Are Australia's automatic weather stations any good?

Part 4: Effect on temperature extremes

Dr. Bill Johnston[1]

scientist@bomwatch.com.au

Summary

Maximum temperature data for Charleville aerodrome from 1943, showed no trend or change attributable to the climate. While aerial photographs, and maps and plans located the original site on the northern side of the aerodrome south of the World War II operations precinct, it was not identified by site-summary or Australian Climate Observations Network-Surface Air Temperature (ACORN-SAT) metadata.

Data from 1943 to 1955 were evidentially reconstructed using rainfall as the covariate, which erased the effect on Tmax of the devastating, long-running inter-War drought. The site (and staff) probably moved from the Aeradio office to the Flight Services Unit near the terminal in 1956 (not 1961 as claimed by BoM). Step-changes in rainfall residuals in 1979, which could not be explained; in 1997, after the automatic weather station (AWS) became the primary instrument; and in 2013 following building new facilities at the current site, caused Tmax to increase 1.75°C (rainfall adjusted). Rainfall reduced Tmax 0.34°C/100 mm and step-changes and rainfall together explained 76.6% of Tmax variation.

Against the background of a non-changing climate, the situation with instrument biases and site changes on the frequency of daily Tmax measured in real-time was more complex. Firstly, because extremes reported by the Bureau are inclusive of site and instrument changes; secondly, the frequency of extremes may increase without affecting mean-Tmax; and, finally, because sensitive, rapid-sampling instruments operating in 60-litre screens are more likely to detect spikes that are not representative of ambient conditions. Further, it appears that daily temperatures approaching about 36°C were moderated to avoid unrealistic numbers of values exceeding 37.8°C (100°F) and also that the highest absolute values were probably instrument-related artefacts.

Ten maximum temperature datasets used to homogenise Charleville for a Tmax step-change in 1996 were also examined. Site-summary metadata was inadequate. No sites were homogeneous; correlation with Charleville in the decades around 1996 was due predominantly to parallel discontinuities caused by site changes such as new buildings and phone towers; undocumented moves; replacement of 230-litre screens with 60-litre ones etc. Use of correlated faulty data to correct faults in ACORN-SAT data has no scientific or statistical merit and should be abandoned.

As there was no change in the climate at either Charleville or the 10-comparator sites there can be no real trend or change in the frequency of extremes either. Calculating Australia's long-term climate trajectory using ACORN-SAT data that has been biased by homogenisation, while simultaneously claiming that extremes and 'records' measured by AWS and 60-litre screens are comparable with long-term data measured using different equipment somewhere else, is disingenuous and damages the credibility of all concerned.

---

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

1. Introduction

Analysis of climate data presents many challenges. For example, coordinates of original meteorological enclosures at 'high quality' Australian Climate Observations Network-Surface Air Temperature (ACORN-SAT) sites at Cairns, Townsville and Rockhampton, which are adjacent to the Great Barrier Reef; and Amberley, Cape Leeuwin and Rutherglen could not be confirmed by site-summary or ACORN-SAT metadata. Factors affecting the quality and comparability of data such as the types of Stevenson screens used, when they were replaced, condition of the site and the quality of instruments was not reported. More recent changes such as when electronic automatic weather station (AWS) probes replaced liquid-in-glass thermometers; 60-litre screens replaced 230-litre ones; reductions in the frequency of site visits and maintenance as staff were made redundant, and possible changes in the way data is processed is also routinely overlooked. As all or any of such changes may affect the quality and reliability of current data compared to historic data, and with temperature records being broken seemingly every day, it is timely to investigate the effect of AWS and 60-litre Stevenson screens on temperature extremes.

The overarching question is: **Are Australia's automatic weather stations any good?**

2. Case study Charleville Qld. (BoM ID 044021)

According to the most recent (v.2) ACORN-SAT metadata[1], observations at Charleville Airport commenced in 1943 and although coordinates of the original site were not provided it was said to have moved in or shortly before April 1961. An AWS installed in August 1990 relocated 150 m east on 12 September, where it (or a replacement) became the primary instrument on 1 November 1996. The site moved to its current location in October 2003. Complicating understanding of the data is that a 230-litre Stevenson screen was used until 13 November 2003, while a small 60-litre screen was installed four years earlier on 8 December 1999[2]. There was no concurrent data for those two screens. Comparing overlap data for the AWS and 60-litre screen at the previous site from October 2003 to February 2006 (site 044221) with that for the current AWS and small screen 500 m away is unlikely to be insightful.

During WW-II the Charleville aerodrome was used by the United States Army Air Force (USAAF) as a heavy-bomber transit base. A 1945 water supply and septic-tank plan located at the National Archives of Australia (NAA ID 1972125, Appendix 1) showed the original enclosure and cloud-base searchlight was about 160 m southwest of the Aeradio office, which was on the northern side of the aerodrome opposite the racecourse (Figure 1).

The enclosure was visible in a high altitude aerial photograph taken in December 1954, together with the 1922 Qantas hangar, fuel-farm, powerhouse, Aeradio office and control tower, which were beside the apron at the western end of the heavy-duty E-W runway (Figure 2). NAA files corroborate that a new Radio Aids building planned in 1950 opened near the airport entrance in 1955. Meanwhile a radio-met building also planned in 1950 was extended in 1955 to accommodate additional (met) staff. Bureau staff shared the office, which was about 60 m east of the current terminal, with the Department of Civil Aviation (DCA) Flight Services Unit (FSU). Later, in 1986, a satellite communications (SatCom) unit was installed nearby.

[1] http://www.bom.gov.au/climate/data/acorn-sat/stations/#/23090

[2] https://cawcr.gov.au/technical-reports/CTR_049.pdf

Figure 1. The earliest Google Earth Pro satellite (17 June 2006) overlaid by a rescaled section of the 1945 water supply extension and septic tank installation diagram (Appendix 1) which shows various operational precinct (OP) buildings including the control tower and radio-met office, the fuel-farm (F), original Qantas hangar (H), which was reportedly demolished in 1956 and the meteorological enclosure and cloud-base searchlight (X). Staff moved to the Flight Services Unit (fsu) probably in 1955/56 and the SatCom unit (sat) was installed in 1986. It is likely the site moved to S_1 also in 1956. A Micromac AWS was installed at S_2 in 1990; which was probably replaced before it became the primary instrument from 1 November 1996. The site moved to S_3 in October 2003. A 60-litre Stevenson screen was installed at S_2 on 8 December 1999; however, the former 230-litre screen was also used until 13 November 2003.

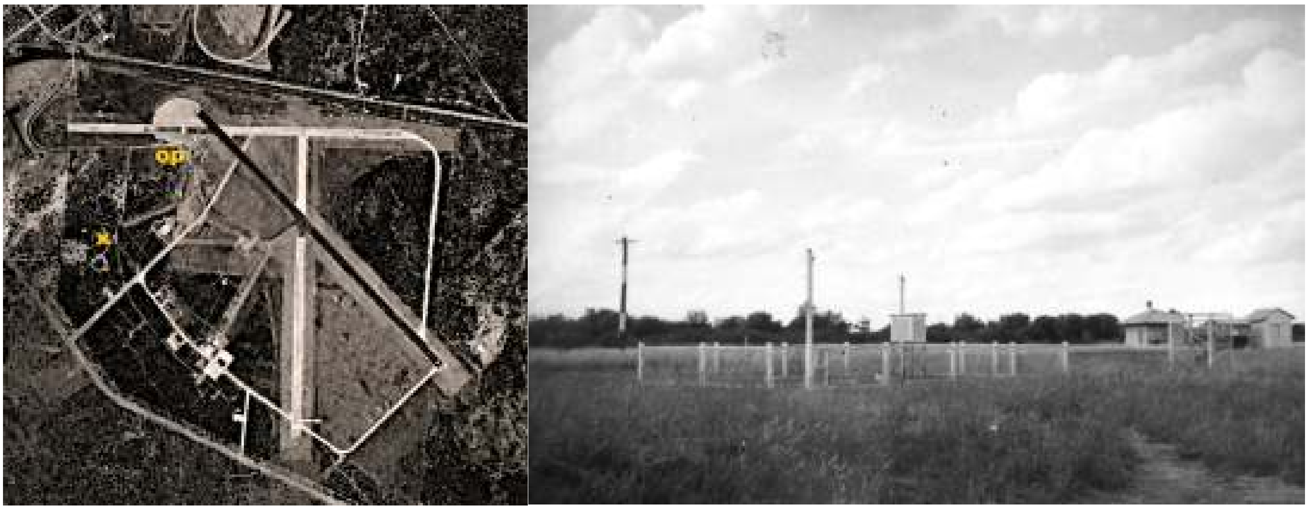


Figure 2. High altitude (25,000 feet (7.6 km)) aerial photograph of the former WW-II USAAF heavy-bomber transit base at Charleville aerodrome taken on 15 December 1954 (left) shows the operations precinct (OP) adjacent to the apron and turning circle and the original enclosure at 'x'. Also photographed in 1954, the weather station with a standard 230-litre Stevenson screen opening to the north, and the cloud-base searchlight to its right. Buildings in the background (L to R) include the power-house and balloon filling (aerosonde) shed with the Aeradio office behind. (Photographs courtesy of the National Library of Australia and the Civil Aviation Historical Society Airways Museum at Essendon.)

It appears from the data that the met-enclosure moved about 580 m southeast in 1956 (not 1961[1]), which was when the original Qantas hangar was reportedly demolished. An A-pan evaporimeter was installed in 1966 and in 1990 after the Micromac AWS was commissioned and about the time the FSU was disbanded, Bureau staff were consolidated in the 1955 Radio Aids building near the airport entrance. While a 1992 aerial photograph shows the 1956 site apparently intact with the AWS operating about 150 m to the east, Google Earth Pro Satellite images shows that by 11 June 2012 the FSU-met office and SatCom were demolished, a new meteorological office, SatCom, car park and autosonde were built about 40 m from the current site and that the former Radio Aids site was re-purposed as the Cosmos Centre.

Relocations and site changes including changed observing practices, replacement of 230-litre Stevenson screens with 60-litre ones, liquid-in-glass minimum and maximum thermometers (Tmin and Tmax) with rapid-sampling AWS sensors and changes in the immediate environment can affect both the 'location' of data (i.e. its level relative to the long-term mean) and day-to-day variability including extremes. However, as some documented changes may not impact on data, while others that might, may not be documented, the problem of assessing trend and change is best approached using unbiased, independent statistical methods. To recap methodology explained previously:

- Naïve linear regression is used to assess significance and strength of the expected negative linear relationship between average maximum temperature (Tmax) and annual rainfall. The coefficient (^{o}C/100 mm of rainfall), its significance (*P*) and goodness of fit or proportion of variance explained (R^2_{adj}) objectively indicate compliance with the First Law theorem. A relationship may be significant but imprecise – it may explain only a small proportion of variation, for instance.
- Re-scaled residuals (i.e. the non-rainfall portion of the signal) rescaled for convenience by adding the dataset grand-mean are evaluated for discontinuities (step-changes or shifts in the mean), which are indicative of embedded site-change effects (site inhomogeneities).
- With segments defined as categories (factors) and rainfall the covariate, step-changes are validated using multiple linear regression (MLR).
- It is expected that rainfall-adjusted segment means will be different (segmented regressions be offset); that slopes are parallel (interaction is not significant thus the segmented response to rainfall is the same); and that MLR-residuals are normally distributed, independent with equal variance across categories.

  Analysis of individual segments evaluates goodness of fit and identifies outliers while within-segment analysis verifies that data consist of untrending segments bounded by step-changes, thus further validating the step-change model.
- Having accounted for site changes and rainfall simultaneously, MLR residuals are examined for trend and change attributable to the general climate.

Naïve linear regression (Figure 3(a)) found that while the relationship with rainfall was negative and highly significant (*P*<0.001), the covariate explained only 48.1% of Tmax variation. Low precision was due to confounded site-related discontinuities in 1979, 1997 and 2013 (b). Data for 1969 and 1970 were affected by missing observations/yr but were not excluded, while data for 1973, which was out-of-range and skewed the response, was omitted from further analysis.

---

[1] CBoM - History of Charleville Meteorological Office (no longer available on-line) (Appendix 2).

ACORN-SAT metadata stated that "Neighbours rather than the overlapping data are used to merge the two main Charleville sites" but the source of the data, the period over which adjustments were made and the methodology used were not provided. Segmented regressions (Figure 3(c) to (g)) strongly suggest data from 1943 to 1955 were reconstructed (predicted) using rainfall as the covariate. As R^2_{adj} in Figure 3(c) = 0.961 and pre- and post-1956 group means were the same (27.7°C) it is highly likely that pre-1956 data were predicted using 1957 to 1978 data. Low precision for the segment from 1979 to 1996 (R^2_{adj} = 0.271 *vs*. 0.431) may have some bearing on the 1979 step-change.

Multiple linear regression (Figure 3 (h)) showed Tmax declined 0.34°C/100 mm rainfall, and rainfall and step-changes together explained 76.6% of Tmax variation (*vs*. 48.1% for rainfall alone). Step-changes caused Tmax to increase by a rainfall-adjusted difference of 1.75°C overall (light-font in (c) to (g)). As segmented time-domain trends were not significant (P >0.05), the overall raw-data trend of 0.26°C/decade was fully explained by rainfall and step-changes and there was no evidence that the climate had changed or warmed.

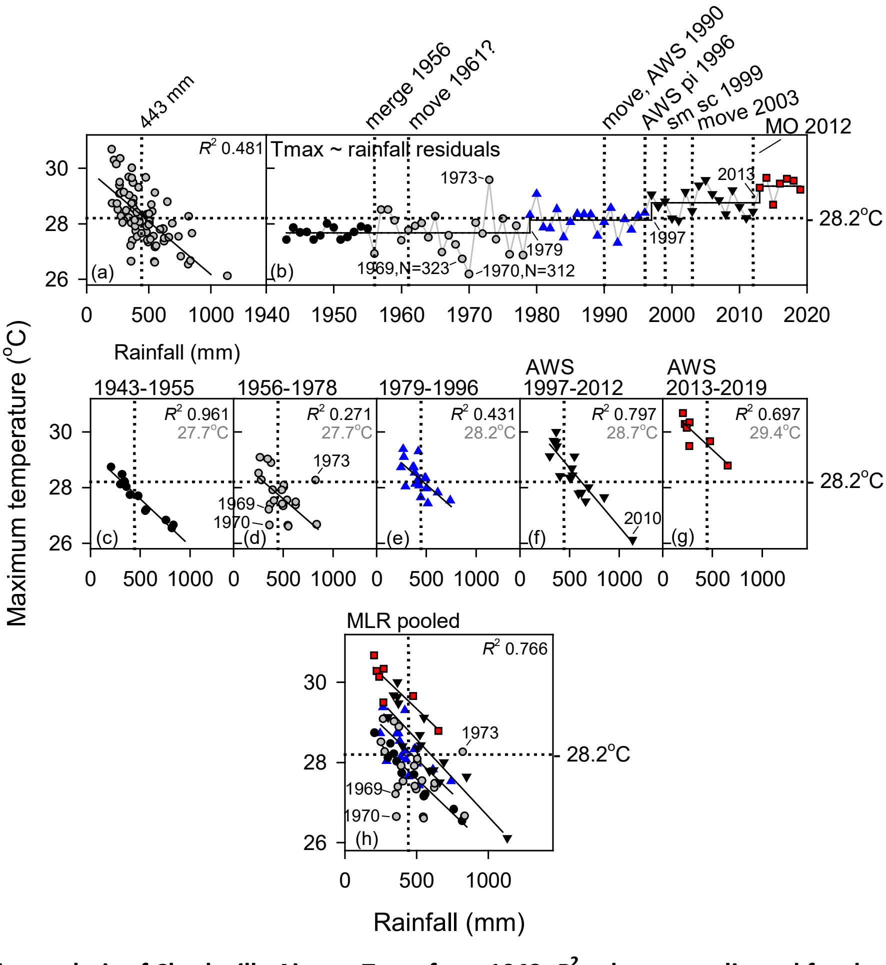


Figure 3. Composite analysis of Charleville Airport Tmax from 1943. R^2 values are adjusted for the number of terms and observations in respective models. The step-change in residuals in 1979 appeared to be associated with a precision change; that in 1997 follows the AWS becoming the primary instrument on 1 November 1996; while the change in 2013 followed disturbance of the site and surrounds when a new met-office, SatCom, car-park and autosonde (balloon launcher) were built. (A 60-litre screen was installed at the previous site on 8 December 1999).

Understanding evolution of the data requires metadata of a high standard. However, despite assurances implied or otherwise, including by in-house peer-reviewers and statements in peer-reviewed journals that 'high-quality' ACORN-SAT data reflect the climate alone, site changes at Charleville aerodrome have not been thoroughly researched. Cross-referenced by Google Earth

Pro satellite images, the original site (at about Latitude -26.4103°, Longitude 146.2561) was southwest of the former Aeradio office and control tower whose foundations are visible in the earliest (17 June 2006) satellite image (at Latitude –26.4092°, Longitude 146.2571°). The image also shows contamination after fuel storage tanks were removed and that the WW-II apron, turning circle and heavy-duty runway had been torn-up.

Confirmed by the 1954 aerial photograph, the earlier 1945 water supply and septic tank plan (Appendix 1) showed the operations precinct layout including the location of the original enclosure. Also it is clear from Figure 3(b) and (c) that data before 1957 were reconstructed; the site re-located south of the FSU office probably in 1956 (not 1961) and it was apparent that Tmax data prior to 1979 were either haphazardly observed or in-filled *ad-hoc* to imitate data that may not have been available. Although Bureau staff allegedly undertook daily observations throughout, there is no explanation for missing data in 1969 and 1970.

3. Effect of AWS and 60-litre Stevenson screens on temperature extremes

3.1 The general climate

Charleville experiences a dry continental/subtropical climate with clear skies, cold nights and dry mild days in winter (20-22°C daytime) and hot, humid, stormy weather, with daytime temperatures often exceeding 35°C in summer. Median annual rainfall (1943-2019) is 442 mm, with 50% of years receiving between 354 mm and 551 mm. The distribution of monthly rainfall relative to potential evaporation is shown in Figure 4.

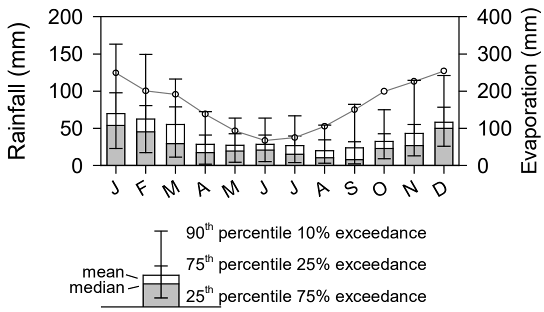


Figure 4. Monthly rainfall distribution and exceedances for Charleville (1900-2019) compared with potential evaporation (0.8 * average A-pan evaporation). Infrequent heavy rain causes mean rainfall to exceed median rainfall (or 50^{th} percentile); also, indicated by the 90^{th} percentile, rainfall in the upper 10% of years is >double that of monthly averages. Noting the different scales, it is unlikely that rainfall in any month would exceed potential evaporation.

Monthly statistics and summaries do not provide a dynamic view of the long-term climate and its propensity to oscillate from runs of relatively moist, cool years to sequences that are dry and warm. Most of the variation in annual Tmax is explained by the water cycle and, consistent with the First Law theorem, the drier it is the warmer it gets (Figure 3). However, as potential evaporation during summer is factors higher, it is only in exceptional years that rainfall could generate appreciable local runoff. For instance, to be effective, that is, to satisfy the preceding month's soil water deficit and current evaporative demand, January rainfall would need to exceed around 300 mm, which as indicated by Figure 4 is highly unlikely.

Calculated as: water stored in the soil from previous months plus current rainfall, minus current month's evaporative demand, the water balance derives a rolling estimate of the status of soil water, which cannot be less than zero. Assumptions are that: (i), plant available soil water is bounded by zero and 100 mm; (ii), as the soil progressively dries the transfer rate to the atmosphere (evapotranspiration) slows; and (iii), rainfall in excess of the assumed soil water holding capacity (calculated as antecedent stored soil water + current monthly rainfall - evaporative demand) is lost as combined runoff/deep-drainage.

Figure 5 shows the climate transitions rapidly from sequences of relatively bountiful years to periods of moisture deficiency every 5 to 7 years, somewhat in synchrony with the El Niño Southern Oscillation (SOI). Negative SOI values (El Niño) are associated with dry years and *vice versa* for La Niña; however, the relationship is qualitative not predictive. Some La Niña years are relatively dry for example, and evidenced by heavy rainfall and flooding in 2010, some El Niño episodes are anomalously wet. Transitions such as in 1924, 1949, 1972 and 2010 are stochastic/unpredictable.

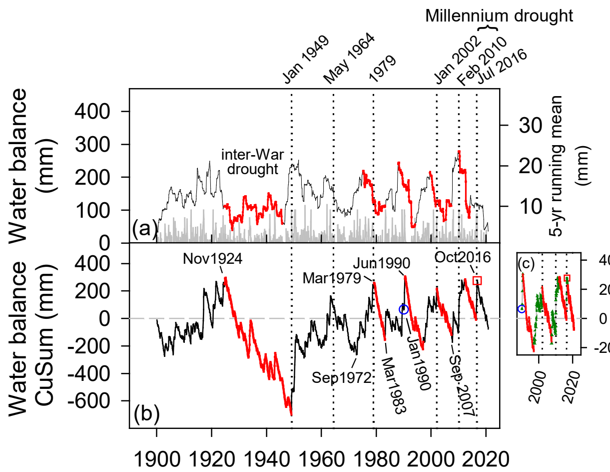


Figure 5. Monthly soil water balance in (a) assumes that the available water capacity (AWC) of the soil is between zero and 100 mm; that potential evaporation is 0.8 times average A-pan evaporation and that evaporation/transpiration declines stepwise from 0.8E_{pan} proportionally as AWC declines below thresholds of 40%, and 20%. Rainfall + residual moisture from the previous month – evapotranspiration, which exceeds AWC is assumed lost as undefined surpluses. Shown by the 5-year running mean (right-hand axis in (a)) droughts of varying intensity/duration occur irregularly every 5 to 7 years. The CuSum plot in (b) shows years that are cumulatively moist (the curve ascends), cumulatively dry (it declines) or about average (it fluctuates horizontally). The inter-War drought from 1924 to 1949 was the longest and most devastating since 1900 and it appears from a water balance perspective that the Millennium drought may have commenced in July 1999 and ended in September 2007. Highlighted for the 30-year period since January 1990 in (c), runs of average to above average rainfall and positive water balances are unusual. Drought periods mentioned in the text are highlighted.

The water balance depicts interaction between rainfall and the environment as a continuous process. Prior or antecedent rainfall may condition following months to generate excessive runoff or *vice versa*; however, water surpluses are rare and typically due to extreme events - tropical cyclones or intense monsoonal troughs that affect wide areas of central Queensland at the same time, and which result in widespread flooding of the Warrego and other rivers that drain to the Murray-Darling and Lake Eyre basins and the Carpentaria and Coral Sea coasts. Of the 16 calculated surpluses since 1900, 13 have occurred since the inter-War drought ended in 1949 while the two most notable events in April 1990 and January/February 2010 accounted for 38% of the estimated total. There is no indication that the severity or frequency of drought is increasing.

3.2 Temperature extremes

Extreme temperatures are often defined as counts or percentages of daily observations per year that exceed thresholds such as 37.8°C (100°F) or 40°C; or, as percentages within classes, for example, 5°C increments from 10°C to 40°C or variations thereof; or percentages above and below statistical quantiles such as dataset medians or quartiles. As the number of daily values per year is effectively fixed, percentage increases in values above a threshold must be offset by reductions in other classes. Percentages of values/year that exceed the dataset median for example must mirror negative deviations. While other metrics may include the number of *consecutive* days/year above thresholds or the calendar date that threshold values are exceeded, except for statistical quantiles that describe distributional properties, thresholds are arbitrary and usually rounded to a convenient number such as 30°C or 40°C. Some scientists have proposed heat-tolerance, comfort and health indices as climate metrics; however, 'feelings' about the climate are subjective, may vary from place to place and can be manipulated to stir fear about the climate.

The overarching question for this section is whether changes and trends in temperature extremes are related to factors such as site relocations, the staged replacement of 230-litre Stevenson screens with 60-litre ones; the introduction of automatic weather stations, which became primary instruments from 1 November 1996; and the dismissal of staff who undertook observations, kept records, checked equipment and maintained sites to a high standard; or whether they reflect the true climate.

3.2.1 Frequency within classes

While frequency within classes (>5 to 10°C; 11 to 15°C; etc. to >40°C) is expected to be random or may possibly trend, unexpected step-changes in means and changes in apparent variation more likely indicate data issues i.e. it is not expected that data consist of non-trending segments or that step-changes have no apparent cause. Figure 6 illustrates the problem for Tmax data within frequency classes 15 to 20°C, which stepped-down abruptly in 1984 and 2017 *vs*. the frequency of values exceeding 37.8°C, which stepped-up in 1979 and 2013.

As daily data are not necessarily uniformly distributed across classes or temporally equally precise, within-class comparisons poorly portray time-wise changes and trends. There is no reason that in Figure 6, increased numbers of values >37.8°C should occur at the expense of those within 15 to 20°C, but not in other classes such as 25 to 30°C for instance. Also, inhomogeneities in subjectively allocated classes are not conveniently analysed simultaneously and the broader the categories the more homogeneous the response may appear.

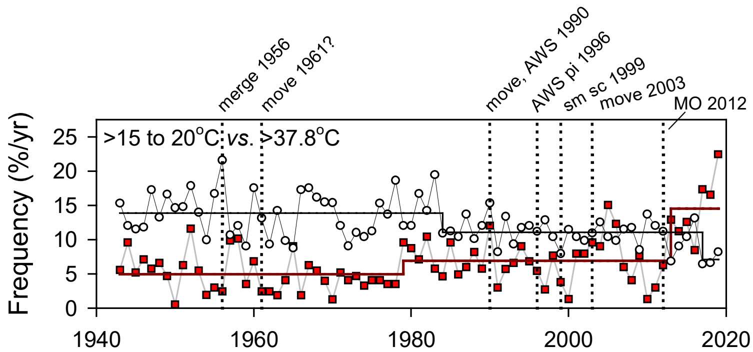


Figure 6. Sustained deviations in the frequency of daily Tmax within fixed temperature ranges must be offset by changes across all or other specific classes. It was unexpected that the frequency of Tmax values exceeding 37.8°C (100°F) (red squares) abruptly increased 2.0% in 1979 and 7.6% in 2013 and that the up-steps were noticeably offset by reductions in the frequency of values within the range of >15 to 20°C of 2.8% in 1984 and 4.0% in 2017 (open circles).

3.2.2 Frequency of deviations from the median

The annual frequency of daily Tmax greater than the daily dataset median of 28.5°C ($Tmax_{>med}$) mirrors frequencies less than the median ($Tmax_{<med}$). Deviations are expected to be random, with

dry years being warmer and *vice versa*; while sustained changes (step-changes or shifts) result from site and/or rainfall changes acting together (Figure 7). As influential factors are confounded, it is not possible to reconcile the shift towards warmer temperatures (and the concurrent reduction in the frequency of cooler days) in 1979 or 2013. Although satellite imagery suggests the 2013 change may be related to disturbances during construction of the new meteorological office, a drying phase also commenced in June 2012 (Figure 5).

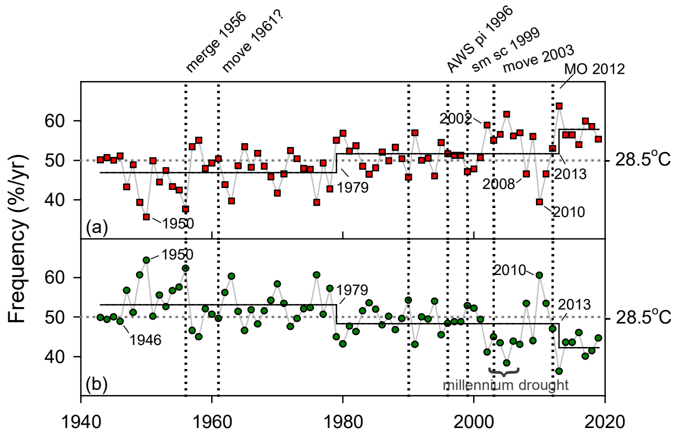


Figure 7. The frequency%/yr of daily values greater than the dataset median (28.5°C) in (a) mirror the frequency of those less than the median (b). The step-change in 1979, which aligned with the start of the 1979 to 1982 drought became embedded in the dataset even though rainfall was near average or above from 1983 to commencement of the next drying cycle in 1990.

3.2.3 Rainfall adjusted frequency

As outlined in Section 2, for a rainfall dependent variable, naïve regression partitions total variation into that due to the covariate and the residual portion that is unexplained by the relationship. Should rainfall alone explain $Tmax_{>med}$, residuals would be normally distributed, random and homogeneous. However, inhomogeneities in residuals indicate a second influential variable was overlooked. As data are equi-ranged and symmetrical about the median they are amenable to being analysed directly using linear regression without transformation.

Applying the same protocols used previously (*vis-à-vis* Figure 3) the highly significant relationship with rainfall (P <0.001) explained 45.9% of variation in $Tmax_{>med}$ (Figure 8(a)). However, step-changes in residuals (b) indicated presence of an underlying factor. While segmented regressions were significant (P <0.05), variation explained (R^2_{adj}) was variable ((c) to (e)). Importantly, ignoring the 1997 datapoint, which skewed the 1979 to 2001 response, segmented regressions appeared to be parallel and offset relative to the *x-y* data centroid (indicated by dotted lines).

Multiple linear regression verified that the response to rainfall was the same across categories and that rainfall-adjusted segment means were different (Figure 8(f)). With the exception of outliers identified previously, MLR residuals were zero-centred, independent, normally distributed and showed no unexplained trend or change.

Excluding data for 1997, which was apparently faulty, the overall $Tmax_{>med}$ trend of 0.133 %/yr consisted of three non-trending segments. Thus while the cause of the 1979 step-change is obscure, the 7% increase in observations >$Tmax_{>med}$ since 1943 was related to site or instrument changes, not the climate.

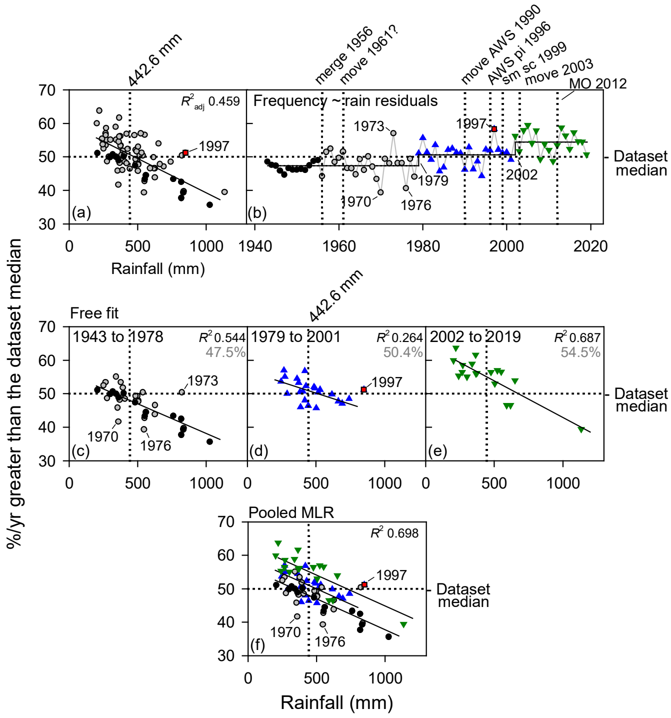


Figure 8. Ignoring out-of-range data for 1997, which skewed the 1979 to 2001 free-fit relationship; but including outliers in 1970, 1973 and 1976, rainfall explained 45.9% of variation in the frequency of daily values (%/yr) that exceeded the dataset median (28.5°C) (a). However, due to step-changes in 1979 and 2002 residuals were not homogeneous (b). Slopes were similar and offset relative to the data centroid (indicated by dotted lines) ((c) to (e)). Evaluated simultaneously, segmented responses to rainfall (f) were the same (slopes were homogeneous) and rainfall adjusted segment means (faint-font in (c) to (e)) were different (*P* <0.001). Site changes in 1979 and 2001 caused the frequency of $Tmax_{>med}$ to increase 7% overall, leaving no trend or change attributable to the climate.

3.2.4 Relative changes in the tails of daily Tmax distributions

As extremes may be cool or warm (not just over-range) and they may occur on any day of the year, counts (N) of daily values/yr less than the 5th and greater than the 95th day-of-year (1 to 366) percentiles (Lo and Hi respectively; Figure 9) also reflect trend and change in extremes. Determined by the dataset and inclusive of all observations, the 2-sided metric is objective and robust to missing observations. While runs of cool or warm years are expected, persistent non-random deviations are due to site changes and rainfall acting together. Expressed as ratios (LoN/HiN), the log transformation causes data to be symmetrical about the transformed mean, which allows parametric statistical methods to be applied without bias.

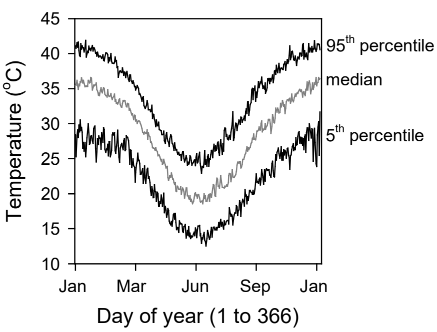


Figure 9. Day-of-year percentiles define robust thresholds for identifying outliers. By definition 5% of all daily values are less than the 5th and 5% exceed the 95th day-of-year percentiles (LoN and HiN).

Predominance of Lo-extremes prior to 1958 shows reconstructed data was cooler than after the site moved to the FSU/met office (Figure 10). While missing observations suggested something happened between 1969 and 1970 and data for 1970, 1973 and 1976 were out-of-range, Hi/Lo ratios were random until 2002 immediately before the site moved in 2003 (Figure 10(b)). The result suggested a small screen response and that raw data were adjusted in-house using overlap data from 13 November 2003 to 28 February 2006 for the previous site (ID 44221).

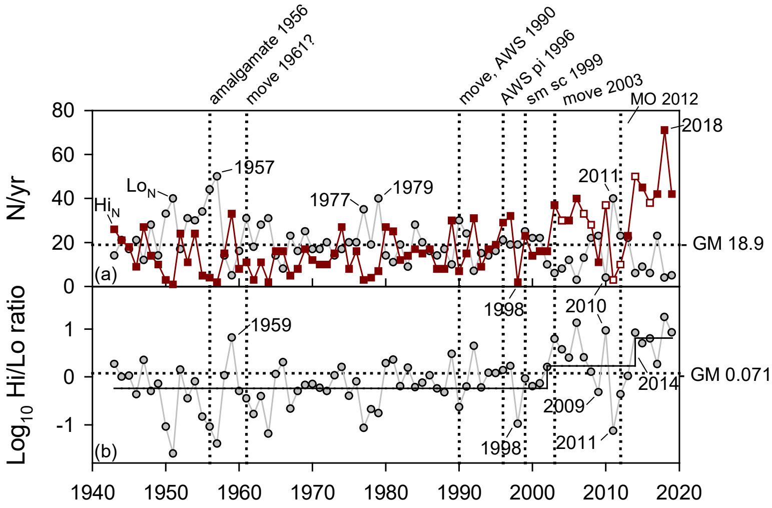


Figure 10. 5.26% of daily Tmax observations/yr were less than the 5th (Lo, grey circles) and 5.07% exceeded the 95th day-of-year percentile (Hi, red squares) (*vs*. 5% expected). Numbers (N) above and below the median are expected average (0.05*365.25) = 18.26 observations/yr. (The grand-mean (GM) across all observations was 18.9/yr.) Although clustered by runs of moist or dry years, variation is expected to be random. Moving the site in 2003 and the strengthening Millennium drought (July 1999 to September 2007) increased numbers of Hi-extremes from about 16 to 36/yr, while of Lo-extremes declined from 18 to <10/yr. Despite near average rainfall in eight of the 16 years since 2003 (unfilled symbols in (a)), upper-range extremes became permanently elevated after 2003.

3.2.5 Percentile differences

Following the unexplained step-change in residuals in 1979, annual Tmax was homogeneous until the AWS became the primary instrument on 1 November 1996. According to ACORN-SAT v.2 metadata and other notes (Appendix 2), the site moved 150 m east of the FSU/met office on 12 September 1990 to where Micromac AWS had been installed in August. Curiously, aerial photographs show an instrument was operating at the new site as early as 28 October 1976 and that the previous site was possibly operational at least until 13 January 1992. (Although it was customary following a site move to continue observations at the 'old' site for several years in order to adjust for the move, comparison data were unavailable.) Assuming metadata is accurate, within the window from 1979 to 31 October 1996 observations were made manually using a 230-litre screen, thus moving the site was the only notable change.

Centred on the move on 12 September (and avoiding when the AWS became the primary instrument) data were split into two equi-length segments of 2190 observations. The first from 12 September 1984 to 11 September 1990 (pre-move) was compared with the post-move tranche from 12 September 1990 to 10 September 1996. Total rainfall received was 2986.8 mm and 2010.8 mm respectively with most of the difference (976 mm) occurring from March to August and October to December. Percentile distributions and differences for each segment are shown in Figure 11.

As rainfall was 976 mm less, as expected average daily Tmax was warmer following the move (28.02°C *vs*. 28.61°C; P <0.05). However, it was counter-intuitive that low-range temperatures (those less than the pre-move median of 28.4°C) were up to 1.5°C *warmer* while upper-range values were up to 0.5°C *cooler*. (Daily Tmin percentile differences (not shown) were randomly higher within the bandwidth of 0.2 to 0.5°C.)

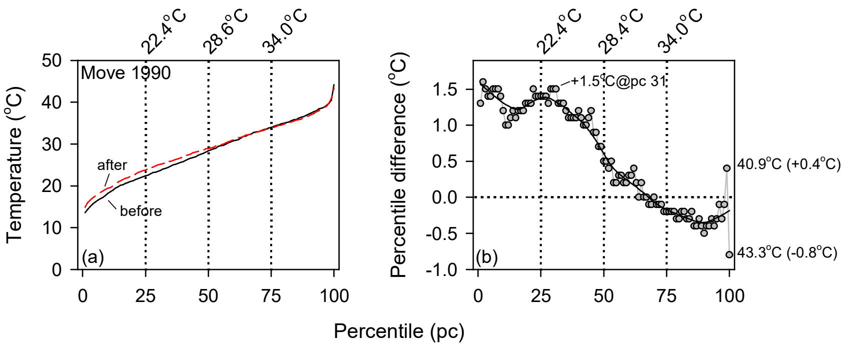


Figure 11. Percentile distributions (a) and percentile differences (new site minus previous site) (b) for equi-length data-segments centred on the move on 12 September 1990 (N = 2190 observations/tranche). Percentile differences are more conveniently interpreted than percentiles.

3.2.5.1 Comparison of 230-litre *vs*. 60-litre screens

Table 5 in: *Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) Observation practices*[1] shows a 230-litre Stevenson screen was used at Charleville from 1949 to 13 November 2003, which is when the site moved to its present location; and that a 60-litre screen used from 8 December 1999 continued to provide data (as station 44221) until 28 February 2006. The site being the same, it can therefore be surmised that data from 12 November 2000 to 28 February 2003 were for an AWS operating with a large screen and that data from 13 November 2003 to 28 February 2006 were for an AWS operating with a 60-litre screen. The first tranche (N=839 *vs*. 832 observations) received 50 mm less rain than the second (1106.4 vs. 1055.8 mm).

Although except for heavy rainfall in November 2000, monthly rainfall differences were not substantial; small-screen daily values <1st quartile (large screen Tmax of 24.5°C) were up to 1.4°C *cooler* than for the large screen with the difference increasing quasi-linearly to plateau about 1°C warmer at the 40th percentile (Figure 12). It appears that the AWS/small screen was off-scale or exceeded its calibration range at daily temperatures approaching 40°C. Although values >90th percentile (small screen Tmax of about 39.0°C) were >1°C and up to 2°C over-range at the 96th percentile they were not culled or moderated. While rain received in February was 106 mm less, it is implausible that differences could exceed 1.5°C (affecting some 40 of the 832 comparison observations).

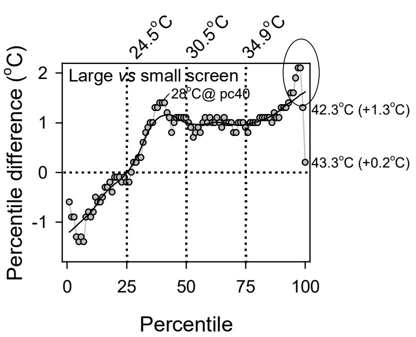


Figure 12. Relative to large screen percentiles (pc), the distribution of small-screen values increased quasi-linearly from about -1.4°C at temperatures typical of April to August, to plateau about 1°C warmer at the 40th percentile (typical values for autumn and spring). Temperature differences were most pronounced in February mainly due to lower rainfall during the small-screen segment. Extreme values of 43.3°C (6 February 2006) and 42.3°C (99th percentile – 24 December 2003, 2,3 and 5 February 2006) were notable outliers to the general distribution and were probably faulty. Values circled appear to have eclipsed the calibration range of the instrument (i.e. values greater than about 39°C went linearly off-scale).

[1] https://cawcr.gov.au/technical-reports/CTR_049.pdf

3.2.5.2 The overlap case

While ACORN-SAT metadata stated the site moved to its current location in October 2003, site-summary metadata indicated the move occurred on the 13 November, which is also the date the 230-litre Stevenson screen at the previous site was decommissioned. A small-screen inter-site comparison dataset is available for the previous site from 13 November 2003 to 28 February 2006 (Site ID 44221).

Rainfall and screen sizes were the same; however, daily Tmax at the new site was consistently 0.3 to 0.5°C warmer across the percentile range (Figure 13). The highest temperature at the new site (44.1°C) was 0.9°C warmer *on that day* than at the previous site, the second highest (43.4°C recorded on 3 days) were warmer by 0.3°C for those days and 0.6°C warmer than the 99th previous site percentile.

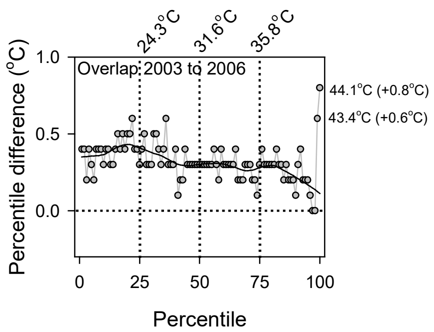


Figure 13. Percentile differences for daily Tmax measured by the AWS at the former site, and temperatures observed at the current site from 13 November 2003 to 28 February 2006 (N = 832 and 839 respectively). Extreme temperatures (44.1°C on 24 December 2003 and 43.4°C on 3, 5 and 6 February 2006 were considerably out of range relative to the general distribution and are likely to be artefacts.

3.2.5.3 Shift in the frequency of upper-range extremes in 2013

Significant rainfall-inclusive step-changes in daily Tmax frequencies greater than the median and the 95th percentile in 2013 (Figures 7 and 10) seemed associated with site disturbances in 2012. However, removing the effect of rainfall (Section 4.2.3) negated the step-change (Figure 8), thus it was *most likely* attributable to the landscape entering the drying phase from June 2012 to April 2016 highlighted in Figure 5. Percentile differences (1 January 2005 to 31 October 2012 *vs*. 1 January 2013 to 31 October 2020; N = 2859 and 2836; rainfall 4364 mm and 2770 mm respectively) show that while daily values typical of winter (<23°C; May to August) stepped-up generally by 1°C, a quasi-linear response ensued which peaked 2.1°C higher at the 86th pre-2013 percentile (36.7°C) before declining unexpectedly (Figure 14).

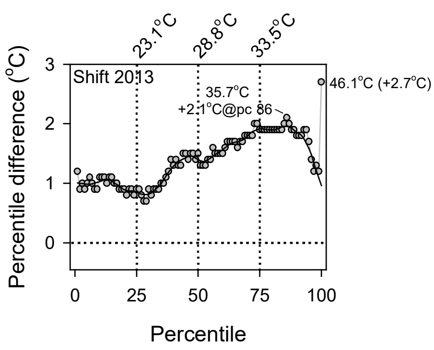


Figure 14. Percentile differences centred on the shift in 2013 show an unexpected turndown above 35.7°C. Percentile differences were calculated relative to those pre-2013.

The 2013 step-change was predominantly caused by low rainfall in 2013 (3rd lowest in 77 years), 2015 (15th lowest), 2017 (lowest on record) and 2019 (4th lowest). However, the quasi-linear inter-quartile range response in Figure 14 and the plateau were inexplicable. Further, the >86th percentile turndown suggests the instrument consistently over-ranged requiring post-2013 values >37.8°C (100°F) to be moderated either *in situ* or during processing. Further, as the highest temperature (46.1°C on 3 January 2014) was aligned approximately with the linear part of the curve it likely reflected an instrument problem. While it is unarguable that summer rainfall was much below average, the high temperature performance of the AWS seemed severely compromised.

3.3 Interim discussion

As the response of daily temperatures to changed conditions are additive, site-changes in 1979, 1997 and 2013 that caused Tmax to increase (Figure 3) caused the frequency of warm days to increase also. Further, temperature extremes (Tmax >95th day-of-year percentiles) and their frequency may also be biased by instruments reporting inconsistently across the range of daily Tmax values, which over the course of the record spanned from 8.1^{o}C to 46.4^{o}C. Although the operation of rapid-sampling platinum resistance probes and the performance of 60-litre Stevenson screens under adverse conditions were of particular interest, the possibility that upper-range temperatures have been misreported by the Bureau cannot be ruled out.

Three inter-woven problems are apparent in the data. Firstly, real-time temperatures are reported inclusive of contributing factors including the weather and the impact of site-related changes. Secondly, daily observations may be affected by instrument biases, including inbuilt or post-acquisition selective censoring of individual values. Finally, most changes likely to affect observations have occurred since 1990.

It is the decades since 1990 that have been relied upon by the Bureau, CSIRO, universities and groups such as the Climate Council and uncritical media such as *The* Conversation, the ABC and *The Guardian* as evidence of accelerated warming and increased temperature extremes. The use of homogenised data to calculate change and trend, and raw data to calculate extremes and records biases understanding of those often-cited parameters.

Taking account of rainfall and step-changes simultaneously (Figure 3) showed the apparent raw-data trend of 0.26^{o}C/decade was spuriously due to site changes in 1979, 1997 and 2013, which caused Tmax to increase 1.75^{o}C over the course of the record. So despite the absence of metadata to cross-reference the 1979 up-step, the question is settled; no trend or change is attributable to the climate.

Against the background of a non-changing climate, the situation with instrument biases and site changes on the frequency of daily Tmax measured in real-time is more complex. For benchmarking purposes, medians and percentiles more objectively define dataset properties than arbitrarily defined classes and thresholds (Section 3.2.1). However, while annual deviations greater than the median stepped-up in 1979 and 2003 (Figure 7), log-transformed LoN/HiN ratios changed in 2003 and 2013. The shift in median deviations was general in nature while those of the LoN/HiN ratio specifically compared tails of data distributions and in both cases data were inclusive of site-factors and rainfall acting together.

Accounting for the significant effect of rainfall on frequencies greater than the median (Section 4.2.3; Figure 8) confirmed that the apparent increase since 1943 was due to step-changes in 1979 and 2002. As MLR residuals were random, normally distributed and independent, no trend or change was attributable to the climate. So the analysis identified 1979 (which was the start of the 1980s drought) and 2002 (which was associated with the Millennium drought (Figure 5)) as possible factors, without indicating underlying mechanisms.

Data distributions may be skewed by observation practices, changes in instruments and rainfall deficiencies/excesses without a change in the mean or median. Thus record-hot summer days can occur while the average temperature of summer is unchanged. As a different instrument may be more sensitive to upper-range extremes or *vice versa*, data may also be skewed independently of other factors by site/instrument changes. For example, according to https://www.austehc.unimelb.edu.au/fam/0601.html, Micromac AWS report hourly data, not maxima and minima like current AWS. Thus as the absolute lowest nor highest daily values were reported, cool-day temperatures after 12 September 1990 were warmer while warm-day values were cooler. The apparent contradiction in Figure 11 is therefore explained by the change in instrument not the climate. A side issue is that according to notes (Appendix 2) after the Micromac AWS was commissioned and staff relocated to the Radio Aids building, daily manual observations may have ceased or become less regular; or they may not have been databased. The Micromac AWS was also probably replaced by an updated AWS before it became the primary instrument on 1 November 1996.

Sensitive, rapid sampling electronic thermometers housed in small Stevenson screens are likely to over-detect data-spikes that do not reflect ambient conditions. A 60-litre screen is also less buffered from being influenced by transient eddies etc. that otherwise would not be sensed by thermometers operating in a screen 380% larger. The internal layout is such that thermometers and probes housed in the same small screen cannot be exposed on the same plane relative to its rear wall that faces the sun, which can also result in instrument bias on warm days.

Centred on changepoints, percentile differences calculated for same-length segments over identical months provide useful insights into the *shape* and *nature* of changes in data distributions. Figure 12 shows the AWS/60-litre screen was predisposed to upper-range bias and that during the course of the small minus large screen comparison, and taking account of low summer rainfall, temperatures in excess of the 90th percentile (37.7^{o}C) were most likely over-reported by the small screen. The comparison of overlap data for the move to the current site in 2003 (Figure 13) found temperatures greater than the 80th percentile (36.9^{o}C) were reduced from 0.3^{o}C over-range to zero as Tmax increased. As the climate is unchanged, the post-2013 >86th percentile downturn (Figure 14) avoided the problem of unrealistically large numbers of observations approaching the old ^{o}F-century benchmark of 37.8^{o}C.

Nevertheless, care is needed in the interpretation of serial percentile differences. For instance in relation to the shift in 1979 (Figures 3, 6, 7 and 8), although pre-shift (1973 to 1978) summer rainfall was higher than from January 1979 to December 1984, significant rain was received in October 1980 (124.2), May 1981 (103.2), March 1982 (116.4), May and November 1983 (119.6 and 164 mm), and January 1984 (164.0). Percentile differences (Figure 15) showed Tmax typical of cooler months (18 to 23^{o}C) were little changed after 1979, while values typical of warmer months: >27.8^{o}C (October to Apri) were uniformly 1.5^{o}C warmer.

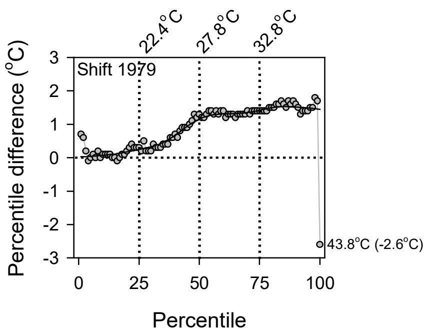


Figure 15. Percentile differenced centred on the shift in 1979 (1 January 1973 to 31 December 1978 *vs*. 1 January 1979 to 30 December 1984; post-shift minus pre-shift).

As observations were done manually using a large screen, percentile differences were consistent with summers being successively dry. Despite that, the highest post-1979 temperature (43.8^{o}C on 19 January 1980) was not as warm as that recorded on 4 January 1973 (46.4^{o}C), hence the negative 100th percentile (the maximum) excursion.

Monitored objectively as soil water deficits by the water-balance CuSum curve (Figure 5) the 1980s drought that commenced in about March 1975 had by March 1979 affected most of eastern Australia. Average Tmax was warmer (Figure 15); however, it's a conundrum that when drought broke in May and November 1983 Tmax did not step-down. While the 1979 up-step was investigated exhaustively using on-line information and objective statistical tools, its cause remained obscure.

Only so much can be gleaned from data alone. A coarse-grained oblique photograph taken by a Royal Flying Doctor Service pilot accessed at the Airways Museum showed the FSU-site with its large Stevenson screen in 1987. The site moved 150 m east to the Micromac AWS on 12 September 1990. A Queensland Government aerial photograph showed both sites were still operational on 13 January 1992 (Figure 16) and according to metadata the 60-litre screen was installed at the new site on 8 December 1999 four years prior to it relocating 600 m west to its present position in October 2003. However the 7 May 2005 site diagram for station 44221 (Charleville Aero Comparison) only shows the large screen, which according to metadata was decommissioned on 13 November 2003!

Adjustments and screen changes potentially affect trend and while it is highly likely data were adjusted for the move in 1990, overlap data is unavailable. Metadata is also silent on the location of the small screen, how adjustments were made, replacement of the Micromac AWS and if raw data were smoothed as far back as the dry years before 1983.

Referring again to Figure 8, it is apparent that moving the site in 2002 shifted the baseline from which upper-range temperatures were measured. Thus the shift provided a leg-up for daily extremes during next dry period that set-in at about the same time as the new met-office was constructed in 2012 (Figure 5). Conflating site moves and changes and the effect of drought on daily temperature 'records' is a fallacy of some consequence for understanding our climate, and as long as sites get warmer or instruments become more sensitive and biased, claims in respect of extremes become more suspect.

Figure 16. A portion of Queensland Government Aerial Photograph QAP5074053 (alt. 2250 m) showing location of the WWII operations precinct (OP), the original site (S1), subsequent sites (S2 to S4), the FSU/met office and SatCom (sc). A triangular array of aerial pylons is visible near S2. Despite having 'moved' to S3 on 12 September 1990, close inspection shows Stevenson screens were still maintained at the previous site at least as late as when the photograph was taken on 13 January 1992. (Features lie to the left of their associated labels.)

Questions remain. Average Tmax stepped-up 0.68^{o}C in 2013 independently of rainfall (Figure 3). However, June 2013 also marked commencement of the drying phase that lasted until spring 2016 (highlighted in Figure 5). While the up-step in *average* Tmax was associated with site disturbances, the increased *frequency* of extremes measured by the AWS/60-litre screen was probably related to over-detection of upper-range values due to over-ranging by the instrument during successive low rainfall years (Figure 6 and Figure 7 *vs*. Figure 8). Analysis shows that while rainfall-adjusted Tmax was unchanged, the AWS and 60-litre screen distorted the distribution of values >95th day-of-year percentiles. Figure 14 indicates the AWS exceeded its calibrated range to peak 2.1^{o}C above pre-2013 percentiles and to avoid unrealistic numbers of observations exceeding 37.8^{o}C (100^{o}F), values >86th percentile (35.7^{o}C) were most likely moderated *in situ* or by post-acquisition processing.

4. Homogenisation of Charleville Tmax for a change in 1996

Commencing with a case study of Tmax data for Mungindi Post Office, the following section briefly examines the ten, mostly manually observed datasets used to homogenise (adjust) for a statistically detected step-change in Charleville Tmax in 1996. The step-change is the same as that shown in Figure 3, which aligned with the AWS becoming the primary instrument on 1 November 1996.

The main issue is whether comparator datasets selected on the basis of being highly correlated with Charleville Tmax are themselves homogeneous, or conversely whether their data embed parallel faults.

4.1 Case study: Mungindi Post Office NSW

Located south of the Barwon River marking the border with Queensland, average monthly temperature for Mungindi Post Office (BoM ID 52020; Latitude -28.9786^{o}, Longitude 148.9899^{o}; Figure 17) is available (with several breaks) from January 1915 while daily Tmax is available from January 1965. Although about the same distance from Charleville (390 km) as London is from Paris (344 km) and further south by 2.46^{o} Latitude, data for Mungindi was used to homogenise Charleville Tmax for a statistically detected change from 1 January 1996. The other change in 2004 was adjusted relative to data for the previous site.

Figure 17. Google Map Streetview of the Mungindi Post Office (left) and, right, the Stevenson screen in the post office yard (Google Earth Pro, 6 October 2015).

Data quality was affected by missing observations; mainly weekends before 1993 otherwise isolated days or contiguous runs (Figure 18). Before metrication in September 1972, data were mainly observed in whole and ½^{o}F.

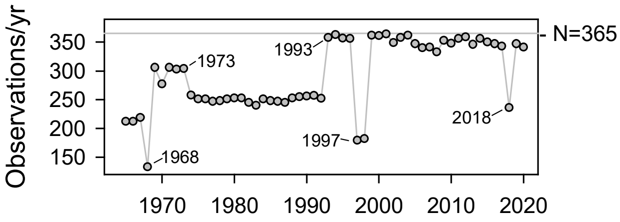


Figure 18. Most years before 1993 were missing weekend observations and isolated days; segments were missing from July to November 1968; July 1997 to May 1998 and July to September 2018.

Rainfall naïvely explained only 36.9% of Tmax variation, partly because mean Tmax was affected by missing observations and also a highly significant step-change in Tmax ~ rainfall residuals in 2004 (Figure 19(a) and (b)).

Ignoring obvious outliers, segmented regressions (Figure 19(c) and (d)) were also highly significant; before 2004 variation was higher (R^2_{adj} was less), which suggested environs surrounding the screen became more consistent or protocols changed (the screen may have moved away from shade or intermittently watered lawns for instance). Multiple linear regression confirmed that segmented regressions were parallel and offset (interaction was not significant and rainfall adjusted means (light-font in (c) and (d)) were different). Rainfall reduced Tmax 0.48°C/100 mm; rainfall and the 1.2°C up-step explained 79.8% of Tmax variation. Regression assumptions were verified and there were no residual changes or trends attributable to the climate. Several years in addition to those indicated (1994, 2002, 2010 and 2017) were identified as possible outliers but were not excluded. (Vertical dotted lines in (b) indicate timing of Charleville Tmax homogenisation adjustments.)

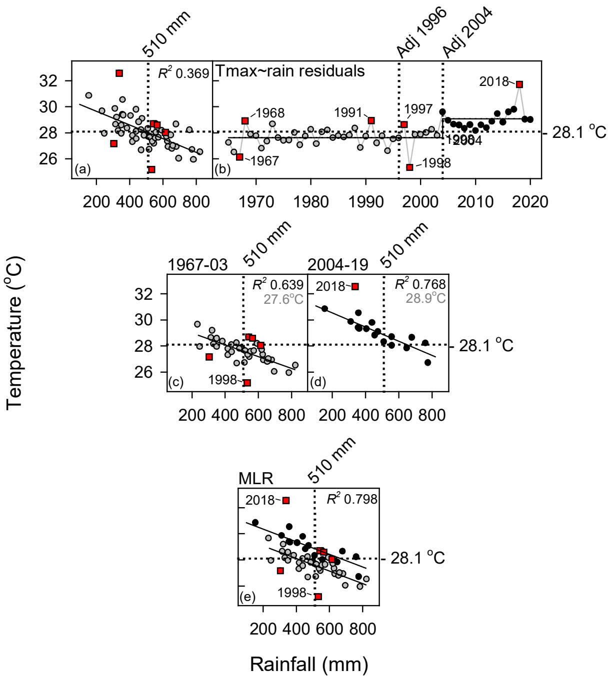


Figure 19. Tmax at Mungindi was highly dependent on rainfall (*P* <0.001) (a) but due to outliers resulting from missing observations (< 300 observations/yr, red squares) and the step-change in 2004 in (b), only 36.9% of Tmax variation was explained.

While metadata, photographs and satellite images could not explain the 2004 step-change; it most likely aligned with replacement of a previous large Stevenson screen with a 60-litre one prior to summer 2003/04. Counts and $\log_{10}$ ratios of daily values <5th and >95th day of year percentiles suggest the screen was probably replaced in 2002 (Figure 20).

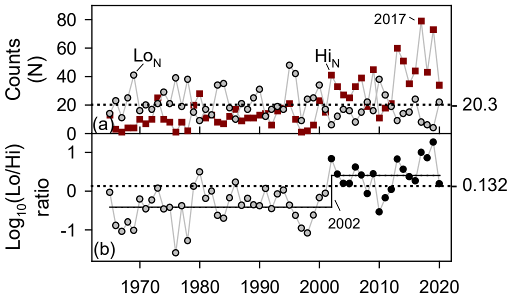


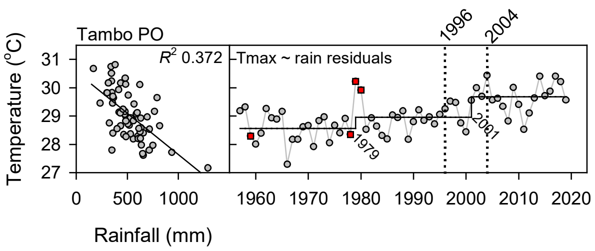
Figure 20. Despite rainfall-related excursions, tails of the distribution of daily Tmax (numbers of observations <5th and >95th day-of-year percentiles) were expected to be random overall. Although confounded with rainfall (2010 recorded the 2nd lowest rainfall since 1965 and 2019 the lowest), the Lo/Hi ratio stepped-up permanently after 2002 (b). Predominance of cooler (Lo_N) values prior to 1980 is suggestive of watering.

4.2 Additional sites used to homogenise Charleville Tmax in 1996

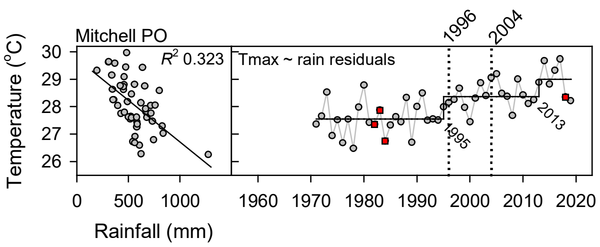
Analysis and commentary relating to the other nine datasets used to homogenise Charleville Tmax for the 1996 step-change are presented in Table 1. In all cases site-summary metadata omitted critical details such as dates when 230-litre Stevenson screens were replaced, if surrounding areas were watered, new buildings etc., which may have impacted observations. Statistical analysis supported where possible by metadata, other documentation, aerial photographs and satellite images is a more reliable and objective analytical approach than reliance on metadata alone.

Like at Mungindi, no sites were homogeneous. Satellite images show the post office yard at Tambo was watered. Site-summary metadata suggests the screen was replaced on 13 July 1998 and it moved away from a tree to the northeast corner of the yard in May 2002. In any event correlation with Charleville in the decades around 1996 was due predominantly to parallel discontinuities. Similarly at Mitchell PO where a step-change in 1996, which was probably due to the installation of a mobile phone tower or screen-size change, perfectly matched that at Charleville, which Mitchell data were used to adjust! Discontinuities related to site changes evident in data for Bollon PO (1994), Blackall (1995), Injune (2002), Longreach (1996), Cunnamulla (1988 and 2002) and Collarenebri (2004) reinforced rather than unbiasedly adjusted Charleville Tmax for the change in 1996.

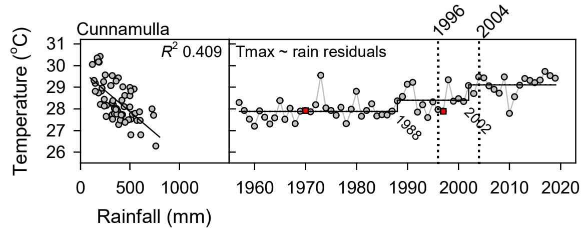
Table 1. Summary analysis of the nine additional datasets used to homogenise Charleville Tmax for a step-change in 1996, which aligned with the AWS becoming the primary instrument. Note that the change in 2004 due to the move and small screen was adjusted using overlap data. Abbreviations include median rainfall (MR), dataset average Tmax (AvTmax), small (60-litre screen) SmSc; aerial photograph, AP and change unconfirmed, ?.



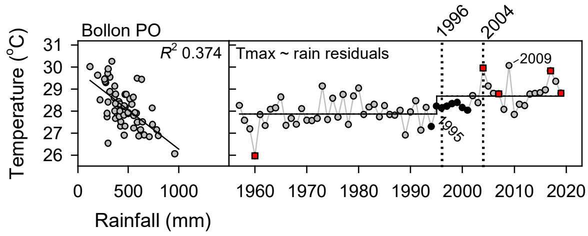
169 km distant from Charelville; MR 486.9 mm; AvTmax 29.0°C; <300 obs/yr (red squares). Rainfall reduced Tmax 0.32°C/100 mm. Rainfall adjusted steps in 1979 (Telstra tower/m?) and 2001 (Sm Sc?) increased Tmax 1.13°C. Step-changes and rainfall together explained 64.4% of Tmax variation and no residual trend or change was attributable to the climate.



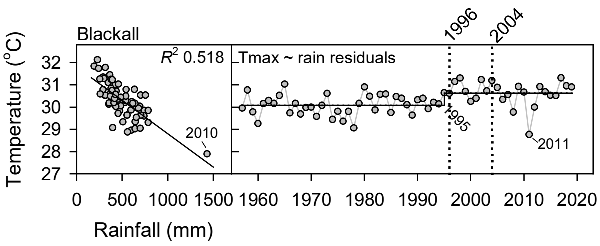
171 km distant; MR 541.4 mm; AvTmax 28.6°C; <300 obs/yr (red squares). Rainfall reduced Tmax 0.30°C/100 mm. Rainfall adjusted steps in 1995 (Telstra?) and 2013 (Sm Sc?) increased Tmax 1.45°C. Step-changes and rainfall together explained 64.4% of Tmax variation and no residual trend or change was attributable to the climate.



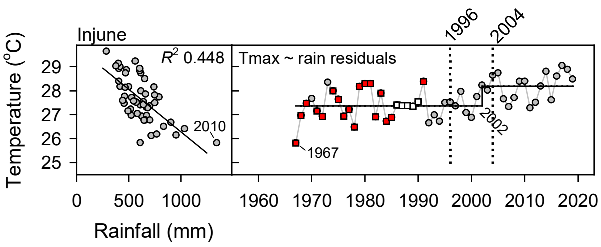
192 km distant; MR 368.4 mm; AvTmax 28.4°C; <300 obs/yr (red squares). Rainfall reduced Tmax 0.39°C/100 mm. Rainfall adjusted steps in 1998 (Telstra?) and 2002 (Sm Sc?) increased Tmax 1.23°C. Step-changes and rainfall together explained 73.1% of Tmax variation and no residual trend or change was attributable to the climate.



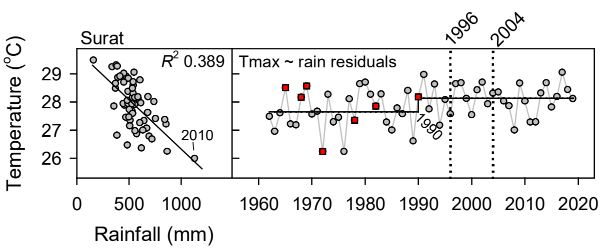
216 km distant; MR 427.6 mm; AvTmax 28.2°C; <300 obs/yr (red squares). Rainfall reduced Tmax 0.36°C/100 mm. Moving the screen 5m from shade in 1994 caused Tmax to step-up 0.81°C. The screen also moved in 1999 but there was no significant impact on data. From 1995 to 2001 (highlighted) data appeared to be predicted/infilled using rainfall as the covariate. Step-changes and rainfall together explained only 55.7% of Tmax variation and no residual trend or change was attributable to the climate.



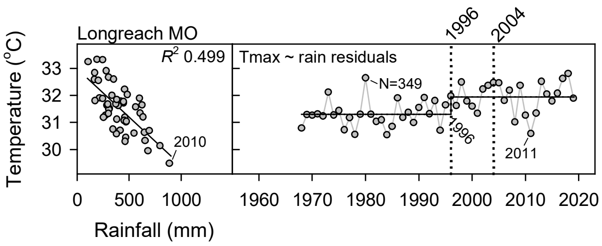
234 km distant; MR 459.0 mm; AvTmax 30.3°C; no <300 obs/yr. Rainfall reduced Tmax 0.31°C/100 mm. Rainfall adjusted step in 1995 (Sm Sc?) increased Tmax 0.56°C. Step-change and rainfall together explained 62.7% of Tmax variation and no residual trend or change was attributable to the climate. Unless the site moved, coordinates of the post office are miss-specified (-24.4240° Lat, 145.4653° Long *vs*. –24.4214°, 145.4672° in current metadata). PO and AP data were joined on 1 April 2001.



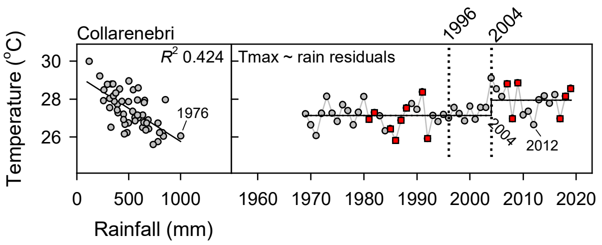
238 km distant; MR 610.3 mm; AvTmax 27.7°C; <300 obs/yr (red squares); data from 1986 to 1990 (white squares) appear to have been infilled/predicted. Rainfall reduced Tmax 0.34°C/100 mm. The rainfall adjusted up-step in 2002 (Sm Sc?) caused Tmax to increase 0.80°C. The step-change and rainfall explained 60.7% of Tmax variation and no residual trend or change was attributable to the climate. (Missing observations, mainly weekends, affected precision).



291 km distant; MR 532.5 mm; AvTmax 27.9°C; <300 obs/yr (red squares). Rainfall reduced Tmax 0.34°C/100 mm. Rainfall adjusted step in 1990 (Telecom) increased Tmax 0.50°C. Step-change and rainfall together explained 46.7% of Tmax variation and no residual trend or change was attributable to the climate. APs show the PO yard was irregularly watered and that Telstra buildings were installed between 1985 and 1992. No change due to the move to the AP SmSc on 4 February 2003.



ACORN-SAT site; 383 km distant; MR 381.0 mm; AvTmax 31.7°C; no N<300 obs/yr. Rainfall reduced Tmax 0.36°C/100 mm. Rainfall adjusted step in 1996 (AWS/SmSc) increased Tmax 0.65°C. Step-change and rainfall together explained 71.8% of Tmax variation and no residual trend or change was attributable to the climate. (The 1996 site change was confirmed by ACORN-SAT metadata).



414 km distant; MR 522.6 mm; AvTmax 27.4°C; <300 obs/yr (mainly weekends, red squares). Rainfall reduced Tmax 0.391°C/100 mm. Rainfall adjusted change in 2004 (verified relocation, SmSc in 2003) increased Tmax 0.85°C. Step-change and rainfall together explained 58.0% of Tmax variation and no residual trend or change was attributable to the climate.

5. Discussion

5.1 Metadata

Lack of reliable metadata for Charleville and the 10-comparator sites used to homogenise Charleville Tmax for the change in 1986 is the main barrier to understanding and analysing respective datasets. As ACORN-SAT site-summary metadata is untrustworthy, independent corroboration is essential. Although only data for Charleville airport was used in this paper, the ACORN-SAT dataset commenced in 1910 using data for the Post Office (actually the school) where according to Simon Torok[1] in 1910 "goats fed under the screen at times"; in 1914 the site moved from the school to a "poorer" site at the post office where 27-years later in 1941, the screen was found to be "too high and faces southwest" (instead of south); in 1943 the site moved "to aerodrome for composite" i.e. data were joined. To account for the join, Torok applied a Tmax homogenisation adjustment of 0.4°C, ACORN-SAT v.1, 1.8°C and ACORN-SAT v.2, 1.64°C. As adjustments affect trend and they are applied arbitrarily the adjustment process lacks rigor and objectivity.

As there is no site plan or other information, factors that may have impacted on the quality of Post Office data (Bureau ID 44022) are unknown. Plans in the National Archives show alterations and additions to premises occurred in 1941 (Barcode 1712629); also, the post office weather station file is open for inspection in NAA's Brisbane office (Barcode 20481565) but is not available online. Analysis found Post Office Tmax stepped-down 0.53°C in 1943 (rainfall adjusted), which negated the response to the worsening inter-War drought. Pending further information about the state of the site and the instruments and screen; from a long-term trend perspective it is unlikely that Post Office data could be considered reliable.

Documentation relating to Charleville aerodrome as relayed by ACORN-SAT is misleading. Observations at the original site south of the former operations precinct (Appendix 1) were made by Aeradio/met staff located in the control building behind the control tower adjacent to the E-W runway. ACORN-SAT metadata stated, "Neighbours rather than the overlapping data are used to merge the two main Charleville sites, as the data indicate substantial changes at both sites in or near 1949". No indication of the nature of those changes was provided by ACORN-SAT. However,

[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.

the nearest neighbouring site was Tambo PO 168 km away where daily data only started in 1957, or 300 km distant at the St George PO and it is dubious that either site's data could usefully predict daily observations for the Charleville aerodrome.

The main climate change event was breaking of the long-running inter-War drought in March and October 1949 and from January to June 1950 (Figure 5), which stimulated grass growth and in-turn fuelled widespread grass and bushfires in 1951.

Cumulatively low rainfall between March 1943 and February 1947 caused widespread devastation. Numbers of sheep in Queensland declined by 10-million … the railways were hauling water as towns ran dry … water was available for only 2-hours per day in Townsville[1]... It could not be the case that the devastating drought highlighted in Figure 5 had no impact on maximum temperature at the Charleville Post Office or airport. Fifteen randomly chosen daily values reported by newspapers between 1944 and 1951 were not the same as those in the current post office dataset either. As Tmax depends on rainfall it was highly unlikely that the step-down in post office data of 0.53^{o}C from 1943 to the end of the record in 1951 reflected the true climate, while infilling of airport data between 1943 and 1956 raises suspicions and questions about the integrity of the Bureau's data handling methods.

5.2 Statistical methods

Analysis protocols used in this and previous BomWatch studies were physically sound and robust. Twelve datasets were analysed, 10 of which were used to homogenise Charleville Tmax for the change in 1996, which aligned with the AWS becoming the primary instrument. Interestingly, while the rainfall-adjusted step-change in 1996 was 0.58^{o}C, ACORN-SAT v.1 made no adjustment, while ACORN-SAT v.2 adjusted by only 0.2^{o}C!

All maximum temperature datasets were confounded with parallel site-related effects. While naïve linear regression accounted for the effect of rainfall on Tmax, step-change analysis of residuals identified non-rainfall discontinuities, which where possible, were attributed using metadata etc. If rainfall fully-explained Tmax, residuals would be random, independent and homogeneous (i.e. be unrelated to their order in the dataset). Sidestepping problems inherent in time-series analysis, multiple linear regression calculated rainfall-adjusted differences without bias and tested that segmented rainfall responses were the same, while additional analysis identified outliers and verified appropriateness of the step-change model.

5.3 Temperature extremes

While Australian temperature trends and trajectories are determined using homogenised data, extremes reported in the Bureau's monthly and annual summaries, and highlighted in daily media are inclusive site and instrument change effects. In the light of previous discussion it can hardly be argued that temperatures reported by AWS operating with small Stevenson screens at the current site would be comparable to those measured using thermometers housed in large screens near the FSU office before September 1990, or daily values that were estimated between 1943 and 1956, including during the height of the inter-War drought. Repeated claims of extreme temperatures, trends in extremes and record-hot days, which cannot be substantiated are misleading and damage the reputation of all concerned.

Of the various methods used to portray temperature extremes, objective data-centric approaches are the least biased and the most amenable to partitioning site-change effects from

[1] J.C. Foley (1957). "Drought in Australia. "Review of records from earliest years of settlement to 1955". Bulletin 43, Bureau of Meteorology, Melbourne.

those due to the climate (Section 3.2.3). As changes in the frequency of daily observations that exceed arbitrarily defined thresholds, such as 37.8°C (100°F) or 40°C, or frequencies within incremental classes could reflect the choice of threshold, observer bias, a faulty (or replaced) instrument or deteriorating site conditions, threshold-based approaches are not recommended.

Shifts in deviations from median-Tmax provide a relative overview of possible site-change effects while values <5th and >95th day-of-year percentiles and their ratio focussed on the tails of data distributions. Accounting for rainfall (Section 3.2.3) found the change in 1979 was *most likely* related to a site change and that the move (and 60-litre screen) in 2003 caused the frequency of daily values greater than the median to step-up permanently. The marked increase in upper-range (>95th percentile) extremes during the Millennium and subsequent drought was related to 2003 up-step. Metadata is inconsistent as to whether the 60-litre screen was installed at the previous site on 18 December 1999 or whether maximum temperature continued to be measured using a 230-litre screen until October 2003 when ACORN-SAT indicated the site moved, or 13 November 2003 as indicated by site-summary metadata; and whether 'raw' data were merged/averaged in 1990 or during the overlap from 13 November 2003 to 28 February 2006. As the large screen was decommissioned on 13 November 2003, only 60-litre screen data was available during the overlap.

Although AWS may be more precise they are also more likely to report spikes and over-range values on warm days especially when unattended under adverse conditions such as at airports and beside dusty tracks and during droughts.

5.4 Percentile differences

Percentile distributions may change due to either a site/instrument change or a change in a related variable. Nevertheless, provided results are interpreted cautiously they provide insights into the effect of a change on the *shape* or *nature* of the Tmax response. Differences may uniformly step-up (down); they may change randomly across the percentile range; they may be due to unexplained excursions; or, they may be systematically related to the temperature being measured. At their extremes, distributions may run off-scale either negative or positive, or show other features suggestive of culling or downscaling of out-of-range values.

Following the various interventions compared in Figures 11 to 14, daily maximum temperature was generally off-scale relative to previous data. Figures also show that individual extremes were also off-scale *relative to the trajectory of percentile differences*. It is unlikely for example that maximum temperature measured on 6 February 2006 (43.3°C; Figure 11) would be 0.8°C under-range while post-1990, 99th percentile values were over-range to about the same extent; similarly in Figure 12 where Tmax on 16 January 2001 (42.3°C) was 1.1°C over-range; in Figure 13 (44.1°C on 24 December 2003), 0.8°C over-range and Figure 14 (46.1°C on 3 January 2014), about 1.7°C over-range. In 1973, Tmax on 4 January (43.8°C) was some 3°C under-range (Figure 15). As individual data values near the extremes of data distributions are more likely to be faulty, their use as climate metrics is highly questionable.

5.5 Homogenisation

Although methods have evolved since 1996 when the first homogenised datasets were used to calculate Australia's warming[1], from the pool of neighbouring datasets only those that are highly

---

[1] Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset. CAWCR Technical Report No. 049 (https://cawcr.gov.au/technical-reports/CTR_049.pdf).

correlated with target site data are used to make adjustments. Data for Mungindi and the nine other mostly manual datasets used to homogenise Charleville for the Tmax step-change in 1996 (Figure 19 and Table 1) owed their correlation to embedded parallel faults – alignment of screen and site changes, site moves, watering, the small screen and AWS at Longreach … Like for Charleville, unbiased analysis found no trend or change at any sites that was attributable to the climate.

Using faulty (correlated) data to correct faults in ACORN-SAT data has no scientific or statistical merit and should be abandoned.

Furthermore, analysis discussed above shows that matching percentile distributions across years whose rainfall was not the same is fraught, especially considering that underlying step-changes in N>$Tmax_{median}$ in 1979 and 2002 (Figure 8), the move to the Micromac AWS (Figure 11), the change in percentile distributions for the small screen indicated in Figure 12 and likely over-ranging by the AWS after 2013 (Figure 14) were not adjusted. And there is also the issue that reconstruction of airport data from 1943 to 1965 effectively erased the cumulative effect of extremely dry conditions on Tmax during the closing years of the inter-War drought.

5.6 Implications

In various guises, for example as raw data, homogenised data, aggregated data such as gridded point or geographically dispersed 2-D and 3-D surfaces, the Bureau's temperature datasets are used widely to inform, discuss, shape public opinion and develop/support policy, as well as for a wide range of scientific studies related to the climate and the biosphere more generally. Many, if not hundreds of peer-reviewed scientific papers published in top-shelf journals are predicated on the notions that (i), data truly reflect the climate; (ii), data are fit for purpose – that conditions under which measurements were made were not biased and that data are carefully curated; and (iii), methods used by the Bureau for quality control including for homogenisation are transparent and replicable across multiple sites, datasets and regions.

However, none of the underlying assumptions relating to the use of Charleville and other sites' data for specific purposes were found to be true. For instance, while all sites thus far examined have changed and moved in ways that affect data, no datasets show trend or change that could be unambiguously attributed to the climate. For all sites, metadata is also inexact and potentially misleading. That pre-1956 data were made-up at Charleville and possibly in-filled/estimated for several years at Bollon and Injune undermines claims that the Bureau's data handling methods are amongst the best in the world.

- That homogenisation uses highly correlated neighbours that embed parallel faults to adjust ACORN-SAT data used by CSIRO in State of the Climate Reports; which, in-turn, is used to inform policy and is reported-on without question by activist organisations such as the Climate Council, WWF and media including *The Conversation* the ABC and *The Guardian* is a scientific travesty of monumental proportions.
- That users of the Bureau's data including CSIRO and professors and students at Australian universities have failed to check soundness of the data they use and publish in peer-reviewed papers seriously undermines their credibility and that of the institutions they represent.
- Finally, failure to undertake due diligence by those purporting to be 'battling' climate change like the Great Barrier Reef Foundation, Farmers for Climate Action, the Murray-Darling Basin Authority, banks like ANZ, lobbyists for the electricity industry, committees

> advising ministers, the Australian Research Council, CSIRO and academies of sciences, fuels widespread distain for institutions the community should be able to trust.

The critical point is there is no change in the climate at Charleville or other sites that have been investigated in this and previous reports. For the people of central Queensland save for a sudden shift in the subtropical ridge such as happened in November 1924 (Figure 5), the climate they have is the climate they will experience for the foreseeable future. Droughts will recur but are not increasing in severity or frequency; their effects are just more visible from space, on the evening news and on social media. It is also no warmer that it ever was and conflating changes in equipment, in particular the use of rapid-sampling AWS-probes housed in infrequently serviced 60-litre Stevenson screens under adverse conditions, with "the climate" is misleading in the extreme.

6. Conclusions

Maximum temperature data for Charleville aerodrome, which commenced in 1943 showed no trend or change attributable to the climate. Location of the original WWII Aeradio office and meteorological enclosure was not identified by site-summary or ACORN-SAT metadata. Data to 1955 were evidentially reconstructed using rainfall as the covariate and therefore the quality of the 'real' data is impossible to gauge. Reconstruction effectively erased the effect of the devastating, long-running inter-War drought. The site (and staff) probably moved from the Aeradio office on the northern side of the airport to the Flight Services Unit office near the terminal in 1956 (not 1961).

Tmax depends on rainfall and a monthly water balance showed the climate transitions rapidly from sequences of years that are dry and hot to sequences that are relatively mild every 5 to 7 years. Drought from November 1924 to January 1949 was the longest, most severe, most unrelenting and most devastating drought experienced since 1900. Local runoff is rare and mostly associated with tropical cyclones and monsoonal troughs that affect wide areas at about the same time. Although the climate could transition as had previously during the early years of the 20th Century, there is no indication that the frequency of drought has or is likely to increase.

The frequency of daily Tmax relative to the dataset median; and counts of values N <5th *vs.* N >95th day-of-year percentiles and their log(LoN/HiN) ratio, provided data-centric indices of trend and change in extremes. Post 1990 site changes, particularly the site move in 2003 caused the frequency of extremes to increase; however, there was no indication the increase was related to the climate.

Percentile differences for segments of equal length over matching months straddling a changepoint visualise the *shape* and *nature* of changes in data caused by the change. However, as rainfall is also a factor they require cautious interpretation.

Headed by a case study of Mungindi Post Office, the ten stations used to homogenise Charleville Tmax for a change in 1998, which was related to the AWS becoming the primary instrument were also analysed. Chosen on the basis of inter-site correlations, most datasets embedded faults that aligned with those in Charleville data. Using faulty data to adjust faults in ACORN-SAT data has no scientific or statistical merit and should be abandoned.

As there was no change in the climate at either Charleville or in respect of the 10-comparator sites there can be no real change in the frequency of extremes either. Calculating Australia's long-term climate trajectory using ACORN-SAT data that has been biased by the homogenisation process; while proclaiming that extremes and 'records' measured by AWS and 60-litre screens are comparable with the long-term data is misleading and damages the reputation of all concerned.

Acknowledgements.

Suggestions on earlier drafts by family members and Chris Gillham (http://www.waclimate.net/), Ken Stewart (https://kenskingdom.wordpress.com/), Geoff Sherrington and David Mason-Jones (https://www.journalist.com.au) were greatly appreciated.

Research undertaken here received no external funding.

Author's note: Although this draft paper has been carefully researched and uses statistical tools, data and other evidence that is available in the public domain, no responsibility is accepted for errors of fact arising from the scarcity of corroborating information, photographs etc.

Charleville data used in the study is given in Appendix 3.

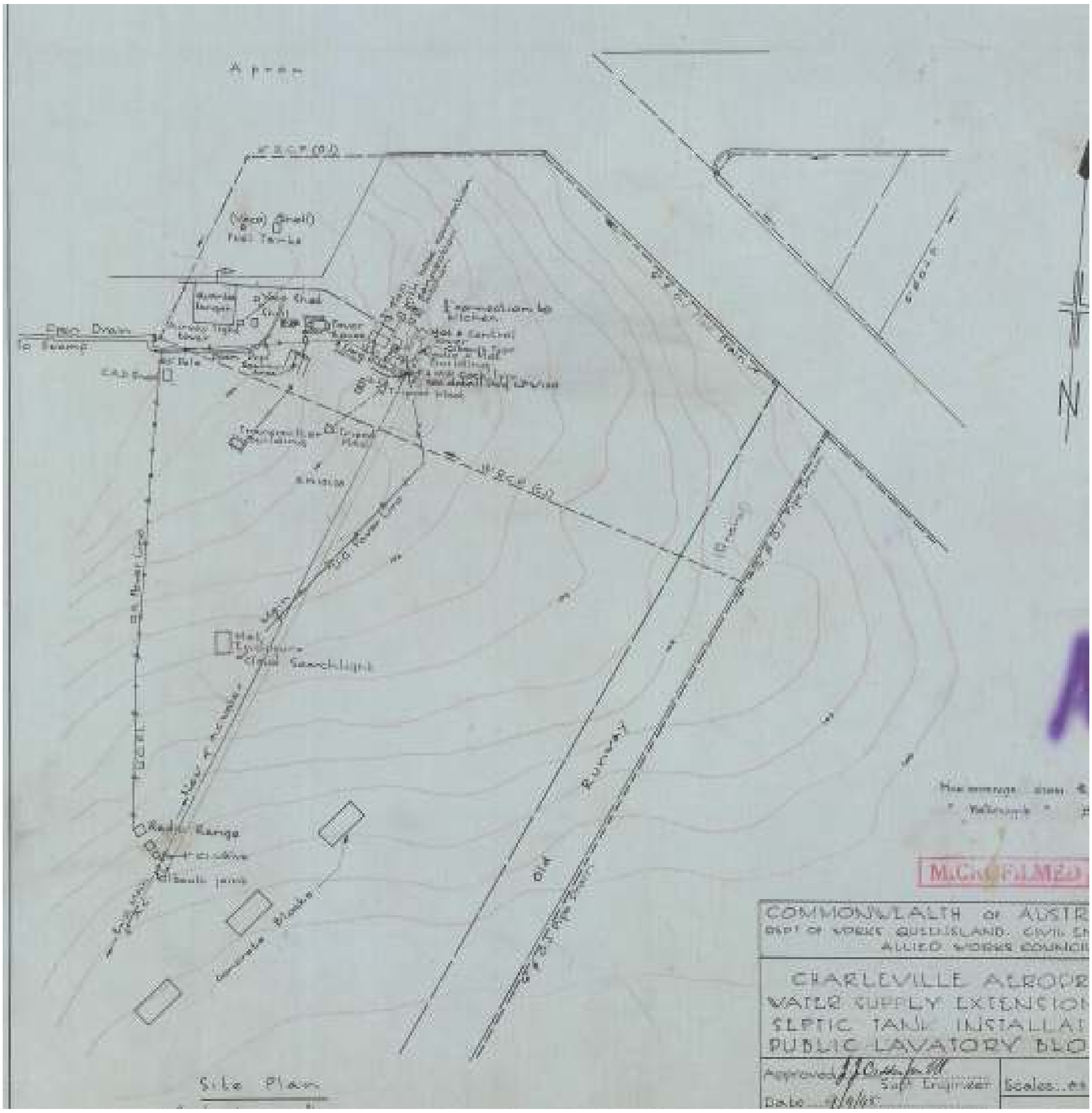
Preferred citation:

Johnston, Bill 2021. Are Australia's automatic weather stations any good? Part 4. Effect on temperature extremes. http://www.bomwatch.com.au/ 30 pp.

Dr. Bill Johnston (version 17 February 2021)

Appendix 1.

Section of Charleville aerodrome water supply and drainage diagram. (NAA Item ID, 1972125)



Appendix 2[1]

CBoM - History of Charleville Meteorological Office

Australian Government
Bureau of Meteorology

History of Charleville Meteorological Office

Home | Warnings & Forecasts | Observations | Satellite Images | Maps & Charts | Climate & History
Contact Us | Offices | Links

Surface observations at **Charleville** commenced at 9am Monday 6th April, 1942 with observation of Upper Winds using Pilot Balloon Theodolite commencing on the same day. The station operated from the Civil Aviation Aeradio Building.

V.J. Bahr, the first Officer-in-charge of the Charleville Office, wrote *"the commencement of Meteorological Services had their genesis in the need to provide vital strategic support connected with the desperate defence of Northern Australia following the rapid advances of the Japanese through South East Asia, Papua New Guinea and the Islands to the North and Northeast of Australia, and the disastrous bombing raids on Darwin on 19 February 1942"*.

The office's role has undergone changes over the years with steady improvement in the equipment used to carry out its functions. The more significant equipment changes/upgrades include: June 1943-Radio Sonde equipment installed; October 1955-Type 277F radar installed in newly constructed RadioAids building near the aerodrome entrance; September 1958 Sferics lightning-detection equipment installed - similar equipment was installed at Brisbane Airport and Townsville Met. offices and these stations formed the Queensland Sferics Network; December 1983-WF 100 Windfinding radar; August 1990-MICROMAC Automatic Weather Station (AWS); May 1991-PC-Cora (computerised radio sonde system).

The forecast and pilot-briefing function of the office moved in 1960 to the new DCA complex which was situated in close proximity to the aerodrome passenger terminal. Met. and Flight Service staff worked in adjacent offices. The RadioAids facility functioned as an upper-air monitoring station at the same time. With the commissioning of the AWS in 1990, all station functions were consolidated into the RadioAids building.

Met. office staff overcame significant personal hardships during the devastating April 1990 flood. Some 2000 residents of the area were evacuated to a large hangar on the airport, prior to further evacuation of many of these people to other towns and cities. From the onset of the flood until early June the Met. staff (total 6 officers) lived and worked at the Met. RadioAids building on the aerodrome, whilst their families had been evacuated out of Charleville.

The Charleville area was the focus for weather experimentation long before the establishment of a weather office in the area. In 1902 the Queensland Government Meteorologist (Professor Clement Wragge) placed Steiger Vortex Guns around Charleville, with a view to firing these into the atmosphere in an attempt to induce rain to ease a drought, which was prevalent at the time. Whilst the experiment failed, the interest surrounding the event is remembered by way of a monument in the city.

This page was created at on

[1] No longer accessible at http://www.bom.gov.au/qld/charleville/history.shtml; accessed 7 March 2017

Appendix 3

Charleville data summarised from daily data downloaded from the Bureau's Climate Data on-line facility, which was used in the study.

Year	Rain	MaxAv	MaxN	MaxVar	ShMaxRes2	Check 1956 ShMaxRes2a
1943	397.5	27.74	359	52.62	1	1
1944	318.0	28.48	365	51.58	1	1
1945	341.3	28.22	364	46.90	1	1
1946	207.0	28.74	364	44.30	1	1
1947	551.4	27.16	365	43.92	1	1
1948	361.4	28.03	362	47.09	1	1
1949	833.4	26.65	361	46.90	1	1
1950	1025.2	25.78	362	38.30	1	1
1951	299.4	28.11	365	46.09	1	1
1952	560.2	27.21	353	55.16	1	1
1953	481.3	27.70	365	43.46	1	1
1954	759.5	26.83	362	36.74	1	1
1955	817.9	26.54	365	44.93	1	1
1956	818.0	25.63	366	47.67	1	2
1957	377.1	28.89	365	51.05	1	2
1958	343.0	29.02	365	45.82	1	2
1959	505.5	28.02	365	40.34	1	2
1960	497.2	27.34	365	54.36	1	2
1961	538.0	27.55	365	39.83	1	2
1962	624.9	27.37	365	33.33	1	2
1963	835.7	26.67	365	39.48	1	2
1964	277.8	28.27	366	38.92	1	2
1965	266.3	29.09	365	49.24	1	2
1966	369.3	27.40	365	44.96	1	2
1967	392.1	27.92	365	48.77	1	2
1968	406.1	27.53	365	55.33	1	2
1969	353.9	27.21	323	41.96	1	2
1970	359.1	26.65	312	35.43	1	2
1971	626.0	27.48	365	43.42	1	2
1972	251.4	28.51	364	37.83	1	2
1973	822.5	28.27	363	38.87	1	2
1974	488.3	27.41	365	38.21	1	2
1975	504.9	28.09	365	38.74	1	2
1976	546.2	26.64	366	36.65	1	2
1977	484.2	27.91	365	46.26	1	2
1978	550.0	26.60	365	52.10	1	2
1979	272.6	29.10	365	49.40	2	3
1980	419.2	29.30	366	47.96	2	3
1981	427.2	28.07	365	53.12	2	3
1982	245.4	28.73	365	50.41	2	3
1983	740.4	27.53	365	53.93	2	3
1984	442.6	27.65	366	40.89	2	3
1985	418.8	28.29	364	46.54	2	3
1986	484.8	28.33	365	48.23	2	3
1987	615.0	27.82	365	44.82	2	3
1988	374.2	28.73	366	45.71	2	3
1989	515.4	27.43	365	54.74	2	3
1990	502.6	27.97	365	61.61	2	3
1991	266.4	29.38	365	34.89	2	3
1992	290.2	28.04	366	42.33	2	3
1993	382.8	28.54	364	40.29	2	3

1994	388.6	28.13	365	40.36	2	3
1995	362.2	28.73	365	45.24	2	3
1996	492.6	28.34	366	41.90	2	3
1997	848.4	27.65	365	38.85	3	4
1998	529.2	28.44	365	51.03	3	4
1999	690.2	28.01	365	34.57	3	4
2000	662.0	27.51	366	37.65	3	4
2001	402.8	28.41	365	42.48	3	4
2002	340.0	29.68	365	40.56	3	4
2003	302.0	29.14	365	46.74	3	4
2004	550.6	29.12	366	44.37	3	4
2005	365.6	30.01	365	52.44	3	4
2006	373.4	29.48	365	51.62	3	4
2007	522.8	28.69	365	44.64	3	4
2008	619.6	27.81	365	37.68	3	4
2009	365.8	29.65	364	41.67	3	4
2010	1133.8	26.12	365	34.31	3	4
2011	590.6	27.79	365	38.59	3	4
2012	503.6	28.35	366	44.37	3	4
2013	220.6	30.29	364	43.01	4	5
2014	476.9	29.66	365	44.88	4	5
2015	267.4	29.50	365	50.26	4	5
2016	651.8	28.79	365	54.74	4	5
2017	203.0	30.68	357	44.40	4	5
2018	270.6	30.35	362	47.39	4	5
2019	240.6	30.14	365	54.35	4	5