raw}$ ~ sh_{Acv2} + rain		
1	35.4[a]	0.0916
2	35.1[a,b]	0.1317
3	34.8[b]	0.1371
4	35.1[a,b]	0.299
5	35.8[a]	0.1993
6	35.6[a]	0.1468
7	35.9[a,b]	0.3452

$Tmax_{AcV.2}$ ~ rainfall R^2_{adj} = 0.189 (All data)

$Tmax_{raw}$ ~ $Sh_{AcV.2}$ + rainfall = 0.408 (All data)

Interaction not significant (P = 0.17); note the lack of discrimination between segments.

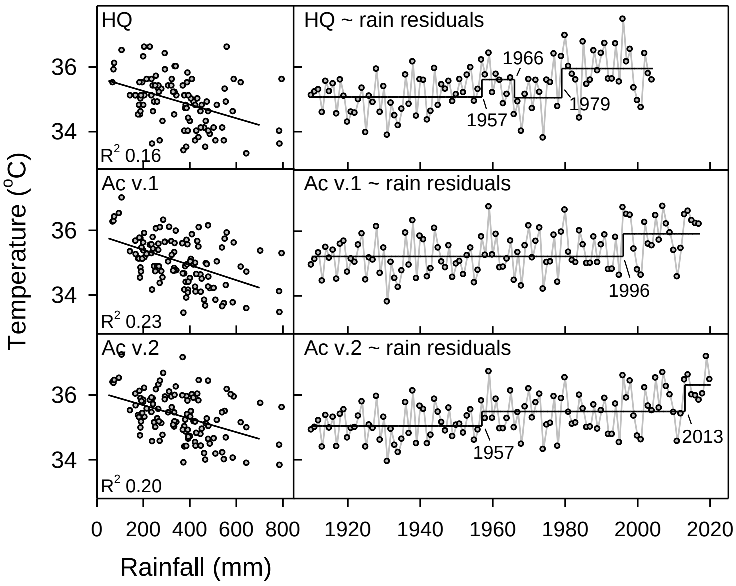


Figure 12. Residual (i.e., non-rainfall) inhomogeneities in homogenised High Quality (HQ), ACORN-SAT v.1 and ACORN-SAT v.2 datasets for Marble Bar. Although marble Bar data were rated poorly (R^2_{adj} for the equivalent Tmax ~ rainfall graph in Figure 5 was 0.297), rainfall explained less variation in homogenised Tmax than raw data. Step-changes in residuals indicate homogenised data were not homogeneous.

4. Discussion

> Methods developed by www.bomwatch.com.au show unequivocally that the climate of Marble Bar has not changed or warmed and that the Bureau's homogenisation methods are inherently faulty, unscientific and should be abandoned.
>
> Read on …

No mention was made in metadata that the original Marble Bar post office had suffered damage from cyclones in 1941 and 1947, that by 1947 it was riddled with termites (http://nla.gov.au/nla.news-article75087856) and that plans were drawn-up for the new post office in 1944. Land was acquired on the corner of Francis and Contest streets and the post office, residence and associated buildings (generator, battery room, fuel shed and garages) were constructed from 1947 to 1948 (Figure 13). The Marble Bar aerodrome also opened in 1948 within the reserve now occupied by the racecourse and a ministerial delegation was flown from Port Hedland by RAAF DC3 in August 1948 to perform the opening.

Unless it moved to temporary premises beforehand, the new post office had not been built in 1944 when based on a statistical test, Torok (1996) and consequently ACORN-SAT suggested the site moved. However, as the 1945 down-step also coincided with the reporting of whole°F (Figure 6) it may be due to post office staff consistently rounding data down to the nearest whole°F, replacement of a worn-out thermometer, or a combination of reasons (Torok listed 'old, poor screen' and 'move' but neither can be confirmed). Also, as the new post office was close to the boundary and the land slopes *uphill* from Francis Street it was unlikely that the site moved "up sloping ground towards the post office" in May 1967. More likely the screen was moved away from the telephone line exchange toward the residence where it was affected by shade and watering during an inspection five or-so years after the exchange was annexed to the to the post office in 1961.

The important point is that site-summary and ACORN-SAT metadata is entirely inadequate for determining if site conditions changed or not, and if they did, whether such changes materially impacted Tmax data. The use of naïve trend analysis alone ignores that moving a site, changing an instrument or the conditions under which observations were made does not change the climate, and should be discouraged.

Figure 13. Google Earth Pro satellite images showing the post office (PO), telephone line exchange (TLE), battery/linesman's office (B), line yard (LY), residence (Res), generator shed (Gen) and garage (G) facing Contest Street in August 2011 (left) and September 2013 after the battery/linesman's office was demolished.

Even though resources such as Google Earth Pro, digitised newspapers and NAA catalogues were unavailable in the early 1990s when Torok did his research, Nicholls and protégés including Blair Trewin and 'experts' who subsequently peer reviewed the Bureau's methods did not consider risks posed by incomplete and faulty metadata on identification of changepoints and therefore on the misattribution of trends as being due to CO_2 or something else. That prior to this research and despite three or more rounds of homogenisation following Torok (1993), BoM scientists were unaware of the 1948 post office highlights fundamental flaws in research underpinning the Bureau's methods.

Application of the First Law theorem (i.e., that Tmax depends on rainfall) to the problem of assessing the fitness or robustness of maximum temperature data provides an unambiguous, replicable means of investigating and apportioning variation due to the deterministic covariable, rainfall, and underlying changes in the data caused by missing variables. Lack of a relationship between Tmax and rainfall indicates the quality of data is questionable. At the other extreme, regardless of how noisy they are, should rainfall fully explain Tmax, residuals would be random, normally distributed with equal variance. However, should residuals embed systematic signals including cycles, trends or step-changes they are not homogeneous, meaning an additional factor is unexplained.

As the effect of rainfall has been statistically removed, residual step-changes (inhomogeneities) are indicative of underlying non-climate effects. Multiple linear regression (MLR) verifies that segmented relationships are offset, and that slope coefficients are the same. Compared to Parafield, Marble Bar data was considerably less precise and more complex, involving both negative and positive excursions from the long-term mean.

Lack of precision in the pooled relationship (R^2_{adj} = 0.503) shows quality of Marble Bar data is relatively poor. While much of the residual variation is synoptic (blows in with the wind), observations are also affected by the quality of the instruments, which seemed to have been rarely replaced, deterioration of the Stevenson screen, site effects such as watering and shade and placement of the current AWS southwest of the Roadhouse in 2000, where it is subject to dusty conditions and weathering. Although not confirmed by metadata, the step-change in 2015 is possibly due to replacement of a screen or refurbishment of sensors etc. Site summary metadata indicates the site was audited in January 2014 and the AWS was re-configured in November 2015. Destruction of local groundcover is likely to materially warm the site going forward (Figure 14).

Expected outcomes of covariate analyses are:

- Tmax is linearly, negatively and significantly correlated with rainfall.
- Rainfall-adjusted segment means are different (regressions are offset, not coincident).
- Interaction between breakpoints and rainfall is not significant (lines are parallel and therefore the segmented response to rainfall is the same).
- Residuals are independent, homogeneous, normally distributed with equal variance (that there is no 'hidden' residual signal).

Figure 14. A January 2022 Google Earth Pro satellite image shows the site scalped of topsoil and vegetation (left), and, right, a general view taken in November 2022 (contributed).

As Tmax and rainfall is the same for each scenario in Table 1, significances and goodness of fit (R^2_{adj}) are inter-comparable and also with analysis presented in Figure 5. However, in contrast to Figure 5, no clear differences were evident between segments defined by HQ or ACORN-SAT changepoints. Homogenisation changepoints were therefore miss-specified.

Although free-fit regressions in Table 1 were statistically parallel (interaction was not significant), segmented means were mostly the same. Thus, segmented relationships were largely coincident. In addition, relationships between homogenised data and rainfall (homogenised Tmax ~ rainfall) and raw Tmax factored on homogenisation breakpoints (Tmax ~ $Sh_{breakpoint}$ + rainfall) resulted in relatively poor fits. For example, while rainfall explained 29.7% of raw Tmax variation in Figure 5(a), applied to the same raw data ACORN-SAT v.2 only explained 18.9% of Tmax variation. Further, factored on the seven ACORN-SAT v.2 changepoints multiple linear regression explained only 40.8% of Tmax variation *vs*. 50.3% for the three-step model. Finally, due to embedded residual step-changes, homogenised data were not homogeneous (Figure 11), consequently trends depicted by each series were spurious.

The four most important experimental outcomes were:

- Consistent with the First Law Theorem, maximum temperature depends on rainfall and the drier it is the warmer it gets. However, while Tmax data was cooler before a Stevenson screen was supplied in 1911 and the site moved to warmer/cooler locations within the town, there is no evidence that the fundamental relationship between Tmax and rainfall has changed since records commenced in 1901. As site changes and rainfall simultaneously explained 50.3% of variation in Tmax ($R^2_{adj} = 0.503$) and there were no residual systematic signals, overall data quality is poor.

  Factors include poor site control (damage and deterioration of equipment, watering, shade etc), imprecision in observations (rounding-down from 1945 to 1960 and at other times mainly during the Fahrenheit era) and site moves and changes that were not documented in site-summary or ACORN-SAT metadata (supply of the Stevenson screen, the move from the old to the new post office, then to the licenced post office, replacement of the 230-litre screen with a 60-litre one, and recent earthworks in the vicinity of the site likely to affect observations (Figure 14).

- Homogenisation methods used by the Bureau, including most recently, those developed by Blair Trewin do not reflect changes in Tmax data. The expectation that homogenisation changepoints would better represent changes determined using objective statistical methods and that resulting data would be homogeneous were not realised. In addition, changes in properties such as the frequency of whole°F, Hi_N/Lo_N ratios, and centred on 2015, probability density functions and differences in cumulative distributions, lend little support to the global warming - climate change thesis.

- Picking homogenisation changepoints based on faulty metadata, and making adjustments using highly-correlated comparator datasets that likely embed parallel faults, creates trends in homogenised data that are unrelated to the climate.
- As homogenisation methods used by the Bureau are biased by the faulty selection of changepoints and the choice of reference series based on correlation of first-differences, and as resulting homogenised data are **not** homogeneous, the ACORN-SAT project is poorly conceived, unscientific and should be abandoned.

Disclaimer

Rules relating to scientific conduct by members of the Australian Meteorological and Oceanographic Society of which the Author is a member, specifically state that "*Members involved in scientific activities should base those activities on sound scientific principles*" and that "*Plagiarism, fabrication or falsification of data, and other misleading behaviour are all unacceptable*" (https://www.amos.org.au/about/rules-and-regulations/amos-code-of-conduct-2/).

Unethical scientific practices including the homogenisation of data to support the global warming narrative undermines trust and is not in the public interest. While the persons mentioned or critiqued may be upstanding citizens, which is not in question, the problem lies with their approach to data, use of poor data or their portrayal of data in their cited and referenceable publications as representing facts that are unsubstantiated, statistically questionable or not true. The debate is therefore a scientific one, not a personal one.

Acknowledgements

Impetus for this research arose from observing that students from primary school to university and beyond are actively being convinced that due to global warming caused by CO_2, the world is facing a tipping-point during their lifetimes, which is not true. Overseen by Neville Nicholls, who commenced as a cadet meteorologist with the Bureau of Meteorology in 1970[1] and later in 1986, was a member of the World Climate Research Programme when Dr John Zillman was Australia's permanent WMO representative, Nicholls oversaw BoM scientist Simon Torok's PhD, which underpinned much of the Bureau's subsequent homogenisation effort. A contributor to the World Economic Forum[2], Nicholls is currently Emeritus Professor at Monash University and he is acknowledged for stirring my interest in the dark-art of data homogenisation. The damage wrought by elite scientists to the integrity of science in Australia, the national economy and the wellbeing and job prospects of future generations in the name of climate change is incalculable.

Development of decimal fraction frequencies as a precision metric resulted from discussions with Chris Gillham (https://www.waclimate.net/). Others who contributed ideas and discussion included Geoffrey Sherrington, Ken Stewart (https://kenskingdom.wordpress.com/) and Tom Berger (http://www.elastictruth.com/). David Mason-Jones is gratefully acknowledged for providing invaluable editorial assistance.

Preferred citation:

Johnston, Bill 2021. Is homogenisation of Australian temperature data any good? Part 2. Marble Bar, the warmest place in Australia. http://www.bomwatch.com.au/ 21 pp.

Dr Bill Johnston

01 December 2022

---

[1] https://www.eoas.info/biogs/P003289b.htm

[2] https://www.weforum.org/agenda/authors/neville-nicholls

Appendix 1.

Sub-agencies or points of contact within the Department of Climate Change, Energy, the Environment and Water.

Australian Alps National Parks
Australian Antarctic Program
Australian Climate Service
Australian Energy Infrastructure Commissioner
Australian Institute of Marine Science
Australian National Botanic Gardens
Australian Renewable Energy Agency (ARENA)
Australia's National Greenhouse Accounts
Australia's Nature Hub
Bureau of Meteorology
Clean Energy Finance Corporation
Clean Energy Regulator
Climate Active Climate Change Authority
Cockatoo Island
Commercial Buildings Disclosure
Energy Ministers
Energy Rating
energy.gov.au
Environmental-Economic Accounting
Great Barrier Reef Marine Park Authority
Independent review of the Environment Protection and Biodiversity Conservation Act
Inspector-General of Water Compliance
National Environment Protection Council
National Pollutant Inventory
National Water Grid Authority
Nationwide House Energy Rating Scheme (NatHERS)
North Queensland Water Infrastructure Authority
Parks Australia
PFAS Australian Information Portal
Physical Environment Analysis Network
Snowy Hydro Limited
State of the Environment 2016
Sustainable Development Goals
Sydney Harbour Federation Trust
Water Quality Australia
Water Rating
Your Home