Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset

CAWCR Technical Report No. 049

Blair Trewin

March 2012

www.cawcr.gov.au

Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset

Blair Trewin

The Centre for Australian Weather and Climate Research
- a partnership between the CSIRO and the Bureau of Meteorology

CAWCR Technical Report No. 049

March 2012

ISSN: 1836-019X

National Library of Australia Cataloguing-in-Publication entry

Authors: Blair Trewin

Title: Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset.

ISBN: 978 0 643 10815 8

Series: CAWCR Technical Report; 049

Other Authors/Contributors: Day, K.A. (Editor)

Notes: Includes index and bibliography references

Enquiries should be addressed to:

Blair Trewin
Centre for Australian Weather and Climate Research:
A partnership between the Bureau of Meteorology and CSIRO
GPO Box 1289, Melbourne
Victoria 3001, Australia

B.Trewin@bom.gov.au

Copyright and Disclaimer

Contents

List of Figures

List of Tables

ABSTRACT

The Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) data set is a long-term data set of Australian daily air temperature, covering the period from 1910 to the present. The purpose of this data set is to provide the best possible data set to underlie analyses of variability and change of temperature in Australia, including both analyses of annual and seasonal mean temperatures, and of extremes of temperature and other information derived from daily temperatures. A full discussion of the motivation underlying the ACORN-SAT data set is contained in a companion report in this series.

The purpose of this report is to describe the ACORN-SAT data set, the techniques involved in its development, and the issues that arise in developing such a long-term data set in Australia. Whilst the issues involved in producing any long-term data set are complex, most of them are well understood and, with appropriate treatment, do not substantially inhibit the development of a data set suitable for characterising long-term Australian temperature trends and variability.

1. INTRODUCTION

High-quality temperature data sets are vital for climate monitoring, and especially the monitoring of climate change. Many temperature data sets exist at the global (Brohan et al, 2006; Hansen et al., 1999, 2001; Smith et al., 2008), regional (e.g. Klein Tank et al., 2002) and national (e.g. Begert et al., 2005; Folland and Salinger, 1995; Hanssen-Bauer and Nordli, 1998; Menne and Williams, 2009; Vincent et al., 2002) scales.

If a temperature data set is to be used for monitoring climate change it is important that it be homogeneous; that is, changes in the temperature as shown in the data set reflect changes in the climate, and not changes in the external (non-climatic) conditions under which the observations are made. Potential non-climatic influences on temperature observations, which are discussed in more depth in Trewin (2010), include changes in local ground conditions around an observation site, changes in instruments and changes in observation procedures. In addition, many station "series" are taken from more than one location, despite often appearing under a single geographical name, with the site moves clearly needing to be taken into account when using the associated data.

Very few century-scale temperature station series are totally free of such influences, and thus careful homogenisation is required in order to produce a homogeneous data set. Whilst site-specific inhomogeneities have only a marginal impact on observed temperature trends at the global scale because they tend to cancel each other out when averaged across large numbers of stations (Jones and Wigley, 2010), they can have a much more substantial effect on outcomes at the local and regional scale, including national data sets where those are based on a relatively small number of stations.

The homogenisation process is a two-stage process: firstly the detection of inhomogeneities in the data, and secondly the adjustment of data to remove those inhomogeneities. The detection of inhomogeneities is a well-explored problem, both in climate science and in the broader statistical literature. Reviews on this topic include those of Peterson et al. (1998) and Reeves et al. (2007). It has also been the subject of the European *Advances in homogenisation methods of climate series: an integrated approach* (HOME) project (www.homogenisation.org), which has included a comprehensive benchmarking assessment of methods for detecting inhomogeneities at the monthly or annual timescale (Venema et al., 2012).

Adjustment procedures have received much less attention, with most data sets applying adjustments on the basis of annual (e.g. Della Marta et al., 2004) or monthly (e.g. Jones et al., 1986) means. However, homogenisation of mean temperatures does not necessarily imply homogenisation of higher-order statistical properties such as variance, or derived statistics which are a function of those higher-order properties, such as the occurrence of extremes. This issue was initially identified by Trewin and Trevitt (1996), who found that in some cases an inhomogeneity, such as a site move, affected different parts of the frequency distribution of daily temperature in different ways.

Most national and international data sets which have been homogeneity-adjusted use adjustments calculated on the basis of annual or monthly means. In most cases these apply a uniform adjustment for each calendar month or across the year, although there have been some data sets (some of which use the term 'daily homogenisation') which apply a different adjustment for each calendar date, normally derived from monthly data (e.g. Brunet et al., 2006; Vincent et al., 2002). Whilst a number of techniques have been developed, as described later in this report, to adjust data at the daily timescale using adjustments which are different for different parts of the frequency distribution, or are otherwise dependent on weather types, the only previous known application of such techniques to a national-level data set is the 1957-1996 Australian daily data set (Trewin, 2001a).

This report describes the methodology used for the construction of the ACORN-SAT data set. This data set makes use of substantial recently digitised data (from paper records) to allow homogeneous maximum and minimum temperature data to be produced at the daily timescale for the period from 1910 to the present, with coverage across the Australian continent. Documentation and traceability of the data and adjustments at all stages, an increasing priority as described in Thorne et al. (2011), are also a high priority in the ACORN-SAT data set.

2. THE AUSTRALIAN TEMPERATURE OBSERVING NETWORK AND ITS DATA AVAILABILITY

Instrumental observations of temperature have been made in Australia since the days of the First Fleet in the late 18th century (Gergis et al., 2009). Various short-term observations (data sets of a few years or shorter) were made up until the middle of the 19th century, either under the auspices of colonial authorities[1] or through private initiatives.

From the 1850s onwards, temperature observations were made in a more systematic fashion. The longest continuous temperature record in Australia commenced in Melbourne in 1855, although with a number of site moves within the central Melbourne area since then, and by the early 1860s a number of sites[2] existed in New South Wales, Victoria and South Australia. The number of sites increased steadily through the second half of the 19th century, and there was reasonable coverage of the eastern mainland by 1890. Progress was slower in Tasmania and Western Australia, where there were very limited observations outside Hobart and Perth before 1900. However, few consistent standards for instruments or observations existed before the late 19th century, making 19th-century data very difficult to compare with more modern records.

---

[1] Until Federation in 1901, Australia consisted of six separate colonies of the United Kingdom of Great Britain and Ireland. On Federation, the six colonies became the current six States, with the current Northern Territory and Australian Capital Territory then being parts of South Australia and New South Wales respectively.

[2] In this report, 'site' refers to a specific place where observations are made, 'location' refers to a general area (e.g. a town). A record for a 'location' may incorporate data from a number of 'sites' over time.

The creation of the Bureau of Meteorology in January 1908 brought all meteorological observations under federal control. This resulted, within a short time, in the implementation of common standards for observations and instrumentation, discussed in more detail below. It also resulted in a rapid increase in the number of sites with available observations. The number of sites then stabilised during the 1910-1940 period, before increasing further from 1940 through to the early 1950s, initially as a result of the Second World War, then with the growth of civil aviation. There have not been dramatic changes in the number of sites since then, although the spread of sites over Australia has become more comprehensive, with observations beginning in the 1950s and 1960s in a number of locations in remote parts of central and northern Australia that previously lacked any coverage. Coverage of high-altitude areas in southeastern Australia has also improved greatly in the last 20 years, mostly through increased use of automatic weather stations.

Figure 1 shows the temperature networks in 1930 and 2010.

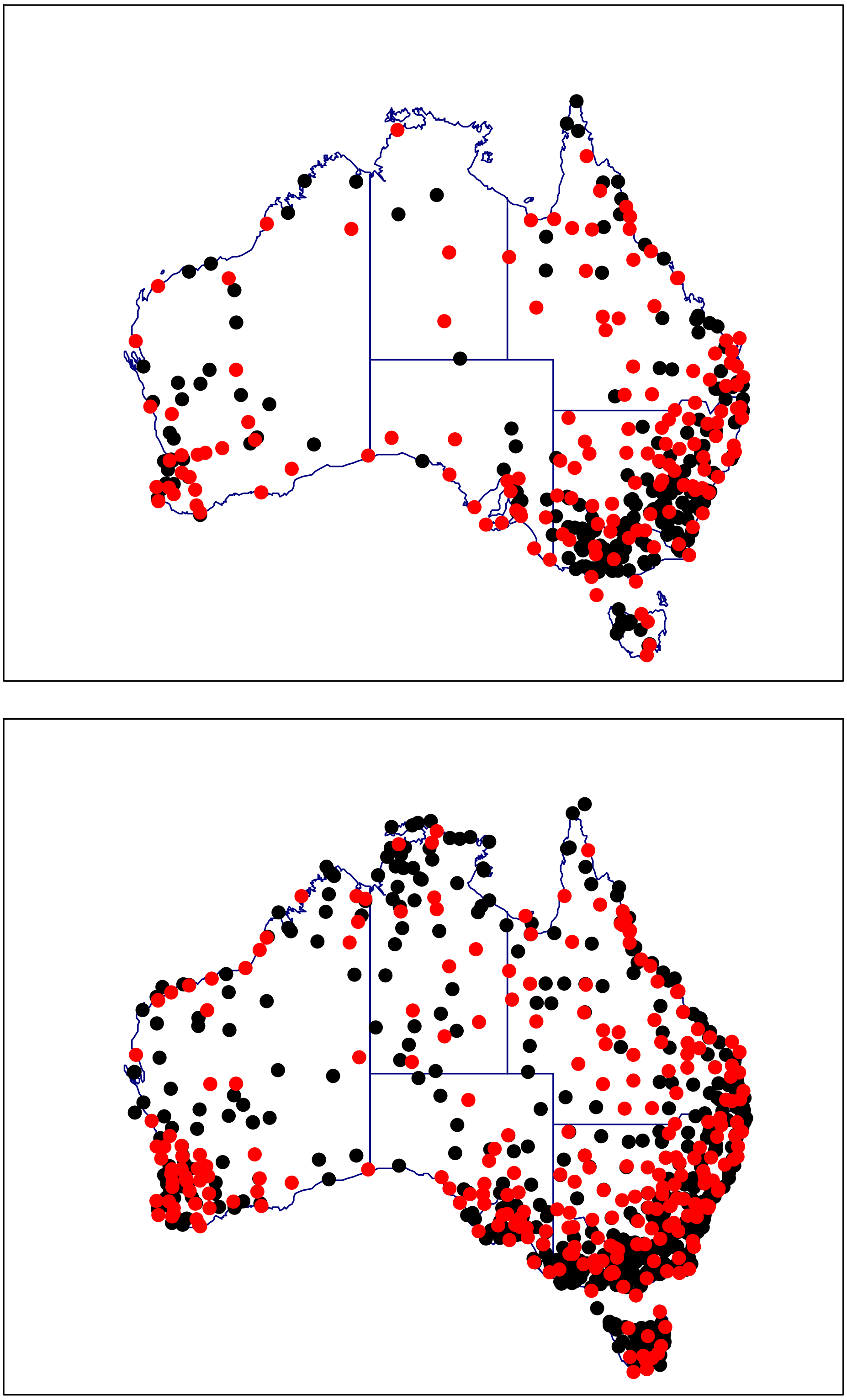


Fig. 1. Australian temperature observing network in 1930 (top) and 2010 (bottom). Sites with 40 or more years of data are shown in red.

2.1 The Australian temperature observing network

There are 1691 Australian sites, with greatly varying periods of record, that have daily temperature data in the Bureau of Meteorology climate database[3]. 761 of these are currently operating[4]. At one end of the scale, 31 sites (16 of which are currently operating) have 100 years or more of available daily data, with 184 sites having 50 years or more of data, and 506 sites 30 years or more; at the other end of the scale, 240 sites have less than five years of data, of which only 29 are still operating. A summary of the number of sites is shown in Fig. 2.

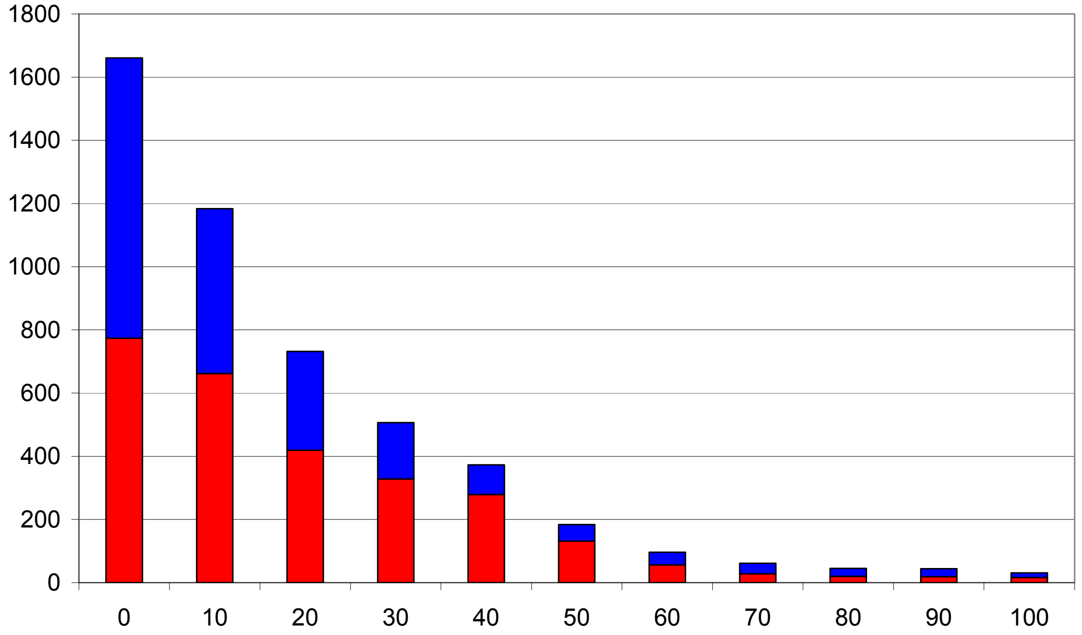


Fig. 2. Number of Australian temperature sites with at least N years of digitised daily temperature data (as of May 2011). The number of sites that are currently open is shown in red.

The distribution of sites is uneven over Australia (Fig. 1; also Jones and Trewin, 2002), with sites most heavily concentrated in the more densely populated parts of southeastern and southwestern Australia and near the east coast. Elsewhere coverage is much more sparse. Many sites are more than 100 kilometres from their nearest neighbour, with two ACORN-SAT locations (Giles and Rabbit Flat) more than 200 kilometres from their nearest neighbour, and substantial areas in the western and eastern interior have no observations at all. Overall, except in the most densely populated regions, network density is substantially lower than in Europe or the continental United States, with Canada being perhaps the best analogue.

Historically, there have been about ten times as many rainfall sites as temperature sites in the Australian observing network, although the data voids in the interior are similar for both elements (Jones et al., 2009), reflecting the difficulties in making observations of any kind in regions with little or no permanent settlement. The disparity between the number of rainfall and temperature stations reflects the historic role of rainfall as a major limiting factor on development in Australia, especially in agriculture.

---

[3] This does not include sites which have monthly data but no daily data, or no digitised data at all – the number of such sites is unknown. All are likely to have closed prior to 1965. The total also excludes offshore island and Antarctic sites.

[4] Defined as having reported temperature at least once between 1 January 2012 and 8 February 2012.

2.2 Types of sites

Several types of sites measure temperature in Australia. In broad terms these can be categorised into three groups. At some locations records may fall into different categories over time – for example, a co-operative site with manually-recorded observations may be replaced by an automatic weather station.

(a) *Bureau-staffed sites*. At these sites observations are made by Bureau of Meteorology employees, who have specific training in observing. Most, although not all, also launch balloon-borne radiosondes that are used for making upper-air observations. All Bureau-staffed sites now have automatic weather stations (AWSs) and report data at a 1-minute temporal resolution. Prior to the introduction of automatic instrumentation, most made full synoptic observations seven or eight times per day, and more limited reports (which included temperature and dewpoint) hourly or half-hourly (although many of these observations have not been digitised – see section 2.5).

(b) *Co-operative sites.* At these sites manual observations are performed by non-Bureau personnel (some paid, some voluntary). Historically many of these sites were located at post offices, although the number of post office sites has reduced in recent decades due to a combination of factors, including the relocation of many sites to airports, development in town centres, and the corporatisation of Australia Post. Numerous sites, including some of the most valuable long-term records, were located at lighthouses or other marine facilities such as pilot stations, and a significant number were operated by other government agencies, particularly state agriculture departments or equivalent (mostly on research farms) and airport authorities. All of these sites measure daily maximum and minimum temperature (the variables of interest for ACORN-SAT); the frequency of fixed-hour synoptic observations varies from one to eight per day, with two (at 09:00 and 15:00 local time) being the most common. The availability of other measurements, such as cloud, wind, dewpoint and terrestrial minimum temperature, varies from site to site.

(c) *Automatic weather stations*. Fully automated weather stations at these sites that transmit data electronically to the Bureau. (At some, the electronic data are supplemented by manual observations of elements such as cloud and evaporation, that are not currently measured by automated methods.) The first automatic weather stations that measured maximum and minimum temperature were installed in the 1980s, although they were not used at any ACORN-SAT location before 1992. These sites generally have the capacity to report at 1-minute temporal resolution, although due to communications limitations many have reported only hourly or half-hourly for most or all of their history, and a few have reported only at three-hourly intervals.

2.3 Changes to instruments and instrument exposure over time

The Stevenson screen has been the standard shelter for thermometers in Australia since the formation of the Bureau of Meteorology as a federal organisation in 1908. It had been progressively introduced through various parts of Australia, particularly Queensland and South Australia, from the late 1880s onwards, but was not in widespread use in New South Wales or Victoria before 1900. Almost all sites had a Stevenson screen in place by 1910, although a very small number of non-standard screens (mostly using the Stevenson screen design but with a metal rather than wooden roof) remained until mid-century.

A wide variety of instrument exposures existed prior to the introduction of the Stevenson screen. Probably the most common alternative exposure was the Glaisher stand, which was not fully enclosed like the Stevenson screen, but had no floor and an open front that faced southward. However, there was little standardisation either within or between states, with many instruments in locations such as improvised wall-mounted screens, underneath verandahs, or in unheated rooms indoors. In general these alternative exposures (except the indoor ones) were substantially warmer than a Stevenson screen for summer maximum temperature, with smaller differences for winter maximum temperatures and for minimum temperatures throughout the year (Parker, 1994). A 60-year set of parallel observations at Adelaide (Fig. 3) showed a warm bias in maximum temperatures measured using the Glaisher stand relative to those measured in the Stevenson screen (Nicholls et al., 1996), that ranged from 0.2 to 0.6°C in annual means, and reached up to 1.0°C in mean summer maximum temperatures and 2-3°C on some individual hot days, most likely due to heat re-radiated from the ground, from which the floorless Glaisher stand provides no protection. Minimum Glaisher stand temperatures tended to have a cool bias of 0.2-0.3°C all year, and the diurnal temperature range thus has a positive bias.

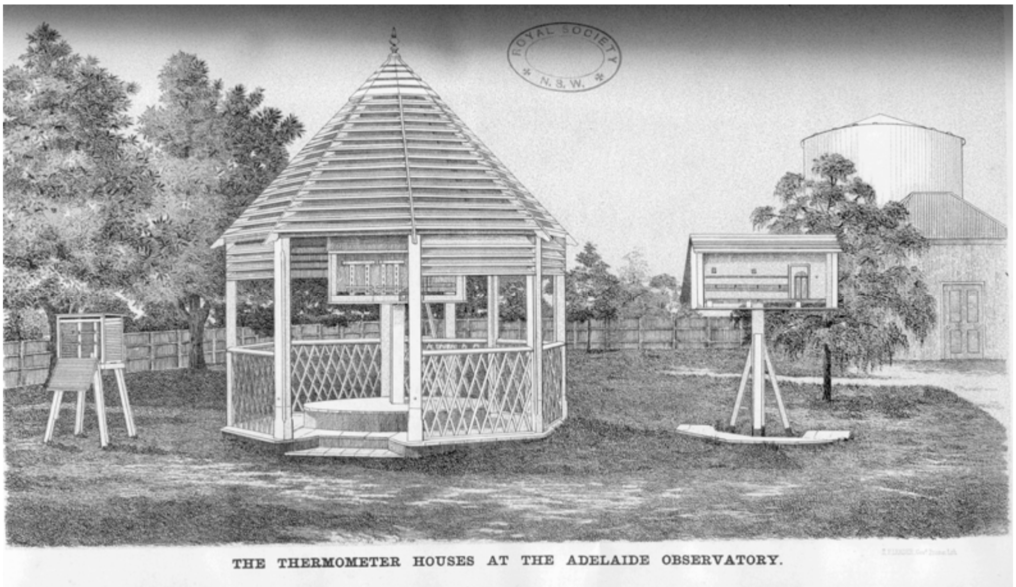


Fig. 3. Illustration of instrument comparison at Adelaide, showing a Stevenson screen (left), an octagonal 'thermometer house' (centre) and a Glaisher stand (right).

At many locations, mercury-in-glass thermometers have been replaced by automated temperature probes. Unlike many countries, Australia retained the traditional Stevenson screen as the standard instrument shelter when installing automated instruments. There is little evidence of any systematic change in mean temperatures with the introduction of automatic weather stations, although in many cases the change also involved a site change, which in some cases did have a significant impact on mean temperatures – a phenomenon that has also been noted in the United States by Guttman and Baker (1996). There is some indication of a small (less than 0.2°C) increase in diurnal temperature range, most likely because of the faster response time of automatic probes relative to mercury-in-glass thermometers. This finding is consistent with international experience (Trewin, 2010).

2.4 Site numbering

All Bureau of Meteorology sites are allocated a six-digit site number of the form TDDNNN, which has the following structure:

- T (type). This has the value 0 for all mainland Australian temperature sites, 2 for offshore islands (and, historically, pre-independence data from Papua New Guinea and some Pacific islands), and 3 for Antarctic and sub-Antarctic locations. The definition of “offshore islands” has been somewhat inconsistent over time, with values of both 0 and 2 being used for different sites on the Bass Strait islands and on the near-coastal islands off the north coast of the Northern Territory.
- DD (district). Australia is divided into 99 districts (some divided into further sub-districts), numbered from 01 (North Kimberley, Western Australia) to 99 (Flinders Island and associated islands, Tasmania). Where T is 2 or 3, DD is set to 00.
- NNN (number). This is a three-digit number, that is now normally allocated sequentially when a site opens. (Two districts have exhausted their allocation of numbers and are now using a value of T = 1 for some sites that measure only rainfall, but no temperature sites are yet affected).

Hence, a site with a number 040842 is a site in district 40 (southeast Queensland).

Current policy is that when a site moves significantly, the old site is normally closed and a new site is opened at the new location under a new number. This policy has varied over time, as has the definition of what constitutes a ‘significant’ move, and there are many past instances of substantial moves (up to several kilometres, or several hundred metres in elevation) without a change of site number.

2.5 Digitisation of temperature data

Large quantities of Australian temperature data have not been digitised, and are only available in paper form, rendering them effectively inaccessible for further analysis at this stage. Historically this has been a major barrier to the development of long-term daily temperature data sets.

The extent to which temperature data are digitised/not digitised is as follows:

(a) *Monthly means.* Most monthly mean maximum and minimum temperature data are digitised. The most significant exception is the period between 1957 and 1964 when a substantial number of sites have no digitised data at any time resolution. Experience with archival material from individual sites also suggests the existence of some undigitised monthly data prior to 1938 (in Queensland), 1925 (in South Australia and the Northern Territory) and 1907 (all states). The absence of such data, should they exist, would have only a minor impact on the overall network but could conceivably be concealing the existence of a few sites with century-long data sets.

(b) *Daily maximum and minimum temperature*. Until the late 1990s, most daily maximum and minimum temperature data were not digitised prior to 1957, with significant amounts of undigitised data also for the 1957-1964 period. At that time, in general, only Bureau-staffed sites had digitised daily data prior to 1957, and only a few major-city sites (Melbourne, Sydney, Adelaide, Brisbane, Darwin) had any digitised daily data pre-1939. In more recent years, substantial effort has been devoted to digitising more pre-1957 data, especially as part of the

CLIMARC project (see below). Large quantities of daily pre-1957 data remain to be digitised, although only a relatively small number of ACORN-SAT locations are in this situation.

(c) *Synoptic (three-hourly) data*. At most Bureau-staffed sites three-hourly synoptic temperature data are fully digitised, except for pre-1955 data at some major city sites. Most sites involved in post-1996 digitisation projects (see below) have their synoptic observations digitised for the period covered by those projects (mostly pre-1957). At most other sites, prior to 1987, only observations at 09:00 and 15:00 local time have been digitised, with observations at other times (if any) undigitised.

(d) *Hourly and higher-resolution data.* Most of these data are generated by automated systems and reach the Bureau in digital form. Some Bureau-staffed sites (especially major cities) made manual hourly observations prior to the introduction of automatic weather stations. Most of these observations are undigitised, although some major city sites have digitised hourly observations back into the 1980s. The hourly and half-hourly data are mostly derived from messages produced for aviation (METARs) and for simplicity will be referred to as METAR data later in this report.

Since 1996 a number of projects have been undertaken to digitise parts of the Australian daily and sub-daily temperature record. The largest of these was the CLIMARC project (Clarkson et al., 2001), a collaborative venture between the Bureau of Meteorology and the Queensland Government. Other projects have involved other agricultural or marine agencies, and some individual researchers. The effect of these projects has been to greatly improve the availability of daily data from the pre-1957 period, particularly pre-1940; the number of sites with available data from 1910 or earlier has increased from five in 1996 to 65 now. (For simplicity data digitised through such projects will be referred to in this report as CLIMARC data). However, large quantities of data remain to be digitised and there are no immediate plans to digitise the bulk of the remaining undigitised data.

Some of the more recently-processed CLIMARC data have not yet been through Bureau of Meteorology quality control procedures and have therefore not yet been incorporated in the main Bureau climate database (Australian Data Archive for Meteorology; ADAM). Such data have been considered for inclusion in ACORN-SAT, whilst being subject to a particularly high level of scrutiny in quality control (see section 6).

3. SELECTION OF THE ACORN-SAT NETWORK

In order to define a set of locations that suitably characterises spatial and temporal climate variability, the available locations must be selected in a manner that optimises a range of relevant criteria. For example, the length of record and the spatial extent of coverage should be maximised from the available network. While the number of sites in a merged or composite record should be minimised, this requirement has to be balanced against the requirements for maximising the length and coverage of the network. Optimisation of the network therefore requires a blend of objective and heuristic methods of selection and treatment, to ensure that the best possible data sets are selected.

Only some of the locations in any meteorological observation network are suitable for use in long-term climate change analyses. Most have too little data (less than 30 years), and some have excessive missing data, have poor sites or unreliable observation quality, or are otherwise unsuitable.

The ideal criteria, that are met at very few places in the world (and none in Australia) for locations used in a climate change analysis include:

- A long period (preferably 100 years or more) of continuous data with few or no missing observations.
- No site changes, changes in observation practices or instruments, or significant changes in local site environment.
- Located well outside any urban areas.

Whilst selecting a network that completely meets these criteria is impossible, it is possible to select a network that is acceptably large to sample climate change and variability adequately and that still approaches these criteria closely enough for its data to be fit for that purpose. Jones and Trewin (2002) found that a network of 100 to 200 locations was sufficient to define temperature variability to a reasonable degree of accuracy over Australia, while Vose and Menne (2004) obtained similar results for the similarly-sized continental United States. In practice, given the number of long-term sites available (Fig. 2), constructing a network of this size requires making use of most of the sites with an acceptably long record, with careful homogenisation required (section 7) to obtain consistent records suitable for use in climate change analyses.

3.1 What is meant by homogenisation and composite sites?

The degree of consistency in a temperature record over time is often referred to as data homogeneity. In any compiled temperature network, several factors can introduce changes in the data (as a function of time) that are not related to physical changes in climate. This subject is dealt with in more detail in section 3 below. Such changes are referred to as inhomogeneities. The detection and correction for artificial changes is termed homogenisation.

The intention in developing the ACORN-SAT data set is to produce long-records of temperature observations at sites which can be used for the monitoring of climate variability and climate change. Most instrumental temperature networks change over time due primarily to the importance of a fixed network for monitoring climate variability and change being less appreciated in the past than in more recent times. In addition, a range of socio-economic considerations, such as changes in demographics and infrastructure, affect the ability to maintain a fixed network over a long period of time.

.

Due to changes in the observing network, it is necessary to merge records from different sites to maximise temporal coverage. The need to composite (merge) sites to form long records can lead to the homogeneity of the record being compromised over time, which must be dealt with through appropriate analyses. Throughout this report, compositing of sites refers to the process of merging nearby sites to create a “single” location series, taking into account differences between the raw data at the sites that are due to the absolute differences in the climate between them, e.g. one site/location might be inherently warmer than another by a few tenths of a degree.

The second issue that arises in producing long, homogeneous records relates to changes occurring over time at individual sites, that can also introduce artificial or non-climate related changes in the recordings.

3.2 Previous networks used for climate change analyses in Australia

Two major temperature data sets have been used for climate change analyses in Australia. They are shown in Fig. 4.

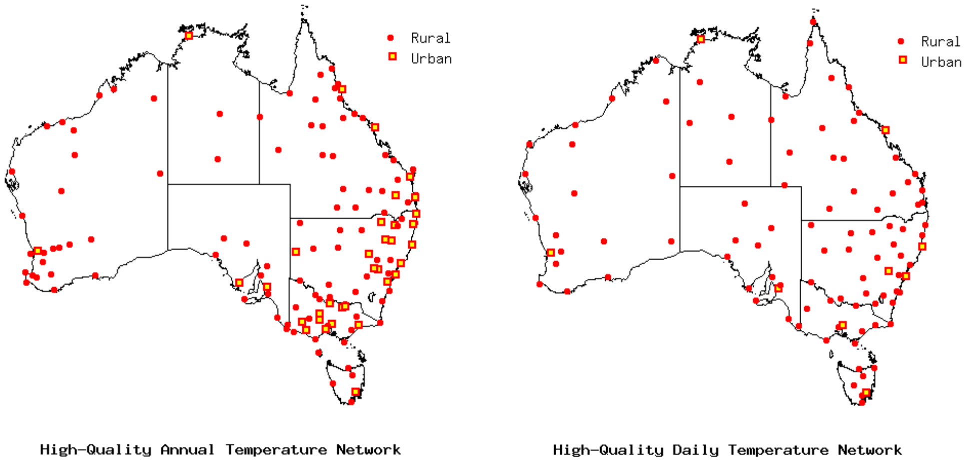


Fig. 4. Previously existing long-term high-quality site networks: (left) annual (Della-Marta et al, 2004), (right) daily (Trewin, 2001a).

(a) *Annual data set*. This set was originally developed by Torok and Nicholls (1996) and enhanced by Della-Marta et al. (2004). The original Torok and Nicholls data set included 224 locations[5], 50 of which were classified as urban, and incorporated all locations with a starting date of 1915 or earlier for monthly data, and still operating at the time of the study. Della-Marta et al. reduced this data set to 134 locations (34 urban), with the others being removed because of closure (without acceptable replacement), excessive missing data or poor site or observation quality. The non-urban component of this data set is the basis, at the time of writing, for annual mean temperature anomalies for Australia reported routinely by the Bureau of Meteorology (e.g. Bureau of Meteorology, 2011).

(b) *Daily data set*. This set was developed by Trewin (2001a), and consists of 103 loctions (four of which were classified as urban, although on a more limited definition than that used by Torok and Nicholls). The core of this data set was based on the network of Reference Climate Stations (RCSs), a network selected by the Bureau of Meteorology (1995), in response to a request made by the World Meteorological Organization in 1990 for its member nations to identify a network of recommended reference climate sites. Sixteen RCSs were not included in the data set, mostly because they had either an inadequate length of record or duplicated other locations, and 25 additional locations were used, mostly to improve geographical coverage. Most of these locations were drawn from the annual temperature data set described above. In contrast with the annual data set, which concentrated on stations that opened in 1915 or earlier, the daily data set focused on locations with the greatest availability of digitised daily data at the time when it was defined (1997-98), resulting in the selection of a number of locations (particularly Bureau-staffed sites) which were not considered for the annual data set because they opened in the 1940s, but had some of the longest available daily data sets at that time. This data set is used for

---

[5] In this context a “location” may include a composite of records from two or more station numbers in close proximity.

the Bureau of Meteorology's monthly anomalies, and for analyses of changes in climate extremes.

An additional data set that is extensively used in routine Australian climate monitoring, but does not currently underlie climate change analyses, is a spatial data set developed as part of the Australian Water Availability Project (AWAP; Jones et al., 2009). This data set incorporates all available data and is not homogenised at the site level, but does take into account changes in network coverage, which is particularly important in topographically complex areas. In the context of the ACORN-SAT project, the major use of the AWAP data set is as a reference against which the ACORN-SAT can be compared (this forms part of a companion report to this volume). It has already been used for that purpose, in evaluating the results of the Trewin (2001a) and Della-Marta et al. (2004) data set (Jones et al., 2009).

3.3 The ACORN-SAT data set

The new ACORN-SAT data set consists of daily maximum and minimum temperature data for 112 locations (Fig. 5). At least 60 locations are available in every year from 1910, at least 85 in every year from 1946, and at least 99 in every year from 1957 (Fig. 6). From 1971 onwards there are no more than two locations out of the 112 missing[6] in any individual year. For the purposes of this report, the ACORN-SAT data set is considered to extend from 1910 to 2009, although ongoing updates are planned, the first of which was undertaken in early 2012 to include data for 2010 and 2011.

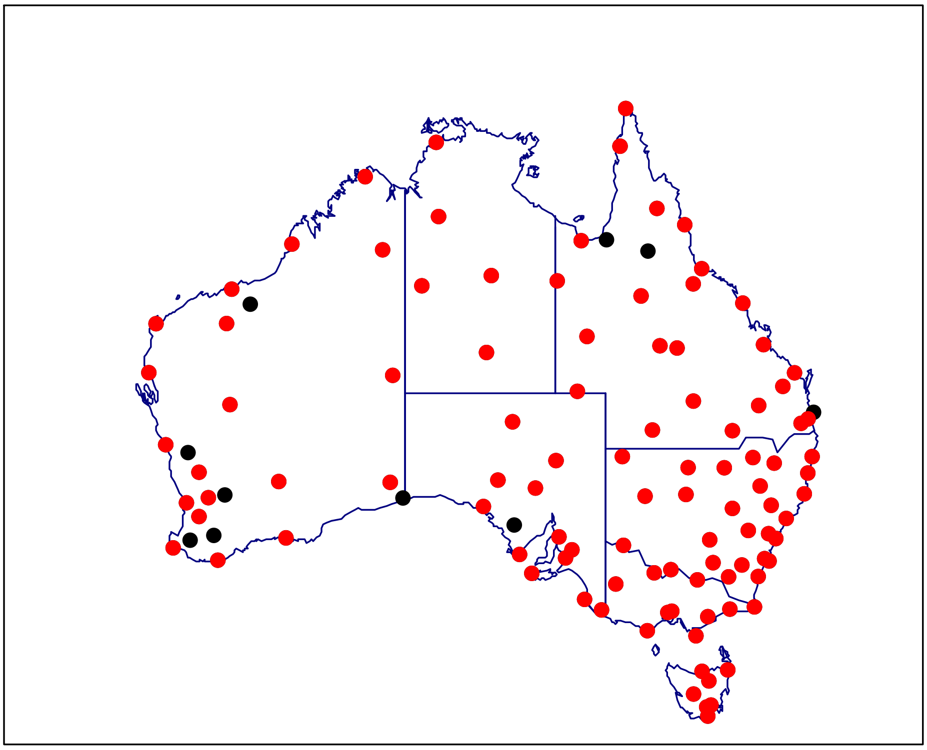


Fig. 5. Locations in ACORN-SAT data set. Locations that were not in the previous daily data set (Trewin, 2001a) are shown in black.

---

[6] In this context, with fewer than 183 observations (50%) available in that year for either maximum or minimum temperature.

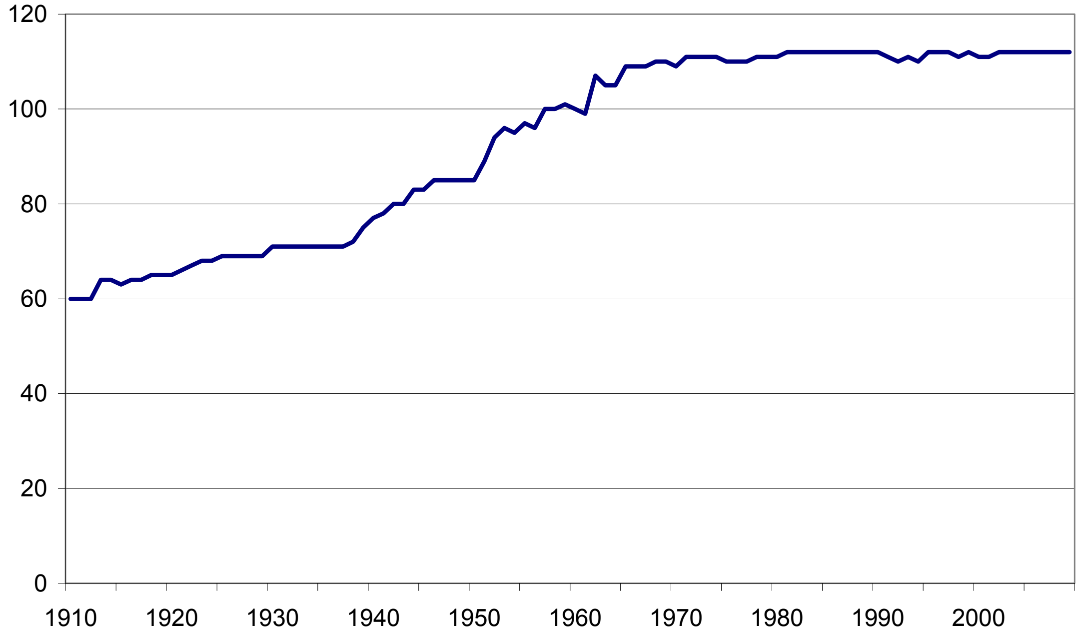


Fig. 6. Number of ACORN-SAT locations with available data, by year.

The Trewin (2001a) network forms the basis for the ACORN-SAT data set. Compared with the Trewin (2001a) set, ten locations have been added and one deleted. The locations that have been added are locations that were in the Della-Marta et al. (2004) data set but not in the Trewin (2001a) set, and now have digitised daily data extending back to 1915 or earlier. The location that was deleted (Tewantin, in southeast Queensland) was found, upon digitisation of its pre-1957 data, to have extremely poor data quality for large parts of the period from 1930 to 1950; furthermore, the site quality was also very poor in the decade preceding a site move in 1996.

The ACORN-SAT data set includes only data from 1910 or later. As discussed in section 1, the Stevenson screen only became near-universal in Australia from 1910 onwards, and most pre-1910 data cannot be considered homogeneous with Stevenson screen data. A separate study is currently under way (Ashcroft et al., 2011) to investigate what use can be made of pre-1910 Australian temperature data, but this analysis is outside the scope of the ACORN-SAT project. At one site (Eucla), a Stevenson screen was not installed until 1913 and the ACORN-SAT data for that location starts in that year.

Table 1 shows a full listing of the ACORN-SAT locations.

Table 1 ACORN-SAT locations. Site number, location and elevation for site in operation in December 2011. Population from 2006 Census (not shown for sites more than 20 km from any urban centre with population above 100; * indicates population is that of a larger urban centre within 20 km of the named town). If first year is shown in italics, further undigitised data are believed to exist.

Site number	Name	Latitude (deg S)	Longitude (deg E)	Elevation (m)	First year of data	Population of urban centre
001019	Kalumburu	14.30	126.65	23	1941	413
002012	Halls Creek	18.23	127.66	422	1910	1211
003003	Broome	17.95	122.23	7	1910	11547
004032	Port Hedland	20.37	118.63	6	1912	11557
004106	Marble Bar	21.18	119.75	182	1910	194
005007	Learmonth	22.24	114.10	5	1975	
005026	Wittenoom	22.24	118.34	463	1951	
006011	Carnarvon	24.89	113.67	4	1910	5283
007045	Meekatharra	26.61	118.54	517	1926	798
008039	Dalwallinu	30.28	116.66	335	1957	593
008051	Geraldton	28.80	114.70	33	1910	27420
008296	Morawa	29.20	116.02	271	1925	597
009021	Perth	31.93	115.98	15	1910	1256035
009510	Bridgetown	33.96	116.14	150	1910	2324
009518	Cape Leeuwin	34.37	115.14	13	1910	1068*
009741	Albany	34.94	117.80	68	1910	25196
009789	Esperance	33.83	121.89	25	1910	9536
010092	Merredin	31.48	118.28	315	1912	2550
010286	Cunderdin	31.62	117.22	217	1950	686
010579	Katanning	33.69	117.56	310	1910	3808
010917	Wandering	32.67	116.67	275	1910	
011003	Eucla	31.68	128.88	93	1913	86
011052	Forrest	30.85	128.11	159	1946	
012038	Kalgoorlie-Boulder	30.78	121.45	365	1910	28242
013017	Giles	25.03	128.30	598	1956	
014015	Darwin	12.42	130.89	30	1910	105991
014825	Victoria River Downs	16.40	131.01	89	1965	
015135	Tennant Creek	19.64	134.18	376	1910	2919
015590	Alice Springs	23.80	133.89	546	1910	21622
015666	Rabbit Flat	20.18	130.01	340	1969	
016001	Woomera	31.16	136.81	167	1949	295
016098	Tarcoola	30.71	134.58	123	1921	
017031	Marree	29.65	138.06	50	1910	70
017043	Oodnadatta	27.56	135.45	117	1940	
018012	Ceduna	32.13	133.70	15	1939	2304
018044	Kyancutta	33.13	135.56	57	1930	513*
018192	Port Lincoln	34.60	135.88	9	1910	13044
021133	Snowtown	33.77	138.22	109	1910	405
022823	Cape Borda	35.75	136.60	158	1962	
023090	Adelaide	34.92	138.62	48	1910	1040719
023373	Nuriootpa	34.48	139.01	275	1957	4414
026021	Mount Gambier	37.75	140.77	63	1910	23494
026026	Robe	37.16	139.76	3	1910	1246
027045	Weipa	12.68	141.92	18	1959	2830
027058	Horn Island	10.58	142.29	4	1950	585
028004	Palmerville	16.00	144.08	204	1910	
029063	Normanton	17.69	141.07	18	1910	1100
029077	Burketown	17.75	139.54	6	1910	173
030045	Richmond (Qld)	20.73	143.14	211	1910	554

Table 1 (cont.) ACORN-SAT locations. Site number, location and elevation for site in operation in December 2011. Population from 2006 Census (not shown for sites more than 20 km from any urban centre with population above 100; * indicates population is that of a larger urban centre within 20 km of the named town). If first year is shown in italics, further undigitised data are believed to exist.

Site number	Name	Latitude (deg S)	Longitude (deg E)	Elevation (m)	First year of data	Population of urban centre
030124	Georgetown	18.30	143.53	302	1910	254
031011	Cairns	16.87	145.75	2	1910	98349
032040	Townsville	19.25	146.77	5	1940	128808
033119	Mackay	21.12	149.22	30	1910	66874
034084	Charters Towers	20.05	146.27	290	1910	7979
036007	Barcaldine	23.55	145.29	267	*1962*	1337
036031	Longreach	23.44	144.28	192	1910	2976
037010	Camooweal	19.92	138.12	231	1939	199
038003	Boulia	22.91	139.90	162	1910	205
038026	Birdsville	25.90	139.35	47	1954	115
039066	Gayndah	25.62	151.62	111	1910	1745
039083	Rockhampton	23.38	150.48	10	*1939*	60827
039128	Bundaberg	24.91	152.32	31	1910	46961
040004	Amberley	27.63	152.71	24	1941	140182*
040043	Cape Moreton	27.03	153.47	100	1910	250*
040842	Brisbane Airport	27.39	153.12	5	1949	1676389
042112	Miles	26.66	150.18	305	1910	1164
043109	St. George	28.05	148.59	199	1913	2410
044021	Charleville	26.41	146.26	302	1910	3278
045025	Thargomindah	27.99	143.81	131	*1957*	203
046037	Tibooburra	29.43	142.01	183	1910	161
046043	Wilcannia	31.56	143.37	75	*1957*	596
048027	Cobar	31.48	145.83	260	1910	4128
048245	Bourke	30.04	145.95	107	1910	2145
052088	Walgett	30.04	148.12	133	1910	1735
053115	Moree	29.49	149.85	213	*1957*	8083
055024	Gunnedah	31.03	150.27	307	1948	7542
056242	Inverell	29.78	151.11	582	1910	9749
058012	Yamba	29.43	153.36	27	1910	5514
059040	Coffs Harbour	30.31	153.12	5	1943	26353
060139	Port Macquarie	31.43	152.87	4	1910	39219
061078	Williamtown	32.79	151.84	9	1942	826
061363	Scone	32.03	150.83	221	*1965*	4624
063005	Bathurst	33.43	149.56	713	1910	28992
065070	Dubbo	32.22	148.58	284	1921	30574
066062	Sydney	33.86	151.21	39	1910	3641422
067105	Richmond (NSW)	33.60	150.78	19	1939	3641422*
068072	Nowra	34.95	150.54	109	*1955*	27478
068151	Point Perpendicular	35.09	150.80	85	*1946*	513
069018	Moruya Heads	35.91	150.15	17	1910	800
070351	Canberra	35.31	149.20	577	1939	368129
072150	Wagga Wagga	35.16	147.46	212	1910	46735
072161	Cabramurra	35.94	148.38	1482	*1962*	60
073054	Wyalong	33.93	147.24	245	*1965*	3191
074258	Deniliquin	35.56	144.95	94	1910	7431
076031	Mildura	34.24	142.09	50	1910	30016
078015	Nhill	36.31	141.65	139	1910	1915

Table 1 (cont.) ACORN-SAT locations. Site number, location and elevation for site in operation in December 2011. Population from 2006 Census (not shown for sites more than 10 km from any urban centre with population above 100; * indicates population is that of a larger urban centre within 20 km of the named town). If first year is shown in italics, further undigitised data are believed to exist.

Site number	Name	Latitude (deg S)	Longitude (deg E)	Elevation (m)	First year of data	Population of urban centre
080023	Kerang	35.72	143.92	78	1910	3780
082039	Rutherglen	36.10	146.51	175	1912	1990
084016	Gabo Island	37.57	149.92	15	1910	
084145	Orbost	37.69	148.47	63	1938	2097
085072	Sale	38.12	147.13	5	*1945*	13336
085096	Wilsons Promontory	39.13	146.42	95	1910	
086071	Melbourne	37.81	144.97	31	1910	3371888
087031	Laverton	37.86	144.76	20	1945	3371888*
090015	Cape Otway	38.86	143.51	82	1910	
091293	Low Head	41.06	146.79	28	1910	4266*
091311	Launceston	41.55	147.21	167	1910	71395
092045	Eddystone Point	40.99	148.35	20	1910	
094010	Cape Bruny	43.49	147.15	55	1923	
094029	Hobart	42.89	147.33	51	1918	128577
094220	Grove	42.99	147.07	65	1952	719
096003	Butlers Gorge	42.28	146.28	667	1944	

3.4 The role of site composites and comparisons

A substantial proportion of the records in ACORN-SAT require the combination of data from two or more site numbers. As noted in section 2, current policy is that a change of site number occurs whenever a significant site relocation takes place. As noted earlier, this policy has not been in place consistently through history; whilst many substantial site relocations in the past have been accompanied by a change in site number, many others have not.

The merging of sites to form a composite record requires one to account for systematic differences in temperature data or recording. The ideal situation is that a change substantial enough to warrant a change of site number is carried out with a substantial overlap between the two sites, sufficient to enable a good comparison between the two sites and appropriate adjustments to be determined. Current practice at the Bureau of Meteorology (although not, at the time of writing, formalised policy) is to carry out parallel observations for a minimum of two years, and preferably at least five, which is the practice recommended by the World Meteorological Organization. Most post-1995 site number changes at ACORN-SAT locations have involved such a period of parallel observations.

There are, however, many cases in the historic record, mostly prior to the mid-1990s when there was no general requirement for parallel observations, where site number changes occurred with no overlap at all, or with no useful overlap. There have also been cases (e.g. at Williamtown (NSW) in 2005) where a comparison was terminated after a few months because changes in the environment around the original site meant that it was no longer a useful comparison location (a comparison is only useful for as long as conditions at the comparison site are representative of those that existed before the comparison began), or where the comparison continued but analysis of the data and/or metadata indicated that there had been a substantial change in conditions at the original site. In these cases records from different site numbers need to be merged without overlap, and are treated in the same way as an inhomogeneity identified from metadata within a record from an individual site number (see section 7).

The site numbering convention for site comparisons has been somewhat confusing over time, partly because of a desire to retain a single site number for upper-air observations at those Bureau-staffed sites carrying them out. Each of the following numbering conventions has been used at one or more locations where an overlap occurs between an old (comparison) and new site:

- Old site retains old number, new site opens with new number.
- Old site switches to new number for the duration of the comparison, new site takes over old number from the start of its observations.
- New site opens under new number then switches to old number after end of comparison.

Table 2 shows the site numbers used for each location in the ACORN data set, and the dates over which they were used.

Table 2 Sites used for ACORN-SAT locations. (* - old site switched to comparison number while number switched to new site; # - comparison site switched to old number after end of comparison)

Location	Sites
Kalumburu	1021 Kalumburu Mission (1941-2005), 1019 Kalumburu (1998-)
Halls Creek	2011 Halls Creek (1910-1952), 2012 Halls Creek (1949-), 2071 Halls Creek Comparison* (1996-2001)
Broome	3002 Broome Post Office (PO) (1910-1944), 3003 Broome Airport (AP) (1939-), 3089 Broome Comparison* (1995-2000)
Port Hedland	4002 Port Hedland PO (1912-1948), 4032 Port Hedland AP (1948-), 4104 Port Hedland Comparison* (1998-2001)
Marble Bar	4020 Marble Bar (1910-2003), 4106 Marble Bar (2000-)
Learmonth	5007 Learmonth (1975-)
Wittenoom	5026 Wittenoom (1952-)
Carnarvon	6062 Carnarvon PO (1910-1950), 6011 Carnarvon AP (1945-)
Meekatharra	7046 Meekatharra PO (1926-1950), 7045 Meekatharra AP (1950-), 7204 Meekatharra Comparison* (1998-2001)
Dalwallinu	8039 Dalwallinu PO (1957-)
Geraldton	8050 Geraldton Town (1910-1941), 8051 Geraldton AP (1942-)
Morawa	8093 Morawa PO (1925-2005), 2896 Morawa AP (1997-)
Perth	9034 Perth Regional Office (1910-1962), 9021 Perth AP (1944-), 9250 Perth AP Comparison* (1997-2001)
Bridgetown	9510 Bridgetown (1910-)
Cape Leeuwin	9518 Cape Leeuwin (1910-)
Albany	9500 Albany (1910-1965, 2002-), 9741 Albany AP (1965-)
Esperance	9541 Esperance PO (1910-1969), 9789 Esperance Met Office (MO) (1969-)
Merredin	10093 Merredin Research Stn (1912-1985), 10092 Merredin (1966-)
Cunderdin	10035 Cunderdin (1950-1996), 10286 Cunderdin AP (1997-)
Katanning	10579 Katanning (1910-)
Wandering	10648 Wandering Shire (1910-2003), 10917 Wandering (1998-)
Eucla	11003 Eucla (1913-)
Forrest	11004 Forrest MO (1946-1995), 11052 Forrest (1993-)
Kalgoorlie-Boulder	12039 Kalgoorlie PO (1910-1953), 12038 Kalgoorlie-Boulder AP (1944-)
Giles	13017 Giles (1956-)
Darwin	14016 Darwin PO (1910-1942), 14015 Darwin AP (1941-), 14040 Darwin AP Comparison* (2001-2007)
Victoria River Downs	14825 Victoria River Downs (1965-)
Tennant Creek	15087 Tennant Creek PO (1910-1970), 15135 Tennant Creek MO (1969-)
Alice Springs	15540 Alice Springs PO (1910-1943), 15590 Alice Springs AP (1944-)
Rabbit Flat	15548 Rabbit Flat (1969-1998), 15666 Rabbit Flat (1996-)
Woomera	16001 Woomera (1949-)
Tarcoola	16044 Tarcoola (1921-1999), 16098 Tarcoola (1997-)
Marree	17031 Marree (1939-), 17024 Farina (1910-1939)
Oodnadatta	17043 Oodnadatta AP (1940-1985, 1994-), 17114 Oodnadatta Police (1985-1991)
Ceduna	18011 Ceduna PO (1939-51), 18012 Ceduna AP (1939-)
Kyancutta	18044 Kyancutta (1930-)
Port Lincoln	18070 Port Lincoln PO (1910-2002), 18192 Port Lincoln AP (1992-)
Snowtown	21046 Snowtown PO (1910-2001), 21133 Rayville Park (1998-)
Cape Borda	22801 Cape Borda (1962-2007), 22823 Cape Borda (2002-)
Adelaide	23000 Adelaide West Terrace (1910-1979), 23090 Adelaide Kent Town (1977-)
Nuriootpa	23321 Nuriootpa (1957-1999), 23373 Nuriootpa (1996-)
Mount Gambier	26020 Mount Gambier PO (1910-1947), 26021 Mount Gambier AP (1942-)
Robe	26026 Robe (1910-)

Table 2 (cont.) Sites used for ACORN-SAT locations. (* - old site switched to comparison number while number switched to new site; # - comparison site switched to old number after end of comparison)

Location	Sites
Weipa	27042 Weipa (1959-1994), 27045 Weipa AP (1992-)
Horn Island	27022 Thursday Island MO (1952-1993), 27021 Thursday Island (1992-1996), 27058 Horn Island (1995-)
Palmerville	28004 Palmerville (1910-)
Normanton	29041 Normanton PO (1910-2001), 29063 Normanton AP (2001-)
Burketown	29004 Burketown PO (1910-2001), 29077 Burketown AP (2002-)
Georgetown	30018 Georgetown PO (1910-2004), 30124 Georgetown AP (2004-)
Richmond (Qld)	30045 Richmond PO (1910-)
Cairns	31010 Cairns PO (1910-1947), 31011 Cairns AP (1942-)
Townsville	32040 Townsville AP (1940-), 32178 Townsville Comparison* (1994-2000)
Mackay	33046 Mackay PO (1910-1950), 33047 Te Kowai (1947-1965), 33119 Mackay MO (1959-), 33297 Mackay MO Comparison* (1995-2001)
Charters Towers	34002 Charters Towers PO (1910-1992), 34084 Charters Towers AP (1992-)
Barcaldine	36007 Barcaldine (1962-)
Longreach	36030 Longreach PO (1910-1969), 36031 Longreach MO (1966-), 36167 Longreach MO Comparison* (1996-1997)
Camooweal	37010 Camooweal (1939-)
Boulia	38003 Boulia (1910-)
Birdsville	38002 Birdsville Police (1954-2005), 38026 Birdsville AP (2000-)
Gayndah	39039 Gayndah PO (1910-2009), 39066 Gayndah AP (2003-)
Rockhampton	39083 Rockhampton AP (1939-)
Bundaberg	39015 Bundaberg PO (1910-1990), 39128 Bundaberg AP (1990-)
Amberley	40004 Amberley (1942-), 40910 Amberley Comparison* (1997-1998)
Cape Moreton	40043 Cape Moreton (1910-)
Brisbane Airport	40223 Brisbane AP (1949-2000), 40842 Brisbane AP (1994-)
Miles	42023 Miles (1910-2005), 42112 Miles (1997-)
St. George	43034 St. George PO (1913-1997), 43109 St. George AP (1997-)
Charleville	44022 Charleville PO (1910-1948), 44021 Charleville AP (1949-), 44221 Charleville AP Comparison* (2003-2006)
Thargomindah	45017 Thargomindah PO (1957-1999), 45025 Thargomindah AP (1999-)
Tibooburra	46037 Tibooburra PO (1910-)
Wilcannia	46043 Wilcannia (1957-)
Cobar	48030 Cobar PO (1910-1965), 48027 Cobar MO (1962-), 48244 Cobar MO Comparison* (1997-2000)
Bourke	48013 Bourke PO (1910-1996), 48239 Bourke AP (1994-1998), 48245 Bourke AP (1999-)
Walgett	52026 Walgett PO (1910-1993), 52088 Walgett AP (1993-)
Moree	53027 Moree PO (1957-1964), 53048 Moree MO (1965-1998), 53115 Moree MO (1995-)
Gunnedah	55024 Gunnedah Research Station (1948-)
Inverell	56017 Inverell PO (1910-1997), 56242 Inverell (1995-)
Yamba	58012 Yamba (1910-)
Coffs Harbour	59040 Coffs Harbour (1943-)
Port Macquarie	60026 Port Macquarie (1910-2003), 60139 Port Macquarie AP (1995-)
Williamtown	61078 Williamtown (1942-)
Scone	61089 Scone Soil Conservation (1965-2000), 61363 Scone AP (1997-)
Bathurst	63005 Bathurst Agricultural Research (1910-), 63305 Bathurst Comparison# (1996-1998)
Dubbo	65012 Dubbo (1921-1999), 65070 Dubbo AP (1993-)
Sydney	66062 Sydney (1910-)
Richmond (NSW)	67033 Richmond (1939-1994), 67105 Richmond (1993-)
Nowra	68076 Nowra (1955-2000), 68072 Nowra (2000-)
Point Perpendicular	68034 Point Perpendicular (1946-2004), 68151 Point Perpendicular (2001-)

Table 2 (cont.) Sites used for ACORN-SAT locations. (* - old site switched to comparison number while number switched to new site; # - comparison site switched to old number after end of comparison)

Location	Sites
Moruya Heads	69018 Moruya Heads (1910-)
Canberra	70014 Canberra AP (1939-2010), 70351 Canberra AP (2008-), 70338 Canberra AP Comparison* (1995-1997)
Wagga Wagga	72151 Kooringal (1910-1950), 72150 Wagga Wagga AP (1948-)
Wyalong	73054 Wyalong (1965-)
Deniliquin	74128 Deniliquin (1910-2003), 74258 Deniliquin AP (1997-)
Mildura	76077 Mildura PO (1910-1949), 76031 Mildura AP (1946-)
Nhill	78031 Nhill (1910-2008), 78015 Nhill AP (2003-)
Kerang	80023 Kerang (1910-)
Rutherglen	82039 Rutherglen (1912-)
Gabo Island	84016 Gabo Island (1910-)
Orbost	84030 Orbost (1938-2011), 84145 Orbost (2000-)
Sale	85072 East Sale (1945-), 85298 East Sale Comparison* (1996-2005)
Wilsons Promontory	85096 Wilsons Promontory (1910-)
Melbourne	86071 Melbourne (1910-)
Laverton	87031 Laverton (1945-)
Cape Otway	90015 Cape Otway (1910-)
Low Head	91057 Low Head (1910-2001), 91293 Low Head (1998-)
Launceston	91049 Launceston Pumping Station (1910-1946), 91104 Launceston AP (1939-2009), 91311 Launceston AP (2004-)
Eddystone Point	92045 Eddystone Point (1910-)
Cape Bruny	94010 Cape Bruny (1923-)
Hobart	94029 Hobart (1918-)
Grove	94069 Grove (1952-2010), 94220 Grove (2004-)
Butlers Gorge	96003 Butlers Gorge (1944-1993, 2008-), 96071 Lake St. Clair (1989-)

3.5 Potential future additions to the ACORN-SAT data set

A number of locations in the current annual temperature data set were not considered for addition to the ACORN-SAT data set at this time because they currently lack digitised daily data prior to 1957. These locations will be considered for addition if and when their pre-1957 data are digitised, which will require an as yet unidentified source of funding.

Whilst experience with the CLIMARC project indicates that it is difficult to define the amount of undigitised data precisely in advance (a number of CLIMARC locations had either much more or much less available daily data on paper than had been anticipated based on the digitised monthly data), it is likely that there are 15-20 further locations that have the potential to be added to ACORN-SAT at some future date.

4. INSTRUMENT SITING AND OBSERVATION STANDARDS

Temperatures measurements can be influenced in a wide variety of ways, e.g. by the nature of the site where the instruments are exposed, by the instrumentation itself, and by differences in observing practices. A detailed review of these influences can be found in Trewin (2010). Some of the more important site influences that affect the homogeneity and representativeness of long-term temperature records in Australia include:

- The local site exposure or environment, including the presence (or absence) of nearby buildings, structures or artificial surfaces (most common in towns or cities, although not necessarily large ones).
- Local topography, including site elevation and, for minimum temperature, whether the site is on flat ground, on a slope, in a valley or on top of a hill.
- Marine exposure at coastal sites: the more exposed a site is to the ocean, the cooler its maximum temperatures will be (especially in summer) and the warmer its minimum temperatures will be (especially in winter).
- Changes in vegetation in the region around the site. In particular, a change from a watered lawn to unwatered grass (or vice versa) was found in numerous cases to have a substantial impact on maximum temperatures during drier periods of the year.

Changes in these factors, either through a site change that places the site in a different environment, or (in the case of the first and last points) a change in the surrounding site environment *in situ*, can create an inhomogeneity in temperature records.

It is important to note that if a site environment is constant, even if it does not conform strictly to standards, that non-conformance in itself will not generate an inhomogeneity - for example, if a site is too close to a building but that building has been there unchanged throughout the period of record.

Other potential sources of inhomogeneity in temperature records include changes in instrument standards, and changes in observing practices. The former has not been a major factor in Australia during the period covered by the ACORN-SAT data set; the Stevenson screen has been the standard instrument shelter throughout and the one major change in instruments, from mercury-in-glass thermometers to automated probes, does not appear, from comparisons carried out at the time, to have had any appreciable impact in itself on temperature measurements. Observation practices have changed over time, the most significant change being a change in observation times in 1964. The impact of this change and two other changes affecting all of, or large parts of, the network (metrication and the introduction of automatic weather stations) is discussed in section 8.

4.1 What standards exist for the siting of temperature measurements in Australia, and how well are they followed?

Standards for instrument exposure and siting in Australia are laid down by *Observations Specification 2013.1* (Bureau of Meteorology, 1997). Among the guidelines are:

- Sites should be representative of the mean conditions over the area of interest (e.g., an airport or climatic region), except for sites specifically intended to monitor localised phenomena.
- The instrument enclosure[7] (if there is one) should be level, clearly defined and covered with as much of the natural vegetation of the area that can be kept cut to a height of a few centimetres.
- The distance of any obstruction should be at least four times the height of the obstruction away from the enclosure. (This criterion is primarily directed at elements other than temperature; for temperature the last guideline is more important.)

---

[7] The instrument enclosure is the area of land on which the instruments are situated.

- The base of the instrument shelter should be 1.1 metres above the ground, with the thermometers approximately 1.2 metres above the ground.
- If no instrument enclosure is provided, the shelter should be installed on level ground covered with either the natural vegetation of the area or unwatered grass, and should be freely exposed to the sun and wind. It should not be shielded by or close to trees, buildings, fences, walls or other obstructions, or extensive areas of concrete, asphalt, rock or other such surfaces – a minimum clearance of five times the width of the hard surface is recommended.

Compliance with these standards has been variable over time due to several factors. Most sites at airports or similar locations meet these standards well. Town centre sites have been more problematical, with the most common issue being inadequate distance from buildings, structures or hard surfaces. (This issue is largely unconnected to the size of the town, with some of the most compromised sites in towns with populations less than 1000). However, with careful homogenisation, most of these records are still usable for long-term analysis, as the issues tend to manifest themselves as step changes (at the point in time when, for example, a hard-surfaced car park is built nearby) and are thus amenable to adjustment in most cases.

Priorities in the selection of the highest-quality sites have changed over time. Up until the early 1990s, automatic weather stations were not considered for long-term data sets. Hence, a major consideration at that time in the selection of Reference Climate Stations (RCSs) was the likelihood of continued availability of reliable human observers, and a number of sites were excluded from the original RCS list because the long-term future of human presence at the site was doubtful. (The best example of this is Palmerville (Queensland), a “ghost town” where the observations were made for many years by the one remaining inhabitant until her death in 1999.) Now that automatic weather stations have become an integral part of the standard network, sites with no regular human presence have become practical locations for long-term climate records (although data losses due to instrument or communications faults, and the limited capacity for a rapid maintenance response, have been a problem at some of the more remote sites since automation).

4.2 Some issues with the distribution of sites and changes over time

Over time, there has been a general trend towards moving sites away from built-up areas. This has been driven by a number of factors, including the decreased role of post offices and the increased role of airports in the observing network, and the increasing recognition of the importance of having good network representation in non-urban environments, especially for elements such as temperature and wind speed, which are susceptible to strong influences from the site environment. The proportion of ACORN-SAT locations located in built-up environments[8] (regardless of the size of the town) has fallen from 67% in 1930 to 38% in 1970 and 20% in 2010, and is likely to fall further over the coming years, as eight of the remaining 22 built-up locations are currently operating in parallel with sites outside the town area that will eventually supersede them. This shift away from urban centres might be considered as a “negative” urban heat island effect, and accounts for some of the systematic differences between the ACORN-SAT data set and analyses derived from raw data, discussed in a companion report (see also section 9).

---

[8] For these purposes a built-up environment is defined as a site which has buildings (other than those explicitly associated with the site, e.g. an airport observing office) within 50 metres.

Numerous long-term temperature data sets in Australia have been at lighthouses or other facilities associated with marine navigation. Apart from occasional issues with wind shaking of instruments (see section 6), many of them are amongst the best long-term sites in Australia, as they are often in remote non-urban locations, many in national parks or other natural areas, and the only nearby structures are those associated with the lighthouse itself, that have typically been there for most or all of the period of observations, and are often heritage-listed, making future changes unlikely. However, by definition, they are in exposed coastal locations, and are therefore sampling conditions that are only representative of the coastal fringe rather than broader continent. There are virtually no century-long temperature records in inland locations that have been outside a built-up environment for their entire history, something which is reflective of the greater historical priority given in Australia to rainfall measurements over temperature observations.

There has been rapid population growth over the last 40 years in many coastal regions of Australia, particularly along the east coast. In some cases it has resulted in sites moving inland as commercial pressures reduce the number of available sites near the coast. However, whilst this tendency has affected the Bureau observing network as a whole, it has had no discernable impact on the ACORN-SAT data set over and above the more general tendency for sites to move to airport sites over time.

4.3 Observation time standards and changes over time

The current Bureau of Meteorology standard for daily maximum and minimum temperatures is for both to be measured for the 24 hours ending at 0900 local time[9], with the maximum temperature being attributed to the previous day. This standard has been in place since 1 January 1964 (except at some automatic weather stations – see below), although the data suggest that it was not until late 1964 or early 1965 that the standard was adopted across the bulk of the network, with a small number of sites not changing until the late 1960s or early 1970s.

The situation in the period from 1932 to 1963 was considerably more complex, both through the existence of multiple standards and the varying ways in which they were implemented. As a general rule, sites adopted the following procedures:

- Bureau-staffed sites, and a few others (mostly lighthouses and similar) that made observations around the clock, used a nominal midnight-midnight day. In practice, in most cases the thermometers were still reset at 09:00 or 15:00, then read at midnight, with the maximum or minimum read at midnight substituted for the value read earlier in the day if it surpassed it. In practice this meant that the 'midnight-midnight' minimum would actually be for the 33 hours ending at midnight, although the impact of this longer observation period is minimal as the number of occasions when the 33-hour minimum differs from the 24-hour value (i.e. where the temperature falls lower between 15:00 and 00:00 than it does in the succeeding 24 hours) is negligible.
- Sites that made observations at 09:00 and 15:00, as most co-operative sites did, reset their maximum thermometers at 09:00 and their minimum thermometers at 15:00. In effect this resulted in minimum temperatures being for the 24 hours ending at 15:00. Maximum temperatures were measured from 09:00 to 09:00, but observer instructions (e.g., Bureau of Meteorology 1925, 1954) were to revert to the maximum measured for the 6 hours from 09:00 to 15:00 if the 24-hour maximum measured at 09:00 the following day was 'close to' the current temperature.

---

[9] Times in this section refer to local clock time unless otherwise stated.

- Sites that only made observations at 09:00 measured maximum and minimum temperatures for the 24 hours ending at 09:00, as per current practice.

Prior to 1932, Bureau-staffed sites appear to have used an observation day ending at 09:00, with practices at co-operative sites as per the 1932-1963 period, except that a small number of sites reset instruments at 21:00 rather than 09:00 or 15:00.

Due to limitations in communications software, in their early years of operation, many automatic weather stations reported maximum temperatures for the 24 hours ending at 12:00 UTC (between 20:00 and 23:00 local time depending on time zone and time of year), and minimum temperatures for the 24 hours ending at 0000 UTC (08:00 to 11:00 local time). This software has been progressively phased out over time and no sites in the ACORN-SAT network are currently affected, although a few have been in the past.

As local clock time is used, except in 1971-72 when standard time was used, there has been an effective shift of one hour in observation time when daylight saving time has been in place. This shift applies for some or all of the October-March period since 1971-72 in the southeastern states (New South Wales, Victoria, Tasmania, South Australia and the Australian Capital Territory), in all states at times during the First and Second World Wars, and for brief trial periods (one or two years) in Queensland and Western Australia. Also, prior to April 1939, solar time, rather than standard time, was used for observations, leading to an effective shift of up to 45 minutes at some locations (notably those close to the western boundaries of their time zones, such as Darwin (Northern Territory) and Burketown and Boulia (Queensland)).

The potential effect of these changes on data at ACORN-SAT locations as a whole is discussed in section 8. A small number of documented instances of non-standard observation times (e.g. the use of an 08:30 observation time rather than 09:00 at Giles in Western Australia prior to 1978) were treated as metadata-defined site specific inhomogenities.

4.4 Units and data precision

The most substantial change during the last 100 years in measurement units or data precision has been the change from imperial to metric units on 1 September 1972 (The potential impact of instrument changes associated with this change is evaluated in section 8). Prior to this date, measurements were, in theory, taken to the nearest 0.1 degrees Fahrenheit, although in practice most observations were rounded to the nearest degree or half-degree. From September 1972 onwards, measurements were made to the nearest 0.1 degrees Celsius; rounding still occurred (Trewin, 2001b), but not to the same extent as occurred in the pre-1972 period.

More recently, in the early years of automatic weather stations, many such sites reported maximum and minimum temperatures rounded to whole degrees Celsius due to limitations in data transmission software. This affected some ACORN-SAT sites at times in the 1990s and early 2000s but no longer affects any of them. Such rounding will not produce any systematic bias in mean temperatures, since values ending in .5 are rounded to the nearest odd number, but does have the potential to affect the number of days above or below thresholds, an issue discussed further in section 7.

4.5 Other relevant observation practices

At many sites (particularly Bureau-staffed sites), both manual and automatic instruments were in place simultaneously in the same screen for a period of time. The automatic instruments became the ‘primary instruments’ as of 1 November 1996, taking precedence over the manual observations when both were available. In effect this means that the effective start date of automatic measurements of maximum and minimum temperature was 1 November 1996 at most Bureau-staffed sites.

5. METADATA AND THEIR USE IN ACORN-SAT

There are numerous issues, as discussed in section 4, that can have an effect on temperature measurements that is not associated with any changes in the background climate. These include issues that are specific to individual sites (such as site relocations, instrument changes or changes in local site conditions), or issues that affect large parts of the observing network (such as changes in observation times).

Documentation of these issues is critical in order to obtain a complete picture of non-climatic factors that might affect temperature observations at a particular location. In the context of the Australian observation network, there is a wide range of potential sources of metadata, although no individual source is comprehensive.

5.1 Types of metadata available for the ACORN-SAT sites

5.1.1 Site-specific metadata

The major sources of site-specific metadata available for the ACORN-SAT sites are a metadata database (SitesDB) and hard-copy station files maintained by the Bureau of Meteorology. Major items of relevance, that form part of the source material for both SitesDB and the hard-copy station files, for the ACORN-SAT project, include:

- Details of site relocations.
- Site inspection reports, including site photographs (or in some earlier periods, sketches) and diagrams. In recent decades these have been approximately annual at ACORN-SAT sites, although there are a few notable exceptions, particularly at locations that are difficult to access. In the historical data they are much less frequent, with inspection reports often a decade or more apart prior to 1960.
- Changes in instruments. These are perhaps the best-documented changes historically, most likely because they involved the expenditure of public funds.
- Instrument calibrations and results of tolerance checks.

SitesDB has been in operation since 1997 and has, since then, been the major store for site-specific metadata. It can only be assumed to be reasonably complete since 1997, although there have been some efforts made to backfill it with the most important historical data (e.g., major site relocations) in some states, especially New South Wales.

For pre-1997 metadata the most comprehensive source is the set of hard-copy station files. Two sets of files are maintained: one by the Bureau's Head Office, and one by the Regional[10] Office in which the site is located. There is considerable, but not total, overlap between the two sets of files (with the Regional Office files being the more complete in most cases), with some documents on one file but not the other.

Earlier sources of metadata include a series of 'Instrument Books', which were hard-copy registers for entering details of any instruments that were issued to sites. Some of them, especially for Queensland, are held in the National Meteorological Library in Melbourne. They mostly cover the period between the late 1880s and about 1920, and are hence most useful, in the context of ACORN-SAT, for determining the date of Stevenson screen installation at each site.

5.1.2 Other metadata

Obtaining metadata that are not specific to individual sites is often more challenging than obtaining site-specific metadata, as there has not always been a specific repository for such information.

Observation manuals (e.g., Bureau of Meteorology 1925, 1954, 1984) are a detailed description of observation practices at a specific point in time, but are less useful in determining the exact date of any relevant changes. Although not directly relevant to the post-1910 period covered by ACORN-SAT, the reports of various late 19th and early 20th century Intercolonial Meteorologists' Conferences (or equivalent), and annual reports of the various government meteorologists, are also a very useful source of information on the pre-1910 evolution of the observation network. More recently, changes in policy have been documented in a series of Observations Instructions (or similar) issued by the Bureau's Observations and Engineering Branch.

However, some issues (e.g., the coding limitations discussed in section 4.3 that resulted in the use of 00:00/12:00 UTC observations for a time at some automatic weather stations) are not documented in any formal publications and are known to the ACORN-SAT project only through the corporate knowledge of individuals involved with the project. (It is highly likely that other changes have occurred historically but without any documentary evidence.) In other cases, the exact date of changes such as the change to an 09:00-09:00 observation day, the use of daylight savings time, or the introduction of metric units for temperature, has been inferred from the design of the forms or field books used by observers in particular years, or obtained as incidental information in other publications (e.g., the Monthly Weather Review issued for each state)[11]. Implementation of network-wide policies has also not always been uniform, a particular case in point being the 1964 shift to a network-wide 09:00-09:00 observation day, as discussed in more detail in section 4.3.

Other information from third parties can also be considered as metadata. This includes maps (particularly topographic maps), and population data, from the Australian Bureau of Statistics, for urban centres near observing sites.

---

[10] The Bureau of Meteorology has seven Regional Offices, in the six states and the Northern Territory. The Australian Capital Territory is covered by the New South Wales Regional Office.

[11] Perhaps the most fortuitous discovery of relevant metadata was that of the observation procedures used during the short-lived implementation of daylight saving time in 1917, which was found by one of the authors in a Brisbane newspaper from late 1916, in the course of unrelated research into severe flooding in Queensland in December 1916.

5.2 Specific metadata used in the ACORN-SAT project

A search was undertaken of the sources discussed above for sites in the ACORN-SAT data set. This built on the work of Torok and Nicholls (1996) and Della-Marta et al. (2004), who carried out a comprehensive search of metadata for the sites in their data set. The information they obtained was available through documentation archived on the Bureau's internal website, with the pre-1996 information also published in Torok (1996). This information was supplemented with the following:

- A search of all available information for the ACORN-SAT sites on SitesDB, mainly to update information to the present day, as well as to obtain information on aspects such as tolerance checks of automatic weather station probes that was not explicitly considered in earlier work.
- A search of hard-copy station files, concentrating on those ACORN-SAT sites that were not included in the Torok and Nicholls or Della-Marta et al. data sets. As part of this effort, visits were made during 2009 to all of the Bureau of Meteorology's Regional Offices (and in some cases to regional offices of the National Archives of Australia), to view their station files and obtain relevant information.
- A search for documentation potentially relevant to network-wide changes (much of which was already known to the authors through earlier projects).
- The author having direct knowledge of the sites (having visited 89 of the 112 ACORN-SAT locations), and the knowledge of other relevant current and former Bureau of Meteorology personnel.

5.3 Limitations of available metadata

The ideal situation is that metadata exist covering all changes that have the potential to affect temperature observations at a site. In practice, metadata usually fall short of this level, or are otherwise difficult to use, for a number of reasons, including:

- Metadata are often incomplete or missing, and can be open to interpretation. This is especially true of metadata that are relevant to a large part of the network at the same time, for reasons discussed in the previous section. It is also an increasing problem as one goes further back in time.
- Metadata have been generally collected to support the data's purpose at that point in time. The purpose of a site may change over time, and as a result, the level of detail or type of metadata collected could also change.
- Extracting relevant information in usable form can be challenging; as noted earlier, much important information is only available in hard-copy documents that require exhaustive searching, with information relevant to a climate record often forming only a small part of a large volume of documentation on other matters relating to a site.
- Metadata can be misleading without additional knowledge – for example, a change in site co-ordinates could arise from a change in site location, or from a resurvey with no actual site move. A particular practice which is relevant here is that up until the 1990s, at some airport sites, the co-ordinates listed for the site were the official airport co-ordinates, normally on the runway, rather than the actual instrument locations; hence, for example, the two sites at Forrest (11004 and 11052) have co-ordinates which are 800 metres apart but are in fact in the same place.
- Metadata often take the form of snapshots at a particular point of time (e.g., quasi-annual inspection reports). This can indicate that a change has occurred at some stage between two specific points in time but not the exact date of the change (for example, if

one set of site photographs shows a new building near the observing site and the previous set does not, this indicates the building was built at some time between the two sets of photographs). This is a particular problem for changes in a site's surroundings – if a site moves the exact date will usually (although not always) be documented, but changes in the surroundings are rarely well-documented, especially outside a 30-metre radius (30 metres is the standard extent of site diagrams).

- Site-specific metadata typically focus on the observing site and its immediate vicinity and do not include changes more distant from the site that may be relevant to temperature observations there. A possible example of this occurs at Canberra, where an inhomogeneity of approximately +0.5°C has been identified through statistical analysis (as described in section 7) for minimum temperature in 1964. There is no documented site-specific change at the site itself at this time, but it coincides with the creation of a substantial artificial lake (Lake Burley Griffin), about 4 km west of the site at its nearest point. It is also possible that changes in observation time may be an additional contributing factor factor.
- Metadata may miss important information or be too imprecise to capture information of climatic relevance. One example occurs at Walgett, where a site move in 1997 was accompanied by a sharp increase in the frequency of extreme low minimum temperatures. A site visit revealed that the observing site is surrounded on three sides by the town's levee bank, thus creating a hollow favourable for trapping cold air under calm, clear conditions. This hollow would not be resolved by standard topographic mapping with a 10- or 20-metre contour interval.

Hence, whilst metadata are an essential part of developing a homogeneous temperature data set, they cannot be used in isolation. The implications of this for the methods used in the development of the ACORN-SAT data set will be discussed further in section 7.

5.4 Accessibility of metadata

A common characteristic of metadata is that much of it is not readily accessible. Even when used internally within the Bureau of Meteorology, significant quantities of metadata are in forms that are difficult to locate, or difficult to obtain and analyse (e.g., hard-copy documents) once they are located.

This problem is compounded for external parties with an interest in metadata, as much relevant information is contained in internal files, or in internal publications, with very little having been published externally. Whilst the station files are public documents and none of the material used in ACORN-SAT is classified information, the logistics of handling hard-copy material make it very hard to obtain for anyone without ready access to the offices of the Bureau of Meteorology and the National Archives of Australia.

These issues are not unique to Australia. Indeed, Australia is in a stronger position than some other countries – for example, Bureau of Meteorology (1997) is one of the most comprehensive set of observations guidelines for a national observation network to be available online, and it is also a substantial advantage that the agency responsible for the nation's official climate record is also the agency responsible for the observations network used to collect the data for that record, something that is not the case in countries such as the United States and New Zealand.

A separate ACORN-SAT station catalogue has been produced which contains detailed metadata for all of the ACORN-SAT locations. It is not, however, possible to release all metadata (e.g., through a bulk scanning of all material in station files), as this includes material covered by privacy laws (e.g., pay records for observers). It is also illegal to publish photographs of military installations without approval, which affects those ACORN-SAT sites that are on Department of Defence land.

6. DATA QUALITY CONTROL WITHIN THE ACORN-SAT DATA SET

Data quality control is a very important part of climate data management, and ensuring that data are fit for purpose (WMO, 2011). Errors can occur in meteorological observations for a wide variety of reasons, the most common being instrument faults, observer errors, errors in data transmission and clerical errors in data processing. The error rate in temperature observations is low – experience with operational quality control procedures at the Bureau of Meteorology in recent years suggests that it is in the order of a few tenths of one per cent – but such a rate still equates to the potential for several tens of thousands of errors in a data set of the size of ACORN-SAT. A somewhat arbitrary distinction is drawn here between short-term issues that affect observations over a finite period (most commonly a single observation, but sometimes persisting for a period of a number of days or weeks), and longer-term influences on a climate record (inhomogeneities) that are discussed further in section 7.

A further issue with data quality control is that there are two possible types of error – that of accepting an observation that is in error, and that of incorrectly rejecting an observation that is correct. The second type of error is particularly important in climate extremes as some of the most extreme events will trigger some checks typically used for detection of errors in data quality control systems (e.g. a check against the highest or lowest values on record), and incorrectly rejecting valid extremes will create a bias in the analysis of extremes.

The Bureau of Meteorology currently uses a computerised Quality Monitoring System (QMS) for climate data. This system subjects data to a series of checks, including checks for internal consistency, spatial consistency and limit/range checks, with data flagged by those checks then being referred to quality-control staff for further investigation, before being either accepted or flagged as suspect. The QMS has been used operationally in its present form since 2008.

The underlying philosophy in the ACORN-SAT project has been to subject data throughout the historic record, to the extent possible, to a level of quality control similar to that currently applied operationally to new data that enter the Bureau's climate data bank. The level of consistent quality control over most of the pre-2008 period, and particularly prior to the introduction of the Bureau of Meteorology's database (ADAM) in 1994, falls well short of current standards, mainly because, before the introduction of modern computer analysis tools, it was very difficult or impossible to perform data-intensive checks such as spatial intercomparisons. Other developments that have increased the capacity to perform quality control include the digitisation of additional data, which increases the amount of comparison data available to verify a record.

A further reason to carry out quality control in the ACORN-SAT data set independent of what has taken place in the operational climate database is to apply a consistent set of rules over time. Some data, for example, have been flagged as suspect in the operational database that would not be considered suspect under current data processing guidelines (e.g., an internal inconsistency that is within tolerance), or would be considered valid under the observation practices that were in use at the time (e.g., where the minimum temperature is higher than the temperature at 0900 the previous day, which is an internal inconsistency under current observation time policies but not under pre-1964 policies).

Experience with the analysis of data has highlighted that most errors in temperature variables tend to be quite random, so that their impact in aggregate when averaging across many sites is very small and can be effectively ignored. However, when analysing the records at individual locations, and when looking at extremes, it is important to remove or correct errors. Errors can also contaminate spatial analyses, particularly at shorter timescales (daily to monthly).

6.1 Quality control checks used for the ACORN-SAT data set

Each of the ACORN-SAT location time series was subjected to the sequence of quality control checks listed below. Any data flagged under these procedures were then subjected to further investigation, as described in the next section. The checks were carried out in the order listed, principally to filter as many of the most obvious errors as possible prior to carrying out the checks 7 through 10, that were the most labour-intensive in following up flagged observations. The theoretical basis of some of the checks (especially 1 to 3, 7, 8 and 10) are similar to those used by QMS. The thresholds used in checks 2, 3, 7 and 8 are based on unpublished results for ‘expected’ variations in good-quality data found during initial development work for QMS. These checks are being used in preference to processing through QMS, as the checks for ACORN-SAT were carried out on all the necessary temperature data by one individual. The combination of these two factors requires specifically designed tools, that allowed the user to make well-informed decisions by using their detailed knowledge of the observing network and local influences contributing to temperature at particular locations, as discussed in section 6.2. This is in contrast with QMS, which has been primarily designed for data managers (and not necessarily climate scientists) to make use of more labour-intensive interactive tools that cover additional observation quantities such as air pressure and wind speed, and would have been too time-consuming to apply to the volumes of data involved in ACORN-SAT.

Through this section, the following naming conventions apply unless otherwise stated:

Tx_d is the maximum temperature for the 24 hour period *from* 09:00 local time on day d, and Tx_{d-1} is the maximum temperature on day ($d-1$).
Tn_d is the minimum temperature for the 24 hour period *to* 09:00 local time on day d. and Tn_{d-1} is the minimum temperature on day ($d-1$).
$T_{hhmi.d}$ is the temperature at hhmi (*hh:hour*, *mi:minute*) local time on day d (for example, $T_{1500,d}$ is the temperature at 15:00 on day d).

1. Internal consistency of daily maximum and minimum temperature

Since the temperature recorded at the time of observation (09:00 under current practice) is an upper bound for minimum temperature on both the day of observation and the following day (i.e. $Tn_d \leq T_{0900,d}$ *and* $Tn_{d+1} \leq T_{0900,d}$), and a lower bound for maximum temperature on both the day of observation and the preceding day (i.e. $Tx_d \geq T_{0900,d}$ *and* $Tx_{d-1} \geq T_{0900,d}$), daily maximum and minimum temperatures must satisfy the relationships:

$Tx_d \geq Tn_d$
$Tx_d \geq Tn_{d+1}$

If one or both of these relationships was violated, both maximum and minimum temperatures were flagged as suspect unless there was strong evidence that any error was confined to one of the two observations.

2. Rapid change and spike check for METAR data

METAR data were checked for excessively rapid changes. The main purpose of this check, that in general only covers the period after automatic weather station introduction and thus incorporates data mostly from the mid-1990s onwards, is to detect instrument faults, since, for automatic stations, the maximum and minimum temperatures are measured using the same instruments as the fixed-hour temperatures (unlike the situation for manual thermometers where they are measured using different thermometers), and hence a fault that affects the fixed-hour temperatures also has the potential to affect the maximum and/or minimum.

The criteria for flagging data were:

Spikes: $| T_{t2} - T_{t1} |$ and $| T_{t2} - T_{t3} | \geq 4$°C, where $T_{t2} - T_{t1}$ and $T_{t2} - T_{t3}$ both have the same sign, and the time pairs ($t1$, $t2$) and ($t2$, $t3$) were each separated by not more than 70 minutes. (Note that this definition incorporates both positive and negative spikes).
Rapid change: $| T_{t2} - T_{t1} | \geq 10$°C, for a time pair ($t1$, $t2$) separated by not more than 70 minutes.

Where METAR data were flagged by one or both of these checks, the maximum and minimum temperatures on the affected day(s) were flagged for further investigation. Figure 7 shows an example of this (Nuriootpa, South Australia, 12 November 2008). The METAR temperature of 15.9°C at 06:30 on 12 November was flagged by this check with a spike value of 4.1°C. The minimum temperature of 15.3°C was then found to be several degrees too low in comparison with both the terrestrial minimum (19.0°C) and surrounding sites.

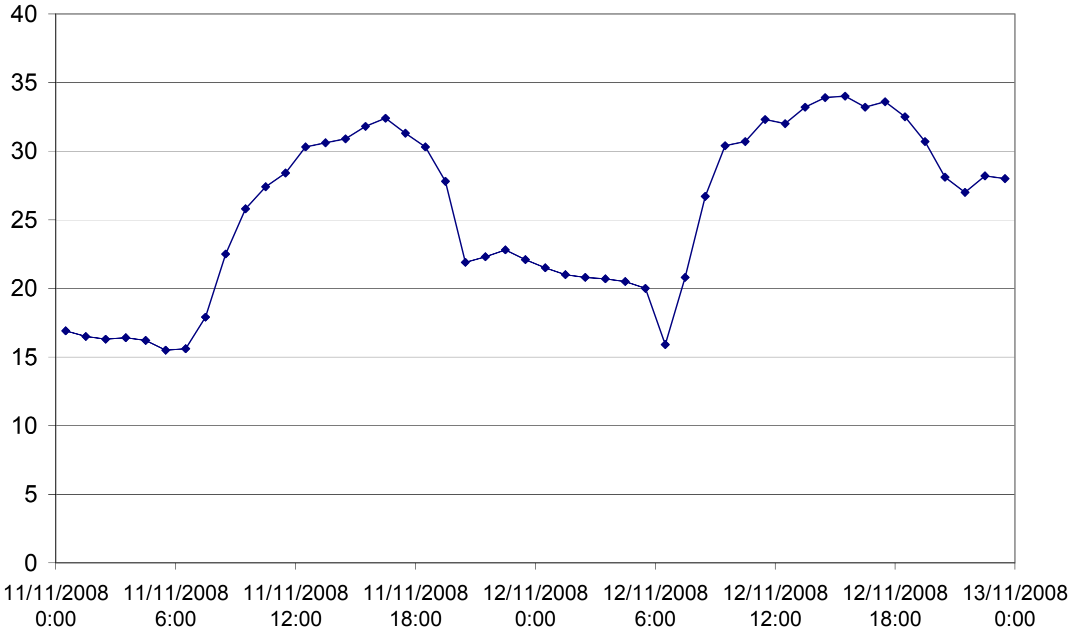


Fig. 7. Example of suspect observation (°C) identified by spike check – Nuriootpa, 12 November 2008

3. Internal consistency of METAR and maximum/minimum temperature data

This check flagged data violating either of the following:

- Maximum temperature 4°C or more above the highest METAR temperature of the day, providing that there was no point during the day when there were more than 70 minutes between METAR temperatures.

- Maximum temperature 1°C or more below the highest METAR temperature of the day. (The tolerance on this test was used because many METAR temperatures, particularly manually observed ones, are only archived to the nearest whole degree.)

Equivalent criteria were used for daily minimum temperatures.

4. Comparison of newly digitised data with previously digitised monthly means

This check was carried out on only newly-digitised data not yet subject to any quality control (see section 2.5). Monthly means calculated using these newly-digitised data were compared with the monthly means already recorded in ADAM and discrepancies investigated to find any obvious errors. In most cases such errors involved shifting of a decimal point (see later section), although one case was also found of observations being entered against the wrong month in the data bank.

5. Review of data already flagged as suspect

All data from ACORN-SAT sites that are currently flagged as suspect or wrong in the ADAM database were subject to similar follow-up checks to those described elsewhere in this section to determine whether the date were actually suspect using consistent criteria. Some of the reasons why data have been flagged in ADAM in a manner inconsistent with current practice include:

- Internal inconsistency between maximum/minimum and fixed-hour temperatures where the inconsistency is ≤ 0.5°C, which is the current tolerance used for manual sites.
- A large number of sites had all maximum and minimum data flagged as suspect for a period, mostly in 1998-99, because of allegedly excessive rounding or 'poor exposure'. Neither would be considered grounds on their own for flagging an observation as suspect now (although they may be pointers to poor observation practices or inhomogeneities which are considered separately), and in any case some of the cases where 'rounding' was indicated as the reason for flagging showed little or no evidence of rounding.
- At some sites, maximum and minimum temperatures are flagged as suspect on all days with very small diurnal temperature ranges (typically less than 1°C). In fact, this does not necessarily indicate an error (for example, if the temperature is falling continuously through the observation time, as might occur if a cool change is passing through the site at or near 0900, the diurnal range can be zero).
- In some cases data were flagged for internal inconsistencies based on current observation time practices that were not in use at the time (for example, a temperature at 09:00 the previous day that was lower than the minimum, which is inconsistent for an 09:00-09:00 observation day but not for an 00:00-00:00 day).
- Some minimum temperatures were flagged as suspect on the grounds of being slightly lower than the terrestrial minimum. Such a difference (although not normally more than 1°C) can legitimately occur, particularly in windy conditions or during the tropical wet season, where daily minimum temperatures sometimes occur during rain/thunderstorms during the day rather than through nocturnal cooling. However, terrestrial minimum temperatures, where available, provided useful supporting data for investigating screen minimum temperatures that were identified as potentially suspect by other checks, as discussed in the next section.

There were also cases where the flagging arose from an issue that did not affect the maximum/minimum temperatures (e.g., cases of internal inconsistency where the fixed-hour observations, not the maximum or minimum, were wrong), or where the observations could be repaired through, for example, shifting maximum temperatures by one day, as discussed in later sections.

6. Spatial check of monthly means

Mean maximum and minimum temperatures for each month at each ACORN-SAT location were checked for spatial consistency, using a cross-validation check. This was carried out using monthly temperature anomalies, calculated using a 1961-90 base period if the location had 25 or more years of data in the 1961-90 period, the current WMO standard for calculation of climatological standard normals, or all post-1910 years of data otherwise.

In the cross-validation check, the interpolated estimated temperature anomaly at the site, T_{int}, was calculated, using all available data other than that from the candidate location, as:

$T_{int} = \sum w_s T_s$, where T_s is the monthly temperature anomaly at location s and w_s is a weighting function:

$w_s = 1/\exp((d/200)^2)$, where d is the distance in kilometres between location s and the candidate location.

If T_{int} differed from the monthly anomaly value at the candidate location by 3°C or more, the month's data was flagged for further investigation.

Figure 8 shows an example of such a case, in which the minimum temperatures for Tibooburra transpired to be 8-10°C too low, with respect to other sites in the region, for several months in late 1931 and early 1932, most likely because of an instrument fault.

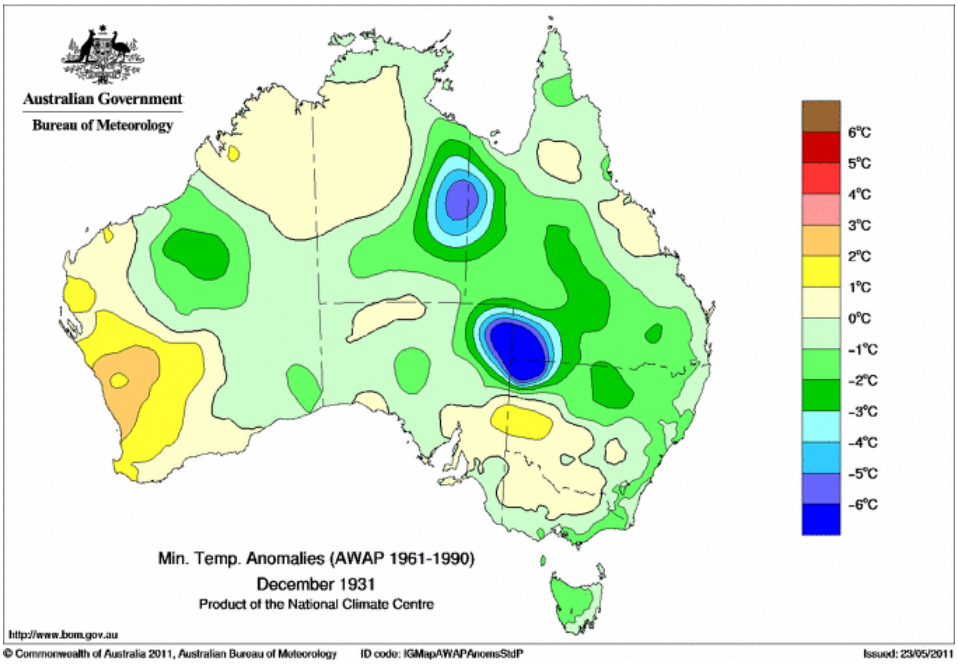


Fig. 8. Monthly minimum temperature anomalies for December 1931, indicating suspect data at Tibooburra (near NSW/Queensland/SA border).

The spatial check was also used where metadata indicated that a site had failed an instrument tolerance check by 1°C or more. In such cases, the spatial check was used to identify the month in which the instrument first went out of tolerance by 1°C or more, and all data measured by that instrument (which, in the case of an automated temperature probe, means all maximum and minimum temperature data) was flagged as suspect from that date until the date of the recalibration.

The major purpose of this check was to identify longer-term data quality issues (most commonly instrument faults) that might not be large enough to trigger checks for individual days but are sufficient to render the data suspect if sustained over a longer period. It was also an early filter that enabled some of the worst observations to be flagged before they were incorporated in more time-consuming checks.

7. Consistency check between maximum/minimum and synoptic temperature observations

This check was for consistency between maximum/minimum and fixed-hour temperature observations. Whilst only a small part of the ACORN-SAT data set has fixed-hour temperatures at hourly or finer resolution as used in checks 2 and 3, almost all of it has at least some synoptic fixed-hour data, and hence this check could be applied to almost all of the data set (although its effectiveness depends on the number of fixed-hour observations available per day). Synoptic fixed-hour observations take the form of observations at some or all of the times 00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00 and 21:00 local time, although prior to 1987, only Bureau-staffed sites (and a few others) have fixed-hour data at times other than 09:00 and 15:00.

The fixed-hour observations used for flagging were those from 09:00 one day to 09:00 the next, consistent with standard post-1964 observation practice. However, for pre-1964 data (and for automatic weather station data using a day ending at 00:00 or 12:00 UTC), the flagged observations were only investigated further if initial screening indicated they were inconsistent with the observation day in use at the time. (The use of a 1 January 1964 cut-off resulted in many 1964 observations being flagged at sites where the changeover does not appear to have been implemented until later in the year).

An observation was flagged if it failed any of the following tests:

- Highest fixed-hour temperature more than 0.5°C above the maximum temperature.
- Lowest fixed-hour temperature more than 0.5°C below the minimum temperature.
- Maximum temperature more than an amount *L* above the highest fixed-hour temperature. *L* was set equal to 5°C if observations were available at 12:00, 15:00 and 1800; 6°C if observations were available at 15:00 but not at one or both of 12:00 or 1800; 8°C if observations were available at 12:00 and 18:00 but not 15:00. In any other case (e.g. observations at 09:00 only, or 09:00 and 21:00) this check was not carried out.
- Minimum temperature more than an amount *L* below the lowest fixed-hour temperature. *L* was set equal to 5°C if observations were available at both 03:00 and 06:00, and 6°C if observations were available at one but not the other. If there were no observations were available at either 03:00 or 06:00 this check was not carried out.

An illustration of this check is given in Fig. 9, for Port Hedland on 28 November 1990. The minimum temperature was 17.5°C and the lowest three-hourly temperature was 28.2°C (at 0600), a difference of 10.7°C. Follow-up investigations also revealed that the terrestrial minimum was 26.0°C, reinforcing the suspect status of the screen minimum temperature on that day.

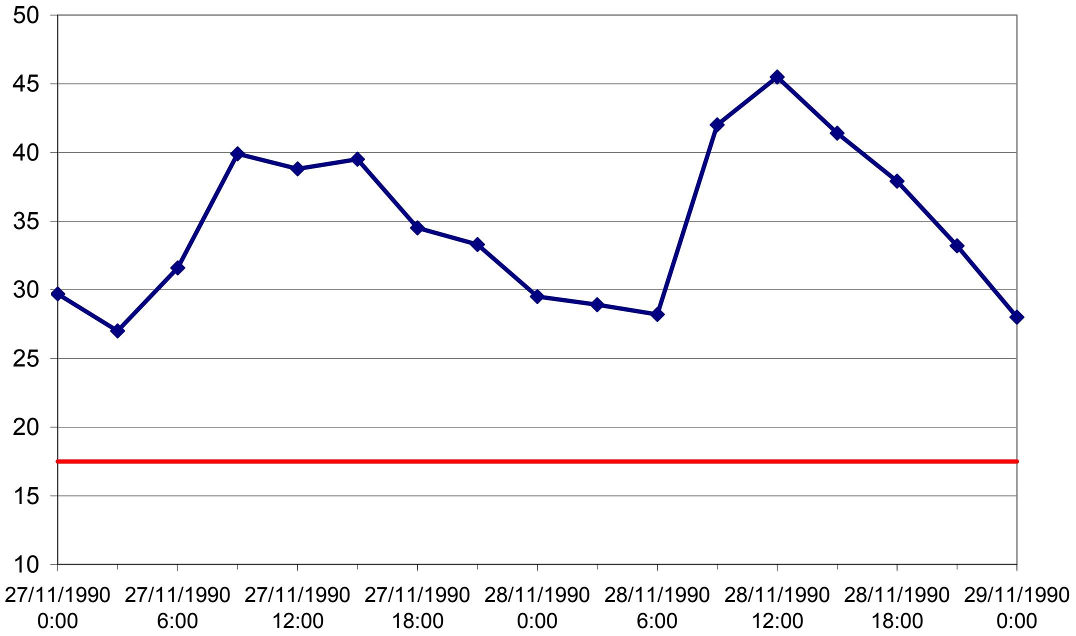


Fig. 9. Synoptic temperature observations (°C) at Port Hedland, 27-28 November 1990 (blue) and minimum temperature for 28 November (red).

This check is most effective in identifying maximum temperatures that are too low, and minimum temperatures that are too high. It is less effective in identifying maximum temperatures that are too high or minimum temperatures that are too low, except at sites with observations from most of the possible three-hourly intervals. It is also relatively ineffective at sites that have observations at 09:00 only. At such sites the spatial check (stage 8) is more effective.

8. Spatial intercomparison of daily data – first iteration

This check compared observations with its nearest neighbours. This is the most powerful check, and is capable of identifying errors, particularly at sites with limited supporting data such as fixed-hour temperatures, that are not detectable by the earlier checks.

The check was carried out by comparing a site's daily temperature anomaly (calculated as in check 6) with a weighted mean of the anomalies of its neighbours. The anomalies were calculated using monthly means at each site as per stage 5. In general the 10 nearest neighbours with some overlapping records with the candidate site were used, regardless of period of record. In some cases, the candidate site had periods when none of its 10 nearest neighbours had data; in these cases the number of neighbours was extended to maximise the number of years for which there was at least one neighbour with available data, and in a few cases nearby sites that had clearly different climates (e.g. an offshore island where the candidate location was inland) were excluded.

The weighted mean anomaly of neighbouring temperature anomalies, T_{int}, was calculated as:

$T_{int} = \sum w_s T_s$, where T_s is the daily temperature anomaly at location s and w_s is a weighting function:

$w_s = c/\exp((d/100)^2)$, where d is the distance in kilometres between location s and the candidate location, and c is the correlation between monthly temperature anomalies at location s and those at the candidate location, or 0, whichever is the greater (the effect of this is that any locations where $c < 0$ were excluded from the weighted mean).

The weighting function used here is weighted more heavily towards nearby sites than that in check 6, reflecting the generally shorter decorrelation length scales for daily temperatures compared with monthly temperatures.

Data were flagged if the temperature anomaly at the candidate site, T, varied from T_{int} by more than a specified value L. L was set at either 4°C, 5°C or 6°C, based on a subjective assessment of network density and local climatological gradients.

Figure 10 shows an example of data flagged by this check, a minimum temperature of 20.6°C at Giles (25°S 128°E) on 11 January 1988. The lowest three-hourly temperatures at the site were 26.2°C at 09:00 the previous day, and 27.5°C at 06:00, while other sites in the broader region mostly exceeded 27°C. (As there are no sites to the north or south of Giles within 500 km, the value affects the analysis over a large area). Such differences, if real, would almost certainly be associated with a thunderstorm at Giles, but no significant rain was recorded there, suggesting that the value on that day was suspect.

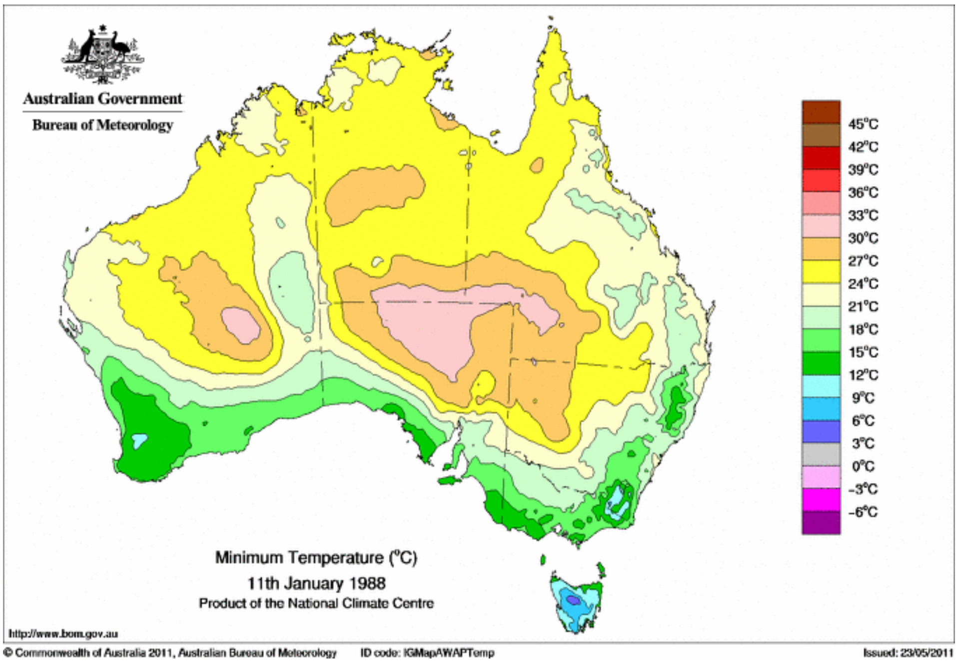


Fig. 10. Australian minimum temperatures, 11 January 1988.

A complication in this check is that in some cases pre-1964, sites in close proximity sometimes used different observation times, occasionally leading to inter-site differences large enough to trigger the spatial check. There are several examples in the ACORN-SAT data set, between 1939 and 1953, of parallel observations between an airport site using a 00:00-00:00 day, and a town centre site using an 09:00-09:00 day (e.g., Alice Springs, Kalgoorlie, Mount Gambier, Ceduna). This is illustrated by Fig. 11 – the maximum temperature of 34.0°C at Ceduna Airport on 12 February occurred before 0900, and had Ceduna Airport used an 09:00-09:00 day as the

Post office did, the three-hourly data suggest its maximum would have been close to the Post Office's 25.0°C, rather than 9°C above it as the daily data suggest.

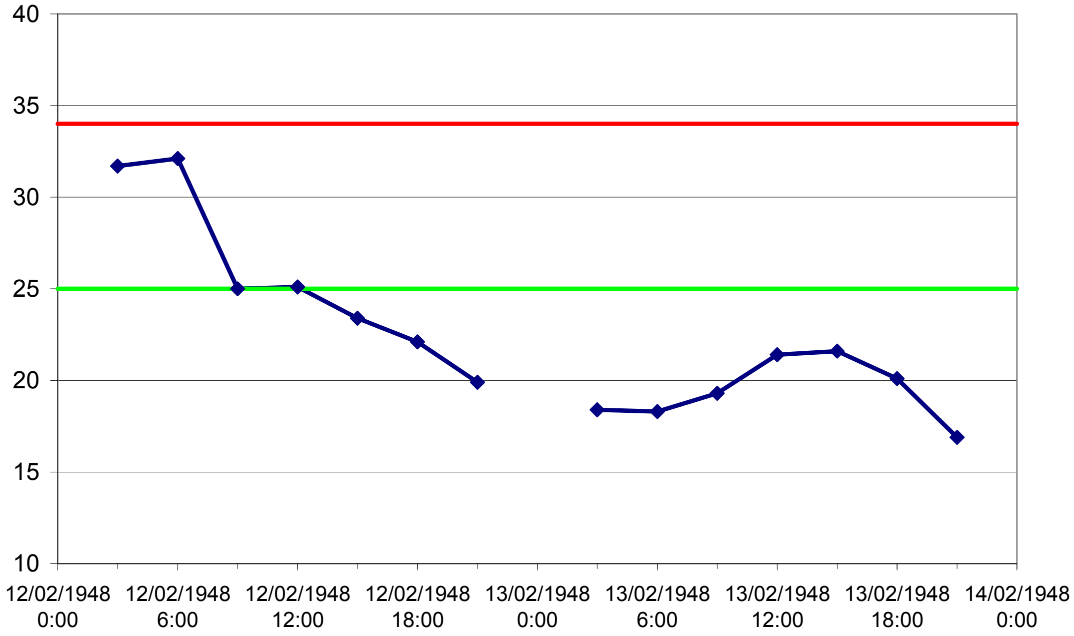


Fig. 11. Synoptic temperatures (°C) at Ceduna Airport 12-13 February 1948 (blue), with maximum temperature for 12 February at Ceduna Airport (red line) and Ceduna Post Office (green line).

9. Spatial intercomparison of daily data – second iteration

At sites where there was only one neighbouring site, or none at all, available for the checks used in stage 8 for some part of the period of record, it was necessary to carry out a second iteration but looking further afield for potential neighbours. The checks in stage 8 were repeated for that period, but on this occasion using the nearest neighbours (up to 10, although more typically 3 to 6) that had data available for some or all of the period and had daily temperatures correlated at a value of at least 0.6 with the candidate site (with correlation values calculated as described in section 6). In this stage of the checking, reflecting the wider areas over which neighbours were chosen, *L* was set to 6°C in all cases.

At three locations (Darwin for most of the period 1910-58, Thursday Island/Horn Island for 1952-56, and Sydney for 1910-39), there were no suitably correlated neighbours and this check could not be applied.

10. Range check

As a final check, the highest and lowest observations at each location were flagged for further verification. The observations so flagged were the ten highest and ten lowest values for daily maximum and minimum temperatures for each of the 12 months (480 observations per location in all).

6.2 Follow-up investigations of flagged data

All observations flagged by the checks described in the previous section were subject to follow-up investigations in order to make a final decision as to whether to accept or reject the value. This was the most time-consuming part of the project as several hundred thousand observations were involved (out of a total of about 7 million observations in the ACORN-SAT data set).

A distinction was drawn in follow-up investigations between observations that had been flagged because of a violation of "hard" limits (e.g., maximum temperature less than temperature at 15:00), for which at least one observation must be wrong, and those that were flagged because of a violation of "soft"[12] limits (e.g., excessive variation from neighbours). In the former case a "guilty until proven innocent" approach was taken, with the maximum/minimum temperature considered suspect unless evidence pointed to the inconsistency arising from another cause (e.g. an incorrect 15:00 temperature). In the latter case an "innocent until proven guilty" approach was taken with marginal observations generally included.

In follow-up investigations all available data were used. They included:

- Fixed-hour temperatures (particularly where they had not triggered the initial check).
- Temperatures from neighbouring locations (particularly where they had not triggered the initial check).
- Inspection of the whole-network analyses produced as part of the Australian Water Availability Project (AWAP; Jones et al., 2009).
- Terrestrial minimum temperatures, where available.
- Dewpoint temperatures, where available (principally for investigation of suspiciously low minimum temperatures). This variable was, however, used with some caution as dewpoint observations are, in general, less reliable than dry-bulb temperature observations.
- Wind, cloud and rainfall observations.

Clearly, the availability (or lack thereof) of such supporting data affects the ability to detect errors, and it is highly likely that some errors will remain undetectable, particularly in the most data-sparse areas.

Also incorporated in the decision-making process was knowledge of the most common errors (see next section) – with data that showed indications of being affected by one of these errors being considered more likely to be suspect - and local climatic influences at each site, particularly those in unusual environments.

An illustration of how this was applied is the case of potential high maximum temperatures at the highly exposed coastal site of Wilsons Promontory. High temperatures at Wilsons Promontory normally require moderate or strong winds from an overland trajectory, that can come from only a narrow range of directions; as such conditions often only exist for a short period of time, rapid temperature rises and falls are quite common at the site and it is therefore not uncommon for the daily maximum temperature to be 10°C or more above the highest three-hourly observation. However, it would be extremely unusual for the maximum temperature to be significantly higher than maximum temperatures at less exposed sites in the region (such as

---

[12] Here "hard" limits refer to limits whose breaching indicates a definite error, while "soft" limits refer to limits whose breaching indicates something unusual but not necessarily an error.

Wonthaggi). For example, Wilsons Promontory recorded a maximum of 22.9°C on 2 January 1971, with fixed-hour temperatures of 12.8°C at 09:00 and 15.6°C at 15:00. This difference would not be grounds on its own for considering the maximum suspect had Wonthaggi and similar sites also recorded maximum temperatures near or above 22.9°C, but in fact Wonthaggi's maximum was 17.5°C and almost all of the southern half of Victoria was below 21°C (Fig. 12), indicating that the Wilsons Promontory value was indeed suspect.

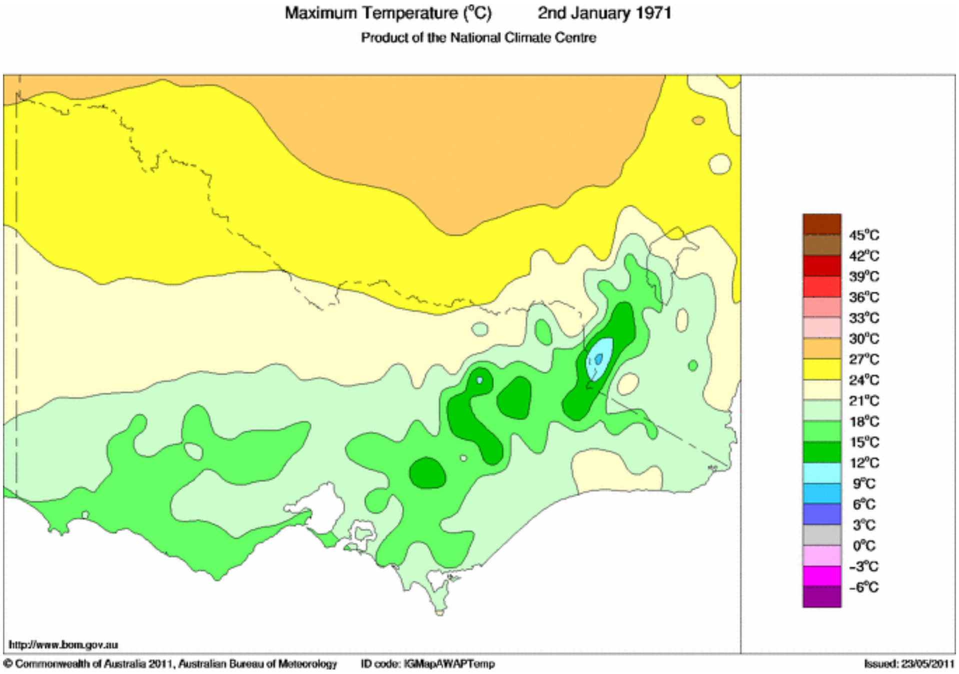


Fig. 12. Maximum temperatures in Victoria, 2 January 1971.

Another example where local climatic knowledge was used was for low minimum temperatures during the tropical wet season. On occasions, minimum temperatures at such locations occur during thunderstorms rather than as a result of nocturnal cooling. As thunderstorms are short-lived and localised, such a low minimum temperature might not necessarily be reflected in either three-hourly temperatures or temperatures at neighbouring sites (bearing in mind that networks are fairly sparse in most tropical areas). Hence, minimum temperatures at tropical sites that had triggered a flag for potentially being too low, but were 20°C or higher and on a day when rain was observed, were generally not considered suspect unless there was further strong evidence (e.g., a terrestrial minimum which was much higher than the screen minimum).

In general, data that were classified as suspect after review were flagged and excluded from further analysis. In general values were only amended if firm evidence existed of the correct value. Forms this could take include:

- Maximum/minimum temperatures derived from one-minute data, in cases where the originally suspect value was triggered by a fault well clear of the likely time of maximum/minimum temperature (e.g., a low minimum temperature caused by a short-lived fault during the afternoon). Half-hourly or hourly data were not considered sufficient for this purpose.
- Documentary evidence of the correct value (e.g., original observation forms) in the case of processing errors. These forms were not checked exhaustively – something that

would be a very time-consuming process – but were checked opportunistically at some locations, especially capital cities.

- Observations where it is highly likely that the value was shifted by a decimal point. Almost all of these involve newly digitised data that were entered as degrees Fahrenheit (as per the original forms), and are obvious on initial inspection (e.g. a value of 70°F that is entered as 7.0°F will show as −13.9°C, lower than the lowest valid observation at any ACORN-SAT location).
- Observations where it is highly probable that the temperatures are shifted by one day.

As a result of these follow-up investigations, 18,400 individual observations and 515 blocks of observations of three or more days were flagged as suspect and excluded from further analysis or amended, while 50 blocks of observations were shifted in time (mostly maximum temperatures brought forward by one day, but also including a few cases of months that had been swapped). The bulk of these issues were between 1957 and the early 1970s. Relatively few errors were identified after the early 1970s (and particularly after the mid-1990s), presumably because of improved quality-control procedures over time, whilst most pre-1957 data were only digitised in the last 10 years and therefore also underwent relatively effective quality control.

It is intended to carry these changes and flaggings across to the main ADAM database but this has not yet occurred at the time of writing.

6.3 Common errors and data quality problems

Some common errors and data quality problems, and characteristic signatures of some of them in the data, include:

- Instrument faults, most commonly with automated probes (as a result of power surges or similar). For automated probes these tend to manifest themselves as a sharp short-term spike in the data.
- Misreading of instruments by a set amount, often 5 or 10 degrees (Celsius or Fahrenheit depending on the era). This corresponds to the difference between large notches on a standard thermometer.
- Clerical errors in data processing.
- Minimum thermometers shaken down by strong winds vibrating the screen. This is almost exclusively confined to exposed coastal sites and is manifested as an anomalously low minimum temperature relative to fixed-hour measurements and surrounding sites (noting that it is rare for such exposed coastal sites to have lower minimum temperatures than neighbours further inland).
- Failure to reset maximum or minimum thermometers, leading to a maximum or minimum being repeated on two (or more) consecutive days.
- Maximum temperatures one day out of alignment. This arises because standard observation practice, as noted earlier, is for maximum temperatures observed at 0900 to be attributed to the previous day, but some observers attributed them on occasions to the day of observation. (Most, but not all, such cases were detected during initial data processing.) This error is typically manifested as a large number of violations of maximum-fixed hour temperature relationships in a short span of time, and usually also a repeated maximum temperature on the first day of the affected period (which is often the first day of a month).

6.4 Treatment of accumulated data

In many cases, if a day of observations is missing, it indicates that, for a manually read maximum or minimum thermometer, the instrument was not reset on the day of the missed observation, and hence the first maximum/minimum temperature recorded following the period of missing data occurred over a period of two or more days. If treated as a single-day observation, such measurements can create a positive bias in mean maximum temperatures and a negative bias in mean minimum temperatures; at sites where observations were regularly missed (e.g., where no observations were made on Sundays, a common practice in the first half of the 20th century) the bias in mean diurnal temperature range can be up to 0.5°C (Trewin, 2001b).

Provision is made for the flagging of such accumulated data in ADAM but the actual flagging has been very inconsistently applied through time. Hence, for the ACORN-SAT data set, it was assumed that the first observation that followed a period of missing data was accumulated over the period of missing data, unless the period of missing data was longer than five days, or an automated instrument was used. Accumulated data are excluded from most downstream analyses but are still useful for some purposes (e.g., highest maximum and lowest minimum temperatures over a given period, such as a year or month).

7. DEVELOPMENT OF HOMOGENISED DATA SETS

A homogenous climate record is one that fluctuates and changes only in response to weather and climate variations (Conrad and Pollak, 1950). Non-climate influences (‘‘climate inhomogeneities’’) are present in most long-term climate records. Inhomogeneities can take a number of forms including abrupt, step change–like behaviour or more gradual, trend-like behavior (e.g., DeGaetano, 2006). An important part of the development of any data set to be used for long-term analyses is the detection and removal of all substantial inhomogeneities in the record.

The ideal temperature record for long-term analyses, as noted originally in section 3, is one where no changes have occurred in the site location, the instruments and methods used for observations, or in the surrounding site environment. At such idealised locations it could reasonably be assumed that any changes in the data reflect changes in the background climate at that location.

In practice, no century-long temperature records meeting such standards exist in Australia. In part, this is a consequence of the observing network in Australia, as in most countries, being developed principally to support weather forecasting; monitoring climate variability has been a secondary priority in temperature observations (although it has been a stronger consideration in the rainfall network, which is roughly ten times the size of the temperature network), and monitoring long-term climate change has only been a significant consideration in network planning in the last 20 years. (At the time of writing, only the United States has a dedicated network specifically for climate change monitoring (Baker and Helfert, 2008), and it is only in the last decade that it has become fully operational.) Previous Australian analyses (Torok and Nicholls 1996, Trewin 2001b, Della-Marta et al. 2004) have all found only a handful of locations with homogeneous records for either maximum or minimum temperature over a period of several decades or more.

Hence, it is necessary to ensure that all ACORN-SAT records are homogeneous. Even for those locations with no relevant documented changes, the metadata describing such changes, as discussed in section 5, may be incomplete. The homogenisation process is a two-stage process; first detecting all potential inhomogeneities, through the use of metadata and statistical methods, and then making adjustments to the data to remove the impact of the identified inhomogeneities and produce a homogeneous record.

The detection of inhomogeneities in a temperature record is a well-developed field of research (see section 7.2) and the methods used in the construction of the ACORN-SAT data set are closely based on those used previously for national-scale networks. However, adjustment of data to remove inhomogeneities at the daily timescale is a much less developed field, with the techniques used in ACORN-SAT not having been used outside Australia for a national-level data set. Previous work in both fields is discussed more fully in the relevant sections.

7.1 What particular issues exist for climate data homogenisation in Australia?

Many of the issues that affect the homogeneity of temperature data in Australia are common throughout the world (Peterson et al., 1998). Site and instrument changes, observation practice changes and changes in the local site environment are all characteristic of climate data sets everywhere, as are the limited accessibility and lack of completeness of metadata.

In some cases issues that significantly affect data in other countries are absent, or less significant, in the Australian data. For example, the retention of the Stevenson screen as the instrument shelter when automated instruments were introduced eliminates one potential source of inhomogeneity, while the existence of national standards for observation times contrasts with the wide range of observation times used in the American network (Karl et al., 1986).

A particular challenge for the homogenisation of Australian temperature data is the sparseness of the observation network in many parts of the country. A dense network is useful in detecting and adjusting for inhomogeneities, as it facilitates the availability of reference series that are well-correlated with the candidate site and can thus be used as an indicator of climate variations at the candidate site (Jones, 1999). (The use of reference series for ACORN-SAT is discussed more fully in section 7.2). Even today, 23 of the 112 ACORN-SAT locations are 100 kilometres or more from their nearest neighbour, and this number has been greater at times in the past, especially prior to 1950. This problem is particularly acute for locations in regions with steep temperature gradients, which in the context of the ACORN-SAT data set mainly affects coastal locations. A related issue is that the limited digitisation of pre-1957 daily data means that many potential reference sites only have monthly data available, which limits the ways in which they can be used to support homogenisation.

The role of site comparisons (initially discussed in section 3) is particularly important in a sparse network; in a denser network, it is more likely that suitable reference sites will be available without an explicit comparison. As discussed in section 3, most major site moves in the last 15 years at ACORN-SAT locations have at least some parallel comparison data available, although those data are not always useful in determining an appropriate adjustment. Comparisons were made less frequently in earlier periods, although some of the new Bureau-staffed sites opened during World War II had long overlaps (10 years or more) with the town-based sites they ultimately replaced.

7.2 The detection of inhomogeneities

Possible inhomogenities in the ACORN-SAT data were detected using a combination of metadata and statistical methods. The use of metadata is preferable if it is available, since it can demonstrate definitively that a change has occurred, and in many cases will also indicate the exact date of the change, whereas even the largest inhomogeneities detected by statistical means will have some level of uncertainty attached to their timing. However, statistical methods are also essential to find those inhomogeneities that are not identified in metadata, for the reasons discussed earlier in this report.

A comprehensive search of metadata, both hard-copy and electronic (see section 5), was undertaken to identify changes at a site that could indicate potential inhomogeneities, with a particular emphasis on site moves and significant developments in the vicinity of the observation site. This procedure includes the merging of records from two site numbers (something which was almost always associated with a site move) whether there was an overlap period or not. All such changes were viewed as potential inhomogeneities at this point. (In practice, some of these changes did not have any significant effect on temperature observations; such non-significant 'inhomogeneities' were filtered out of analyses during the adjustment process, as described in section 7.7.)

For statistical detection of inhomogeneities, the principal problem is that of determining where a breakpoint exists in a time series that is larger than can be attributed to chance (to a defined level of confidence). This problem has a well-developed statistical literature. Most of the techniques that have been used in constructing large-scale climate data sets fall into two broad categories: the standard normal homogeneity test (SNHT) of Alexandersson (1986), and two-phase regression (TPR), originally developed for climate use by Easterling and Peterson (1995), and with a number of refinements since, that have been implemented in the widely-used RHtest software suite (Wang et al., 2010). In a review Reeves et al. (2007)[13] found that the two methods achieved a broadly similar overall level of performance, with their ranking depending on user priorities (for example, accurately detecting the date of a breakpoint, or minimising the number of false alarms).

Since the ability to detect a breakpoint in a time series is a function of the ratio of the size of the breakpoint to the standard deviation of the data, a common technique to improve the signal-to-noise ratio is to apply statistical tests to the difference between the time series at the candidate site and that of a reference series that is representative of the background climate at the candidate site (Peterson et al., 1998). For locations that are highly correlated and have similar variance, this will tend to greatly reduce the variability in the test series by removing background noise arising from interannual climate variability, thereby increasing the signal to noise ratio for any tests.

Reference series are commonly constructed as a weighted mean of data from neighbouring locations. This method was used in the development of previous Australian temperature data sets (Della-Marta et al., 2004; Trewin, 2001a) and has also been used for Australian data sets for evaporation (Jovanovic et al., 2008) and cloud (Jovanovic et al., 2011). However, the use of such a reference series depends on the assumption that the reference series itself is broadly homogeneous around the time of the potential inhomogeneity being investigated, something that may not hold if, for example, there is a substantial change in the composition of the reference

---

[13] A more extensive benchmarking of different detection methods was carried out by Venema et al. (2012) but this was published too late to have any influence on the ACORN-SAT project as it currently stands.

series (e.g., a neighbouring site opening or closing), or a change affects a substantial proportion of the reference series around the same time as it affects the candidate site (e.g., an observation time change).

An alternative approach used by Menne and Williams (2009) is to undertake a series of pairwise comparisons, with the candidate site being compared one-by-one with its neighbours (that will indicate breakpoints both at a candidate location and at each neighbour), then using an iterative procedure to isolate which breakpoints are most likely attributable to the candidate site rather than one or more of its neighbours.

The method used for statistical detection (but not adjustment) of potential inhomogeneities in the ACORN-SAT data set broadly follows the method used by Menne and Williams (2009) for the continental United States, to which readers are referred for a full description. Details of the implementation of the method in the Australian context, and of deviations from the original method, are described in the following paragraphs.

The main deviations from the Menne and Williams method are:

- It is assumed that, where breakpoints are found from different neighbour series that are separated by no more than 1 year, these refer to the same inhomogeneity, rather than evaluating confidence intervals for each breakpoint date independently. This assumption, which simplifies the analysis considerably, is based on the results from the simulations reported in Menne and Williams that found that, for the vast majority of simulations, the breakpoint date was correctly simulated to within ±1 time step.
- All breakpoints are treated as step changes with no anomalous trend (model M3 of Menne and Williams). This assumption is based on their findings that the method was only moderately effective in reliably identifying more complex breakpoint models (e.g., an anomalous trend superimposed on a breakpoint). Note that a separate check for anomalous trends at locations potentially affected by urbanisation is described in section 9.
- Rather than using a fully iterative procedure as described in Menne and Williams to determine which breakpoints in difference series were attributable to the candidate site and which were attributable to the neighbour, pairwise differences were assessed for each site pair in the 41×41 matrix (candidate and 40 neighbours). A breakpoint in the difference series was considered to be attributable to the candidate rather than the neighbour if the number of breakpoints for that year in candidate-neighbour difference series for the candidate was greater than, or equal to, that for the neighbour. This also simplifies the analysis as the homogenisation is only being carried out for a relatively small proportion of the network (i.e., there is no specific need to homogenise the neighbours too).

For each candidate site, testing for inhomogeneities was carried out separately for time series of mean maximum and minimum temperature anomalies for annual means, and for seasonal means for each of the four calendar seasons (Dec-Feb, Mar-May, Jun-Aug, Sep-Nov). This procedure was followed because, in some cases, an inhomogeneity will vary seasonally – for example, if a site moves from a coastal location to one further inland, the difference in maximum temperatures between the sites is likely to be greatest in the warmer months and least in winter (Fig. 13). In such cases, a monthly difference time series will have an annual cycle that may affect the detection of breakpoints; the testing of seasonal series independently is to ensure the detection of any cases where there are opposite inhomogeneities in summer and winter, that could cancel out in the annual mean data. Data at monthly or longer timescales have been found to be more stable in the detection of inhomogeneities than data at shorter timescales (Trewin et al., 2010), while the testing of annual and seasonal anomalies separately addresses potential issues arising from an annual cycle in the variance of monthly or seasonal temperatures. However, the use of data at seasonal or longer timescales for detection of inhomogeneities does preclude the statistical detection of inhomogeneities in daily variance or other higher-order statistical properties that are not accompanied by a change in mean, unless supported by metadata.

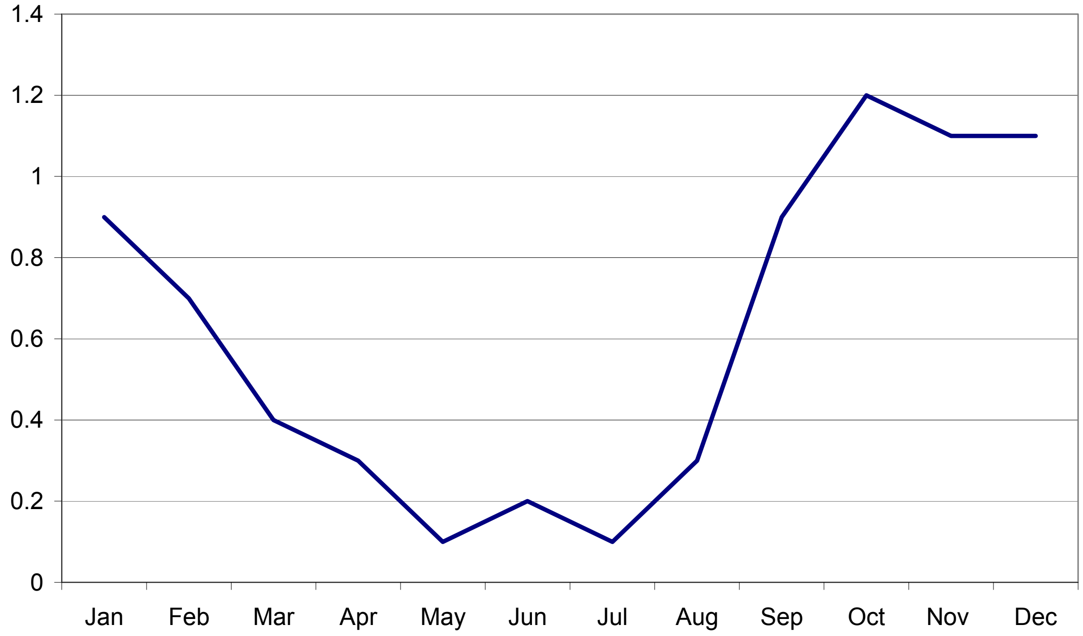


Fig. 13. Difference (Airport minus Hill Street) in mean monthly maximum temperatures (°C) between Port Macquarie Airport (060139) and Port Macquarie Hill Street (060026) during overlap period (1996-2002).

For each candidate site, 40 neighbouring site time series were chosen from all available Australian sites with some overlapping data with the candidate site. These neighbouring sites were initially chosen as the 40 best-correlated sites (using monthly anomalies for the period of overlap with the candidate site, and considering maximum and minimum temperatures separately) from amongst a pool drawn from the 150 nearest neighbours by distance. If this procedure resulted in a candidate site having fewer than 7 neighbours out of the 40 with available data in any given year, the 41st and subsequent sites were substituted for those sites in the original 40 with the least data, until at least 7 reference sites were available in each year. (Menne and Williams used a pool of 100 sites rather than 150, but this would not have been sufficient to achieve at least 7 available sites in all cases for Australia).

Once those significant breakpoints in candidate-neighbour difference series that were most likely attributable to the candidate had been identified (see third dot point above), the number of neighbouring sites that generated such breakpoints was checked. The breakpoint was considered to be potentially significant if this number of sites exceeded a specified threshold, bearing in mind that a small number of "significant" breakpoints could occur by chance (for example, if 40 neighbour sites are available in a given year then a mean of two difference series would be expected by chance to generate a breakpoint for that year significant at the 95% level). The threshold used was two sites if there were fewer than 5 sites with usable[14] data, three if there were 5-9, four if there were 10-19 and five if there were 20 or more; these thresholds were chosen as a number of sites for which there was less than a 5% probability that "significant" breakpoints in difference series could occur by chance. If potentially significant breakpoints were found in two or more consecutive years, the breakpoint was attributed to the year for which the greatest number of neighbour sites generated breakpoints in the difference series.

The potentially significant breakpoints from the annual and seasonal time series were consolidated. An inhomogeneity was considered to be significant if it was identified in the annual time series, or in at least two of the four seasonal time series (to within ±1 year). If both criteria were satisfied then the year of the inhomogeneity in the annual time series took precedence.

Finally, the inhomogeneities identified by metadata were consolidated with those found by statistical methods, with the metadata-identified inhomogeneity taking precedence if it occurred within 2 years of a statistically-identified inhomogeneity. All inhomogeneities were presumed, for the purpose of further analysis, to have taken place on 1 January unless a specific date could be identified from metadata.

7.3 Adjustment of data to remove inhomogeneities – an overview

Once potential inhomogeneities have been identified, the next step is to adjust the data to remove the effects of the inhomogeneity (normally by adjusting data prior to the inhomogeneity to make it homogeneous with the most recent data, although the reverse is also possible) and make the data set homogeneous. The practice of homogenising to the most recent data has clear advantages for ongoing monitoring as it allows new data to be simply appended to the location time series (until such time as the next inhomogeneity occurs).

Most adjustment techniques that have been used in large-scale climate data sets have used either a uniform annual adjustment (e.g., Della-Marta et al., 2004), or adjustments calculated for each of the 12 calendar months (e.g., Jones et al, 1986). These adjustments have typically been calculated by comparing location means, or their difference with a reference series, before and after an inhomogeneity. An advance on these methods involves using interpolation between monthly adjustments to produce a set of calendar-date adjustments that follow a smooth annual cycle (e.g., Brunet et al., 2006; Vincent et al., 2002); such methods are sometimes referred to as involving "daily adjustment", but do not use weather-dependent or distribution-dependent daily adjustments.

A characteristic of such methods is that they seek to produce a data set whose monthly or annual means are homogeneous. It does not, however, follow that such an adjusted data set also has homogeneous higher-order statistical properties, such as variance or the frequency of extremes. This issue was identified by Trewin and Trevitt (1996), who noted that the temperature difference between sites could be weather-dependent, with, for example, ridge-valley minimum temperature differences typically being larger on cold nights (which tended to be calm and clear) than on warm nights (which tended to be cloudy and/or windy). Figure 14 illustrates one such example, for summer maximum temperatures at Albany Airport (which is about 10 km inland) and Albany Town (which is on the coast); the coastal site is between 1 and 2°C cooler than the airport on most days in the bottom half of the frequency distribution, but the

---

[14] For this purpose, a neighbour site with 'usable' data is one which has data for the year under consideration and at least 3 years before and after.

difference progressively increases on hotter days, reaching 6-8°C on days between the 95th and 99th percentile. The difference then collapses to near zero on the very hottest days, when strong northerly winds are sufficient to override sea breezes even at the coast – this creates further complications for adjustment methods, as discussed in section 7.8.

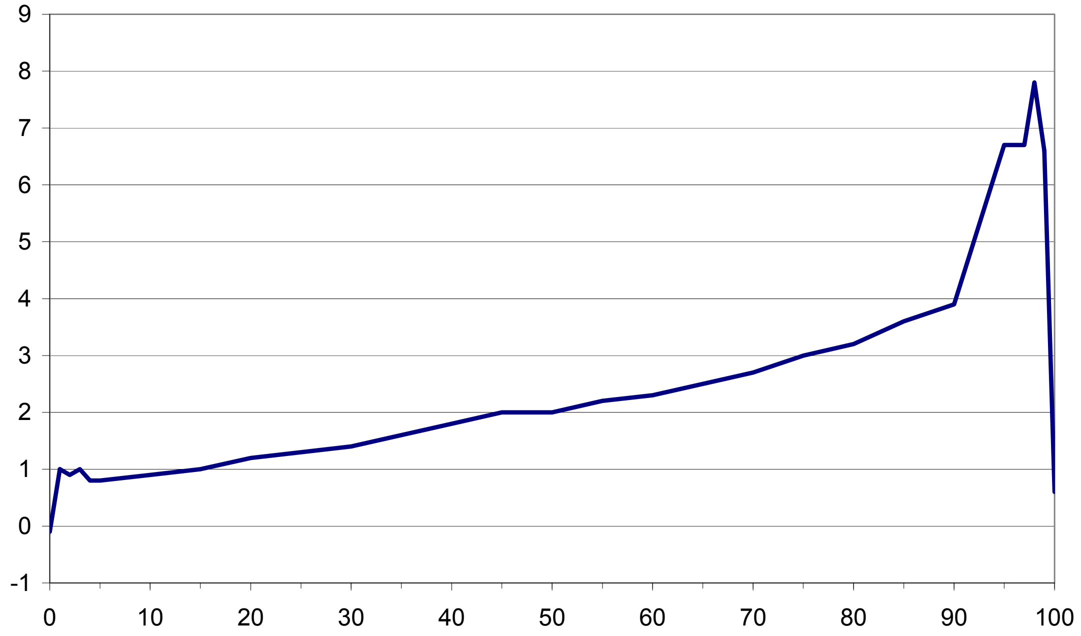


Fig. 14. Differences (°C; Airport minus Town) between percentile points of summer maximum temperature at Albany Airport (009741) and Albany Town (009500) during overlap period (2002-2009). The 0th and 100th percentiles indicate the lowest and highest summer maximum temperatures recorded during the overlap period.

In recent years a number of attempts have been made to address this problem of weather-dependent inhomogeneities. Some have involved explicitly testing the homogeneity of higher-order statistical properties, such as mean daily variability (Wijngaard et al., 2003) or exceedances of percentile-based thresholds (Allen and DeGaetano, 2000), while others have sought to homogenise daily data across the full range of the frequency distribution, by matching percentile points in the frequency distribution (Della-Marta and Wanner, 2006; Trewin, 2001a) or by other methods (Brandsma and Können, 2006; Wang et al., 2010); Trewin (2001a) was the first known attempt to produce a homogenised national-scale data set at the daily timescale. This is currently an active area of research, in particular through the European COST Action on 'Advances of homogenisation methods of climate series: an integrated approach' (COST, 2009).

7.4 The percentile-matching (PM) algorithm

In this section a detailed description of the percentile-matching (PM) algorithm for data adjustment is given, with two main variations (PM95 and PM99). This algorithm is similar conceptually to those used by Trewin (2001a) and Della-Marta and Wanner (2006), although there are some differences, principally in the details of generating transfer functions. It has also recently been applied to the related problem of matching climate model output to observations (e.g. Grose et al., 2010).

The PM algorithm takes two forms. The first, simpler, form is for the case of merging data from two sites where there is a useful overlap between sites. The second, more complex case, is where there is no overlap (or an overlap too short to be useful), and the adjustment is a two-step process involving the use of neighbouring sites.

7.4.1 The overlap case

In cases where two sites have at least one year of overlap with at least 50 observations in common for each set of three consecutive months of the year, the algorithm involves the following steps.

(a) For each site, calculate daily temperature anomalies from the mean. These anomalies were from a daily climatology, calculated by linear interpolation from the monthly means (attributing the monthly mean to the middle day of each month), with the monthly means calculated using a 1961-1990 base period if there were at least 12 years of data in the 1961-1990 period, or all years of record otherwise.

(b) Select the period used for data-matching between the sites. (In most cases this was the full period of overlap, but a subset was chosen in some cases where an inhomogeneity had been identified at one or both of the sites during the overlap period.) For this period, for each of the 12 calendar months, select the daily anomalies for all days that have data for both sites, and which are either in the target month or the month either side (for example, if March is the target month, February-April data are used). For example, if there are 5 years of overlap with complete data, this will result in a comparison set with between 450 and 460 paired observations for each target calendar month.

(c) For the comparison set of daily anomalies, for each site, calculate percentile points for the 1st, 2nd, 3rd, 4th, 5th, 10th, 15th,…., 90th, 95th, 96th, 97th, 98th and 99th percentiles.

(d) Reconvert the anomaly percentile points to temperatures by adding the monthly mean for the target month.

(e) Define a transfer function using the percentile-point pairs as fixed points. This takes two forms. For the PM95 variation, the fixed points for defining the transfer function are the 5th through 95th percentiles; for values below the 5th percentile, the temperature difference between the sites is assumed to be the same as the difference between the 5th percentiles (and similarly for values above the 95th percentile), with linear interpolation between the two nearest fixed points being used for values between the 5th and 95th percentiles.

For the PM99 variation, linear interpolation is used for the full range between the 1st and 99th percentiles, with the inter-site difference for values below the 1st percentile assumed to be the same as the differences between the 1st percentiles (and similarly for values above the 99th percentile). A PM90 variation can also be defined using the 10th and 90th percentiles as the limits of linear interpolation.

(f) Convert values at the old site to equivalent values at the new site using the transfer function to produce a composite record, with the new site taking precedence where data exist at both sites.

7.4.2 The non-overlap case

The majority of adjustments required did not have any overlapping data, either because they did not involve a change of site number, with an inhomogeneity that was identified either through metadata or by statistical methods, or because a change of site number occurred with no overlap or with insufficient overlap (generally less than one year) to define a transfer function.

In the non-overlap case, the algorithm operates as follows:

(a) Identify a set of *N* neighbouring sites with sufficient overlapping data with the candidate location both pre- and post-inhomogeneity (a minimum of 50 observations for each set of three consecutive months of the year).

(The way in which the neighbours were identified in the ACORN-SAT data set is described in later sections).

(b) For each neighbour separately, define a transfer function between the candidate site pre-inhomogeneity and the neighbour, using the method for the overlap case as described above. The period of comparison was generally the five calendar years prior to, but not including, the year of inhomogeneity (e.g., if the inhomogeneity was in 1994, 1989-1993 data were used), although in some cases this was shortened if there was a known inhomogeneity during the five-year period.

(c) Also for each neighbour separately, define a transfer function between the neighbour and the candidate site post-inhomogeneity, also using the method for the overlap case described above, with the period of comparison generally the five calendar years following, but not including, the year of inhomogeneity.

(d) For each value at the old site, create an ensemble of *N* estimated equivalent values at the new site, using the *N* pairs of transfer functions defined in steps (b) and (c) (each of which converts a value at the old site to an equivalent value at a neighbour, then the equivalent value at that neighbour to a value at the new site).

(e) Calculate the 'final' equivalent value as the median of the ensemble of N estimates defined above.

Figure 15 illustrates this process, based on a site move at Kerang (Victoria) in 2000 and using Swan Hill, Echuca, Deniliquin and Bendigo as the nearest neighbours. Figure 15(a) demonstrates a transfer function based on Swan Hill; based on matching frequency distributions, a July minimum temperature of 0°C at the pre-2000 Kerang equates to −1.2°C at Swan Hill, which in turn equates to −0.5°C at the post-2000 Kerang. Hence, using Swan Hill as the neighbour, a temperature of 0°C at Kerang pre-2000 would be adjusted to −0.5°C to be homogeneous with the post-2000 period. Figure 15(b) then shows how the four neighbours are composited to produce the final transfer function, that shows adjustments near −0.4°C near the middle of the frequency distribution, −0.8°C at the lower end but only −0.1°C at the upper end, indicating that the site move had far more impact on temperatures on cold nights than on warm ones (a quite common situation).

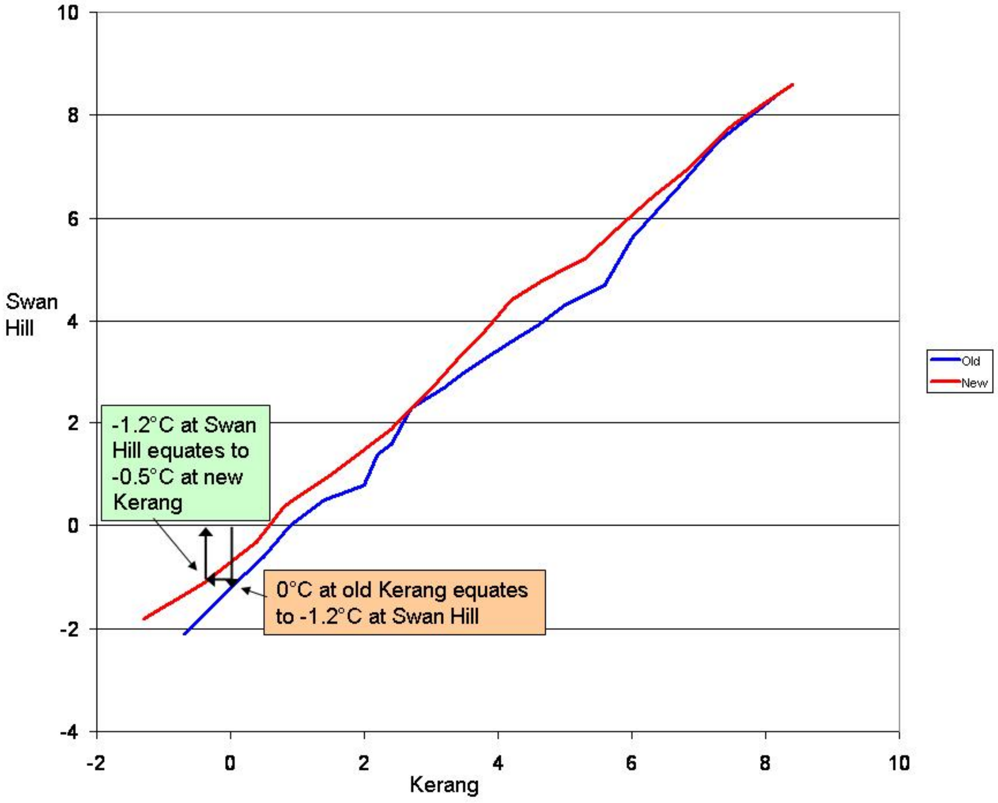


Fig. 15 (a). Example of two-step adjustment procedure – winter minimum temperatures (°C) at Kerang and Swan Hill.

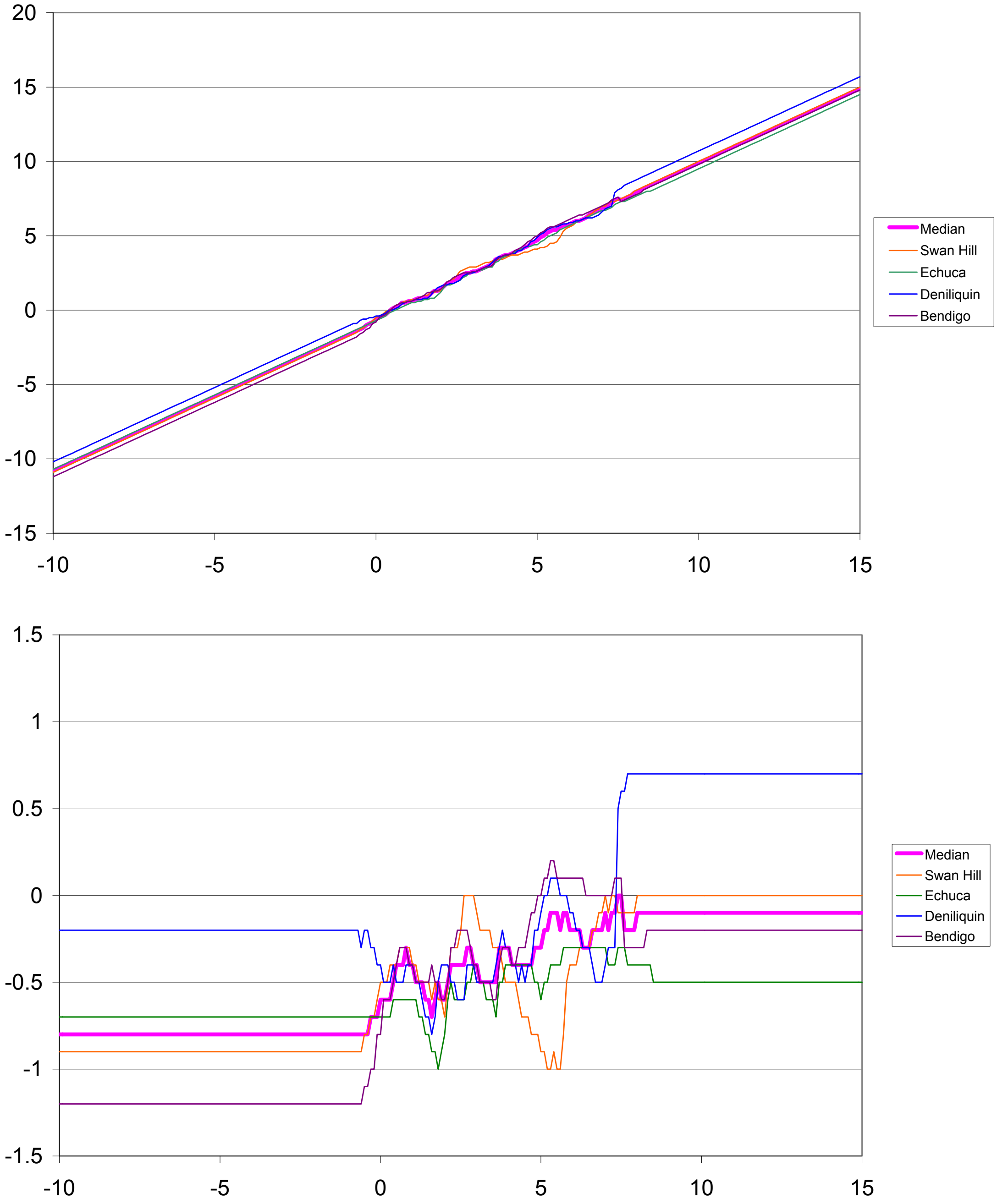


Fig. 15(b). (top) Example of transfer function for winter minimum temperatures (°C) at Kerang; (bottom) transfer function expressed as inter-site differences.

7.5 Monthly adjustment method

An adjustment method was also defined using monthly data, for use in method evaluation (see below), and for cases where insufficient neighbours existed with available daily data for the PM algorithm to be used. This method has similarities with the adjustment method used in the annual data set of Della-Marta et al. (2004).

These monthly adjustments were calculated as follows:

(a) An interpolated estimated monthly anomaly was calculated for the location of the candidate site for each month in the candidate site's record. This anomaly was calculated as a weighted mean of the monthly anomalies from a set of neighbouring sites other than the candidate sites, with the weighting function:

$w_s = 1/\exp((d/100)^2)$, where d is the distance in kilometres between site s and the candidate site.

Since a common base period is important for this interpolation, each site monthly anomaly was calculated using a 1961-1990 base period. Where the site did not have at least 12 years of observations in the 1961-1990 base period, first a mean was calculated for all years of record, then it was corrected to a 1961-1990 equivalent using those neighbours that did have at least 12 years of observations in the 1961-1990 base period (e.g., if a site had data from 2004-2009, its January mean maximum for that period was 32.9°C, and those neighbours that did have sufficient 1961-1990 data had a mean January maximum anomaly over the 2004-2009 period of +0.4°C, the corrected 1961-1990 equivalent mean at the site was (32.9−0.4) = 32.5°C).

(b) For each of the 12 calendar months, and for (in general) the five calendar years prior to the inhomogeneity, calculate the difference between the anomalies at the candidate site and the interpolated estimate from neighbouring locations.

(c) Repeat step (b) using the five calendar years following the inhomogeneity.

(d) Calculate the monthly adjustment as the difference between the two differences calculated in steps (b) and (c). (For example, if the candidate site has January maximum temperature anomalies 0.3°C lower than its neighbours pre-inhomogeneity and 0.4°C higher post-inhomogeneity, the adjustment for pre-inhomogeneity January maximum temperatures is +0.7°C.)

7.6 Evaluation of different adjustment methods

Various adjustment methods were evaluated to verify that the best possible methods were being used. The evaluation was carried out using a set of 16 ACORN-SAT locations (Table 3) that had a period where two sites had overlapping data, and where no known inhomogeneity existed during the overlap period at either site. The overlap data covered time spans ranging between 4 and 11 years during the 1992-2009 period.

Table 3 Site pairs used for adjustment technique evaluation. (Locations in italics not used for the 1930 network comparisons)

Location	Site numbers (old/new)	Period of parallel observations	Length of parallel observations (yrs)	Mean temperature difference (new – old, °C) and classification	
				Maximum	Minimum
Kalumburu	1021/1019	1998-2005	6.5	-0.05 (small)	-1.14 (large)
Cunderdin	10035/10286	1996-2007	10.5	0.09 (small)	-0.55 (medium)
Wandering	10648/10917	1998-2003	4.2	-0.34 (medium)	-0.65 (large)
Port Lincoln	18070/18192	1992-2002	9.9	-0.76 (large)	-1.18 (large)
Burketown	29004/29077	2001-2009	7.7	0.91 (large)	-0.34 (medium)
Birdsville	38002/38026	2000-2005	4.7	-0.23 (medium)	0.20 (small)
Gayndah	39039/39066	2003-2009	6.4	-0.28 (small)	-0.25 (small)
Brisbane AP	40223/40842	1994-2000	5.8	-0.23 (small)	0.23 (medium)
Miles	42023/42112	1997-2005	7.6	-0.34 (medium)	-0.05 (small)
Thargomindah	45017/45025	1999-2005	5.7	-0.51 (medium)	0.74 (large)
Port Macquarie	60026/60139	1995-2003	7.6	0.62 (large)	-1.43 (large)
Dubbo	65012/65070	1993-1999	6.9	-0.41 (medium)	-0.16 (small)
Deniliquin	74128/74258	1997-2003	6.1	-0.14 (small)	-0.49 (medium)
Nhill	78031/78015	2003-2008	5.5	0.34 (medium)	0.92 (large)
Sale	85298/85072	1996-2008	9.2	-0.16 (small)	-0.71 (large)
Launceston	91104/91311	2004-2009	4.9	0.69 (large)	-0.10 (small)

The locations were classified into the categories of large differences between pairs (difference in annual means during the overlap period ≥ 0.6°C), medium differences (difference in annual means during the overlap period 0.3-0.6°C, or a difference in seasonal means ≥ 0.3°C for at least two of the four seasons, or ≥ 0.5°C for one season), and small differences (all other locations). For maximum temperature four locations were categorised as having large differences, six locations medium differences and six locations small; for minimum temperature the numbers were seven, four and five respectively.

In each case, a potentially inhomogeneous 'test' series was created by switching from the older site to the newer site at the start of the overlap period. Data from the period after the switch were then adjusted to be homogeneous with the older data (the reverse of, but effectively equivalent to, the process using for the ACORN-SAT homogenisation). The accuracy of the adjustment was then evaluated for the overlap period, using the continuation of the old site as the 'truth'.

The methods that were evaluated were as follows:

(a) No adjustment (the 'control' case)
(b) The PM99 algorithm, using the nearest sufficiently-correlated[15] neighbours (up to 10) with available daily data around the time of the inhomogeneity.
(c) As for (b), but considering only those neighbours that also had available daily data in 1930. (This was to test the performance of the method in the sparser networks typical of earlier years).
(d) As for (b), but using the PM95 algorithm.
(e) As for (b), but using the PM90 algorithm.
(f) As for (d), but using a maximum of 5 neighbours.
(g) Monthly adjustments, using the nearest sufficiently-correlated neighbours (up to 10) with available monthly data around the time of the inhomogeneity.

---

[15] Correlation of monthly temperature anomalies with the candidate station 0.6 or greater.

(h) As for (g), but considering only those neighbours that also had available monthly data in 1930.
(i) The quantile-matching (QM) algorithm used in the RHtestsV3 software (Wang et al., 2010).

Only 13 locations were used for tests (c) and (h) due to a lack of available 1930 data at the sites and/or their neighbours.

The metrics that were used for evaluation were:

- The daily root-mean-square (RMS) error.
- The proportion of all observations where the simulated value was within 0.5°C of the actual value.
- The percentage difference between the actual and simulated number of days with maximum and minimum temperatures above the 90th percentile, and below the 10th percentile (calculated using a 1961-90 base period, using the index definition of the ETCCDI - http://cccma.seos.uvic.ca/ETCCDI/list_27_indices.shtml).
- The difference between the actual and simulated highest and lowest value of the maximum and minimum temperature for each of the 12 months during the overlap period.

The results of this evaluation are shown in Tables 4 and 5, and Figs. 16 and 17. The key points to emerge from the evaluation are as follows:

- All adjustment methods fail to consistently outperform the control case (no adjustment, method (a)) for locations with small inhomogenities (e.g. Fig. 16), suggesting that 0.3°C is near the lower limit for the size of inhomogeneity that can be adjusted for with useful skill.
- The QM method (i) performs more poorly than the PM family ((b) through (f)) for almost all metrics, and only outperforms the control case for maximum temperature for locations with large inhomogeneities, although for minimum temperature it also outperforms the control case for locations with medium inhomogeneities. It should be noted that the QM method does not use reference series for adjustment, and is hence likely to perform poorly when applied in situations where there are rapid changes in the background climate; the results for maximum temperature in this evaluation are driven to a large extent by particularly poor results for four locations in inland New South Wales and Queensland, a region where five-year mean maximum temperatures warmed by up to 1°C over the 1992-2009 period.
- There is no appreciable difference between the daily (PM; (b) through (f)) and monthly ((g) and (h)) adjustment methods for the RMS and proportion within 0.5°C metrics. However, the daily methods outperform the monthly methods substantially in simulating extremes, especially for locations with large inhomogeneities, with the partial exception of extreme high maximum temperatures, for which in some cases the two method classes perform similarly.
- The PM95 method (d) performs similarly to PM99 (b) on the first three metrics, but is generally much better than PM99 in simulating the highest and lowest values. It is likely that this reflects instability in the transfer functions towards the ends of the distribution when the 1st and 99th percentiles are used; in most cases those percentiles are based on only a few observations, and may therefore be vulnerable to data quality issues at neighbour sites, or effects of unusual weather events. The PM90 (e) method performs marginally worse than PM95 on most measures.
- The five-neighbour case (f) performs marginally worse than the 10-neighbour case (d) on most measures.

- Across the set of evaluation locations as a whole, the relative performance of daily and monthly methods using the 1930 network ((c) and (h)) was similar to that using the more recent network ((b) and (g) respectively). However, at some individual locations where the availability of neighbouring monthly data in 1930 was much better than that of neighbouring daily data, the monthly adjustment method outperformed daily methods.

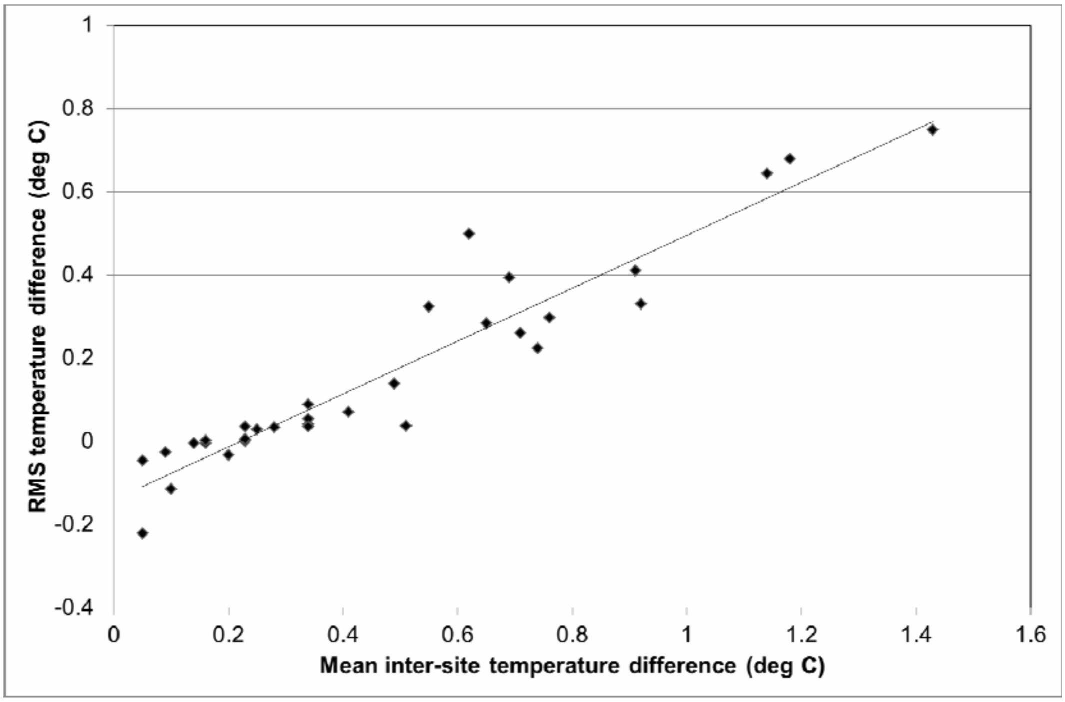
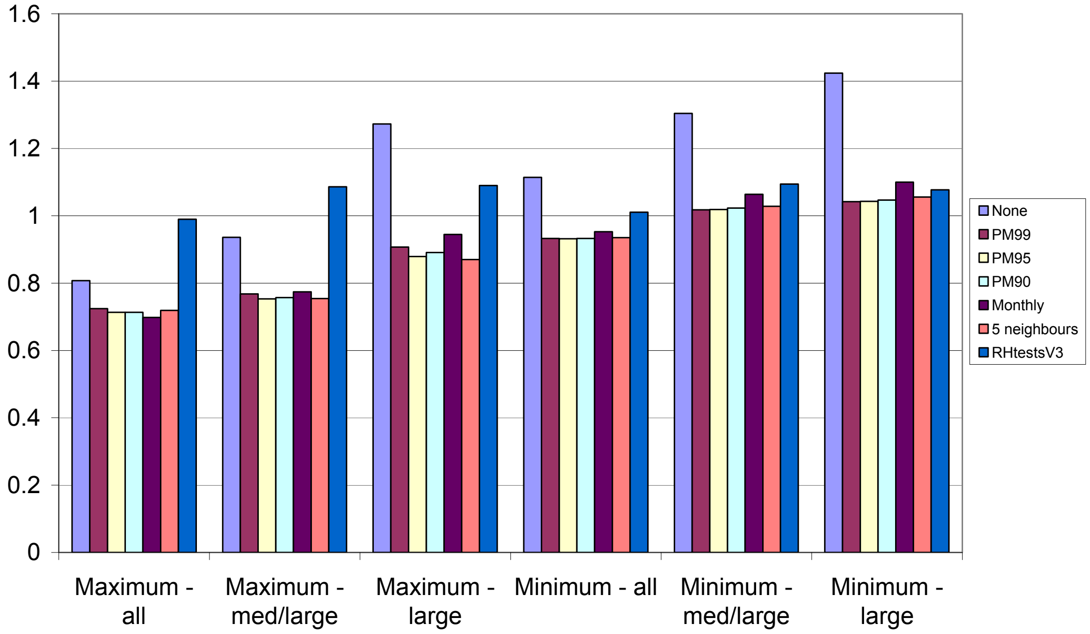


Fig. 16. Difference in RMS errors (no adjustment − PM95 algorithm) for station pairs, by size of mean temperature differences between paired stations.

Based on this evaluation, the PM95 method with 10 neighbours was selected for use in most cases. Details of the implementation are given in the following section.



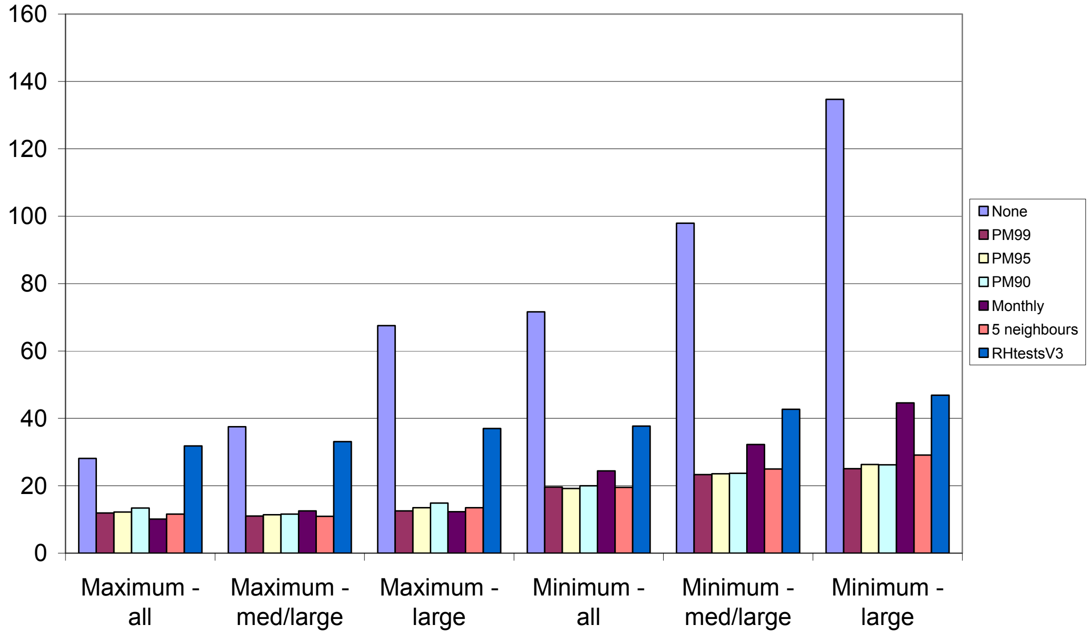


Fig. 17. Performance measures of adjustment techniques across different classifications of site pairs: (top) RMS error (°C), (bottom) % error in count of days with maximum above 90th percentile and minimum below 10th percentile.

Table 4 Comparison of adjustment methods, current network (M – medium; L – large).

Test	Variable	Station type	Adjustment method						
			None	Daily PM99	Daily PM95	Daily PM90	Monthly	5 neighbours	RH-tests
RMS (°C)	Max	All	0.808	0.711	0.704	0.703	0.702	0.705	0.991
		M/L	0.930	0.754	0.742	0.746	0.766	0.744	1.080
		L	1.263	0.888	0.865	0.879	0.931	0.856	1.077
	Min	All	1.126	0.913	0.908	0.910	0.950	0.911	1.021
		M/L	1.321	0.991	0.988	0.993	1.061	0.998	1.108
		L	1.451	1.003	1.000	1.006	1.092	1.009	1.099
Prop within 0.5°C	Max	All	0.491	0.568	0.568	0.572	0.588	0.567	0.331
		M/L	0.426	0.554	0.557	0.558	0.565	0.549	0.307
		L	0.280	0.520	0.528	0.525	0.536	0.531	0.385
	Min	All	0.399	0.467	0.472	0.469	0.449	0.474	0.396
		M/L	0.310	0.429	0.433	0.429	0.392	0.431	0.380
		L	0.262	0.417	0.423	0.418	0.368	0.423	0.381
Indices count (mean % error)	Max 10th p'ile	All	28.7	15.4	14.6	12.8	17.1	13.4	55.2
		M/L	37.1	14.6	14.4	12.5	20.8	14.1	64.5
		L	65.1	18.1	15.2	15.7	38.5	15.8	81.1
	Min 10th p'ile	All	75.4	17.4	16.0	18.4	28.5	16.3	40.6
		M/L	103.4	21.3	19.8	22.3	37.8	19.7	47.0
		L	143.4	23.5	20.6	23.9	52.9	20.1	53.3
	Max 90th p'ile	All	28.3	11.3	11.8	11.3	11.0	12.1	31.3
		M/L	37.5	9.3	9.8	9.8	12.4	10.1	33.2
		L	67.3	10.4	12.4	12.2	11.8	12.1	37.2
	Min 90th p'ile	All	16.6	11.7	9.9	9.5	19.9	9.8	20.4
		M/L	20.2	13.0	10.4	10.2	22.3	10.2	19.6
		L	23.2	15.1	11.1	10.2	22.9	10.5	14.9
Extremes error (°C)	Lowest max	All	0.46	0.60	0.50	0.49	0.49	0.50	0.71
		M/L	0.44	0.58	0.45	0.55	0.48	0.54	0.74
		L	0.55	0.68	0.49	0.56	0.66	0.54	0.71
	Lowest min	All	1.19	0.78	0.69	0.75	0.85	0.70	0.89
		M/L	1.48	0.82	0.76	0.84	1.01	0.78	1.02
		L	1.64	0.81	0.71	0.83	1.05	0.71	0.94
	Highest max	All	0.76	0.73	0.61	0.63	0.65	0.63	0.90
		M/L	0.87	0.75	0.58	0.65	0.69	0.68	0.98
		L	1.22	1.01	0.70	0.63	0.94	0.66	0.85
	Highest min	All	0.54	0.66	0.58	0.59	0.69	0.60	0.69
		M/L	0.58	0.69	0.58	0.59	0.74	0.61	0.69
		L	0.63	0.79	0.63	0.65	0.83	0.66	0.69

Table 5 Comparison of adjustment methods, 1930 network.

Test	Variable	Station type	Adjustment method			
			None	Daily PM99	Monthly	RH-tests
RMS (°C)	Max	All	0.804	0.763	0.742	1.005
		M/L	0.904	0.824	0.817	1.097
		L	1.298	0.970	1.038	1.127
	Min	All	1.162	1.010	1.020	1.041
		M/L	1.348	1.060	1.106	1.111
		L	1.440	1.045	1.128	1.073
Prop within 0.5°C	Max	All	0.514	0.535	0.550	0.332
		M/L	0.450	0.508	0.507	0.302
		L	0.304	0.493	0.498	0.398
	Min	All	0.380	0.408	0.407	0.379
		M/L	0.293	0.384	0.371	0.367
		L	0.269	0.394	0.358	0.391
Indices count (mean % error)	Max 10th p'ile	All	28.7	17.9	21.1	59.0
		M/L	37.1	20.8	23.5	69.3
		L	65.1	27.1	37.4	101.0
	Min 10th p'ile	All	75.4	20.1	30.4	34.3
		M/L	103.4	16.9	37.6	37.4
		L	143.4	14.1	47.1	41.6
	Max 90th p'ile	All	28.3	10.1	12.1	25.4
		M/L	37.5	12.2	14.7	26.4
		L	67.3	13.5	15.5	18.2
	Min 90th p'ile	All	16.6	12.5	18.9	19.6
		M/L	20.2	13.2	21.0	19.1
		L	23.2	11.2	24.5	11.3
Extremes error (°C)	Lowest max	All	0.46	0.67	0.52	0.74
		M/L	0.44	0.69	0.52	0.76
		L	0.60	0.73	0.66	0.78
	Lowest min	All	1.21	0.90	0.89	0.86
		M/L	1.48	0.89	1.03	0.95
		L	1.58	0.85	1.00	0.84
	Highest max	All	0.78	0.90	0.70	0.94
		M/L	0.84	0.99	0.75	0.98
		L	1.26	1.37	1.11	0.81
	Highest min	All	0.58	0.72	0.74	0.73
		M/L	0.60	0.70	0.79	0.73
		L	0.64	0.75	0.92	0.73

7.7 Implementation of data adjustment in the ACORN-SAT data set

The final implementation of data adjustment was carried out using the following rules:

- Other than in the specific cases outlined below, the PM95 method was used. Neighbours (up to 10) were selected in descending order of correlation[16] with the candidate site, with a lower correlation limit of 0.6.
- If fewer than 3 sufficiently-correlated sites existed with daily data around the time of the inhomogeneity, or the 10th best-correlated site with monthly data was better correlated than the 3rd best-correlated site with daily data, the monthly adjustment method was used.
- In the event of a 'spike' (defined as an inhomogeneity, followed by another inhomogeneity of opposite sign within 3 years), monthly adjustments were used to adjust the data during the 'spike' period to the period before the 'spike', as daily adjustments were considered to be too unstable for such short-period corrections. The 'spike' period was also excluded from the overlaps used for the calculation of transfer functions for other inhomogeneities.
- If no sufficiently-correlated sites existed with either monthly or daily data around the time of the inhomogeneity, RHtestsV3 (which does not require a reference series) was to be used. This provision was not ultimately required for any ACORN-SAT location.
- The size of adjustments generated by the techniques above was checked by comparing the means of pre- and post-adjustment data for the five calendar years prior to the inhomogeneity. The adjustment was only implemented if the means differed by at least 0.3°C on an annual basis, 0.3°C (not necessarily of the same sign) in at least two of the four seasons, or 0.5°C in at least one season (the last two provisions are in order to capture inhomogeneities that have different impacts in different seasons which might cancel each other out in an annual mean[17]). If the difference failed to satisfy one or more of these criteria the inhomogeneity was considered to be too small to justify adjustment.
- Adjustments were not made for statistically-identified inhomogeneities within three years of the start or end of the time series, or for metadata-identified inhomogeneities within two years of the start or end of the time series, due to the limited reference data available for adjustment. (The option exists, although it was not applied to the 1910-2009 ACORN-SAT data set, of temporarily excluding one or more locations from downstream analyses should it become apparent that a large inhomogeneity has occurred in the last two to three years).
- A special case was the 1965 move at Albany from the town to the airport, with no overlap. A site was re-established in Albany town in 2002 and monthly data indicated it was approximately homogeneous with the pre-1965 town site (and much better-correlated with the airport than any other neighbour). The post-2002 town site was

---

[16] For the purposes of this section, the 'correlation' value is calculated as follows: correlations between daily temperature anomalies are calculated for each of the 12 months, and the final 'correlation' value is taken as the 6th highest (i.e. near the median) of these 12 monthly correlations.

[17] For example, in the period of overlapping data from 1997-2001 for two sites at Snowtown (South Australia), there is only 0.1°C difference in annual mean maximum temperature between the two sites, but the new site is 1.3°C cooler than the new site in winter, and 0.9°C warmer in summer. A possible explanation for this seasonal difference is that the new site is in cropland where the predominant ground cover is green vegetation in winter and bare soil in summer, whereas the old site, in town, has similar ground conditions all year.

therefore used as a proxy for the pre-1965 site, with the adjustments for the 1965 move calculated using a transfer function derived from the 2004-09 overlap. (An analogous situation existed, at first glance, at Bundaberg, where there were observations at Bundaberg Airport from 1959-1970 prior to the long-term move from the town to the airport in 1990, but the 1959-1970 airport observations were found not to be homogeneous with the post-1990 observations and hence were not used in calculation of the 1990 adjustment).

After the first round of homogenisation, the homogenised data sets were evaluated, using the following tools:

- An examination of time series of annual means, highest and lowest values, and number of days above the 90th and below the 10th percentiles in the raw and adjusted data for each location.
- Mapping of trends in mean annual and seasonal maximum and minimum temperature, and in the number of days above the 90th and below the 10th percentiles. Locations that showed trends which differed substantially from those of other ACORN-SAT locations in their region were subject to further scrutiny. (For example, Fig. 18 shows an anomalous upward trend in the frequency of extreme cold nights at Nhill (Victoria), possibly indicating inadequate adjustments for an inhomogeneity associated with a site move at the start of 1995).
- The generation of time series of area-averaged data using the beta version of the ACORN-SAT data set, and a comparison of these with other area-averaged data sets (e.g. the AWAP data set and the annual data set of Della-Marta et al. (2004)).

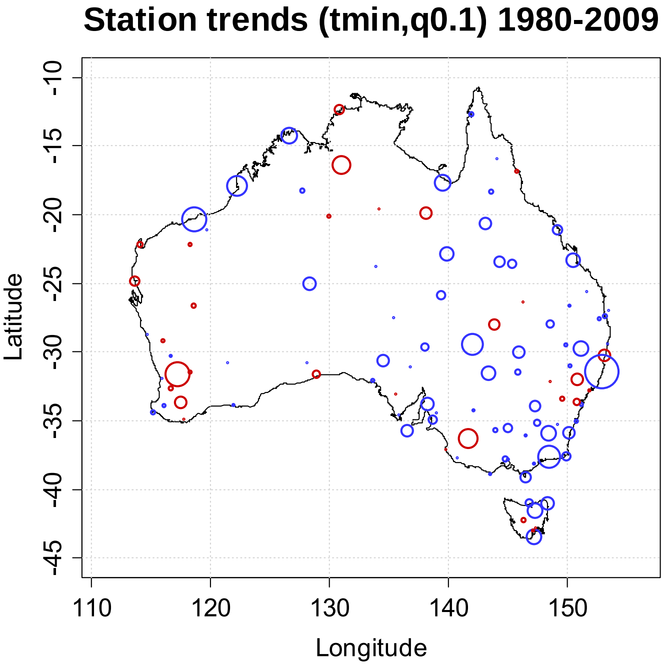


Fig. 18. Trends (1980-2009) in number of days with minimum temperature below the 10th percentile in the initial iteration of the ACORN-SAT data set – red indicates positive trends and blue negative, with the radius of the circle being proportional to the magnitude of the trend.

In a number of cases, this evaluation identified locations whose homogenised data produced apparently anomalous trends in one or more indicators relative to their neighbours. In such cases, a comparison was undertaken between the homogenised data from these locations and homogenised data from other ACORN-SAT locations in the region, then particular breakpoints whose adjustments appeared doubtful were reassessed, using different sets of reference stations and different sets of years before and after the breakpoint for calculation of transfer functions. The most common issues identified by this process were cases where a station had a breakpoint at a similar time to several of its neighbours (in which case the adjustment was recalculated without those inhomogeneous neighbours), or where temperature differences between a site in the period immediately prior to a breakpoint were not representative of the broader pre-breakpoint period, e.g. where a site's condition deteriorates rapidly in the 1-2 years prior to a move (in which case the adjustment was recalculated with a different reference period).

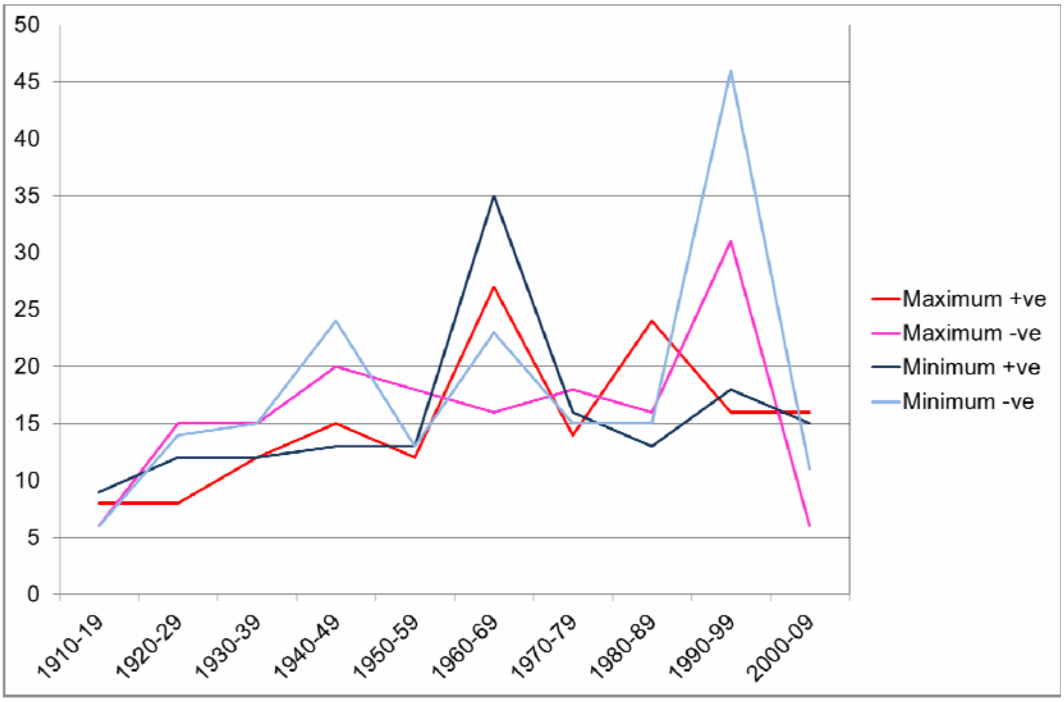
This was continued as an iterative process until the number of spatial anomalies was minimised.

The frequency distribution of adjustments, and the distribution of positive and negative adjustments over time, is shown in Table 6 and Fig. 19. There is an approximate balance between positive and negative adjustments for maximum temperature but a weak tendency towards a predominance of negative adjustments (54% compared with 46% positive) for minimum temperature. As shown in Fig. 19, negative adjustments of minimum temperature predominated over positive adjustments mainly in two decades, the 1940s and 1990s, which were both periods when substantial numbers of ACORN-SAT locations had site moves from town centres to airports (some of which involved a change of site number, and some of which did not).

Details of the timing of all inhomogeneities, whether or not they were supported by metadata, and the transfer functions used for adjustment will be made available with the ACORN-SAT data set.

Table 6 Summary of adjustments carried out in ACORN-SAT data set.

	Maximum temperature	Minimum temperature
Total number of adjustments	315	345
Number of adjustments supported by metadata (number related to site moves in brackets)	160 (138)	171 (141)
Number of positive adjustments	153 (49%)	154 (46%)
Number of negative adjustments	160 (51%)	184 (54%)
Number of 'spike' adjustments (excluded from percentages above)	2	7
Number of adjustments > +1°C (number metadata-supported in brackets)	9 (6)	9 (3)
Number of adjustments < −1°C (number metadata-supported in brackets)	8 (7)	24 (17)



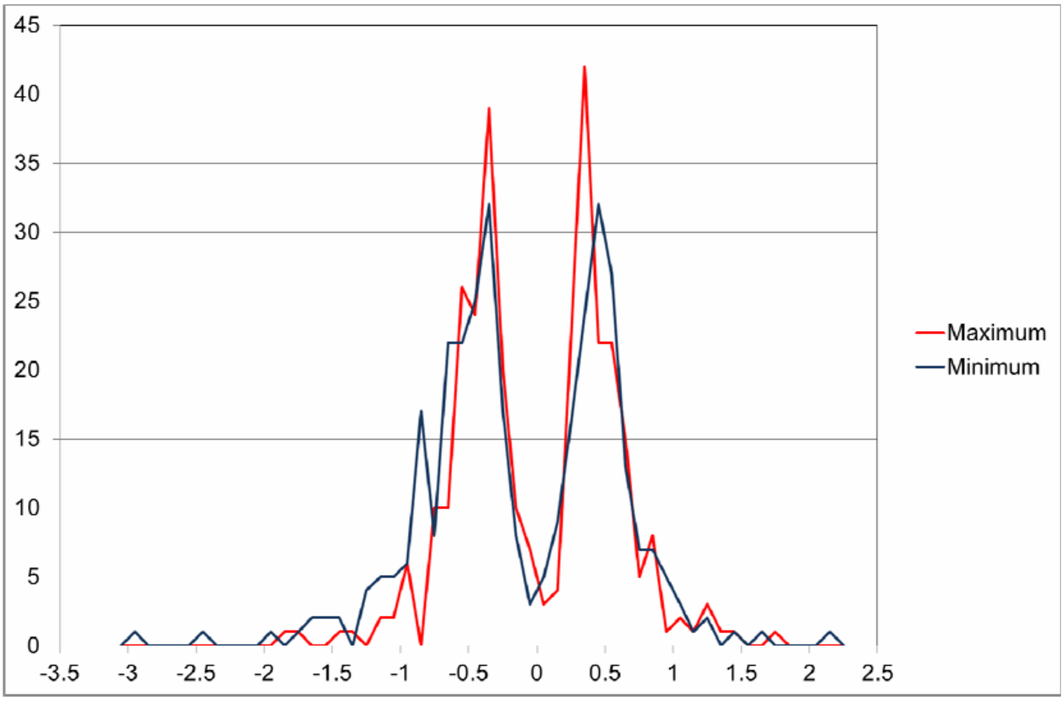


Fig. 19. (top) Number of positive and negative adjustments by decade, for maximum and minimum temperature. (bottom) Frequency distribution of mean annual adjustment size (°C) for maximum and minimum temperature.

7.8 Identification of locations whose extremes were not homogenisable

At a small number of locations, the PM95 method could not adequately homogenise some extremes. All such locations were coastal locations where the observation site moved from a highly exposed coastal site to a site further inland. The issue is most acute for summer maximum temperatures; on coasts with a strong land-sea temperature contrast, the coastal-inland temperature difference tends to increase with increasing temperature, before collapsing to near zero on the very hottest days as offshore winds override any marine influence (Fig. 14). The 95th or 99th percentiles are insufficient to resolve this behaviour at some locations (where extreme heat with strong offshore flow occurs on much less than 1% of days), leading in some cases to highly unrealistic adjustments for the most extreme values. Figure 20 shows an example of this at Albany – the differences between the airport and town sites for the 95th and 99th percentiles of summer maximum temperatures were between 6-8°C (Figure 15), and hence pre-1965 extreme high temperatures at the town were typically adjusted upwards by this amount, but on the very hottest days the difference between the two sites is in reality near zero (Fig. 14), resulting in a few very unrealistic values (e.g., a value from the town site of 44.8°C in 1933 was adjusted to a clearly unrealistic 52.5°C). The method produces a realistic time series for mean temperatures but a highly unrealistic one for extremes.

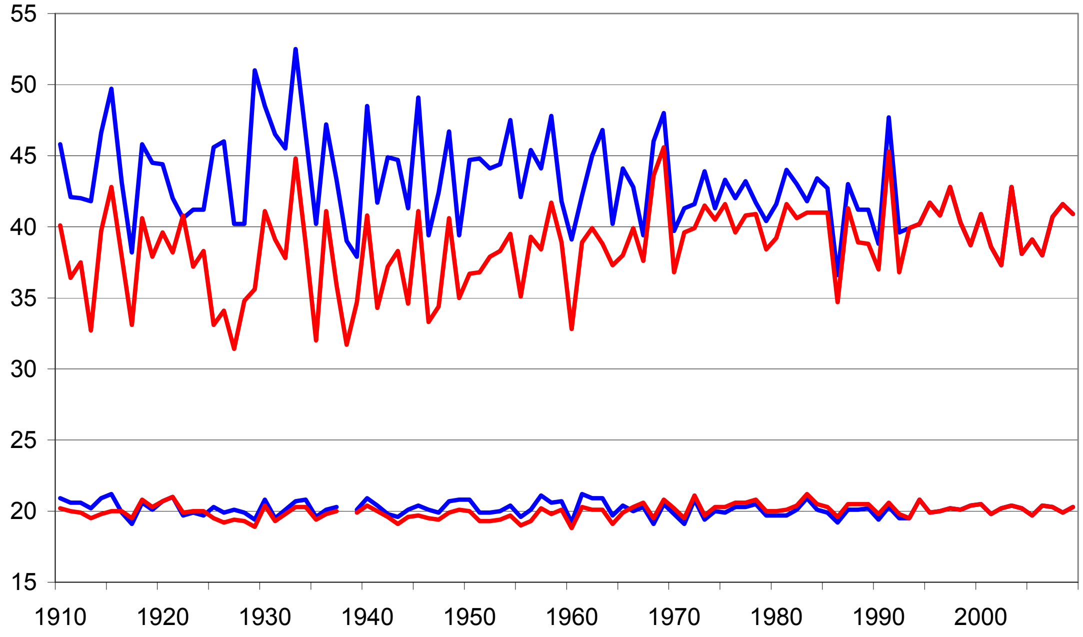


Fig. 20. Maximum temperatures (°C) at Albany for unadjusted (red) and homogeneity-adjusted (blue) data – mean annual (lower lines) and highest value in each year (upper lines).

To assess this problem at each location, a time series was developed for each year, for maximum and minimum temperature, of the following indicators:

- (Highest value in each year – 95th percentile value for each year)
- (Lowest value in each year – 5th percentile value for each year)

At locations where there was evidence of an inhomogeneity of at least 2°C in either or both of these time series, the location was considered to be one whose extremes were not homogenisable. These locations are excluded from downstream analyses that involve extremes, but are still considered reliable for analyses involving means, as only a very small proportion of values (less than 1%) are considered unreliable and they have little impact on means. Four locations were affected: Albany and Port Macquarie for extreme high maximum temperatures, Eucla for extreme high minimum temperatures, and Horn Island for extreme low minimum temperatures.

7.9 Corrections for data precision

As noted in section 3, there are some examples, notably from the early years of automatic weather stations, where data are rounded to the nearest degree. Such rounding will have no systematic effect on mean temperatures. It does, however, have the potential to affect the number of days above or below thresholds (Zhang et al., 2009), particularly in climates with low daily temperature variability: for example, if temperature is rounded to the nearest whole degree, a count of the number of days below 15°C will exclude days in which the temperature (to the nearest 0.1 degree) is between 14.5°C and 14.9°C inclusive (Fig. 21).

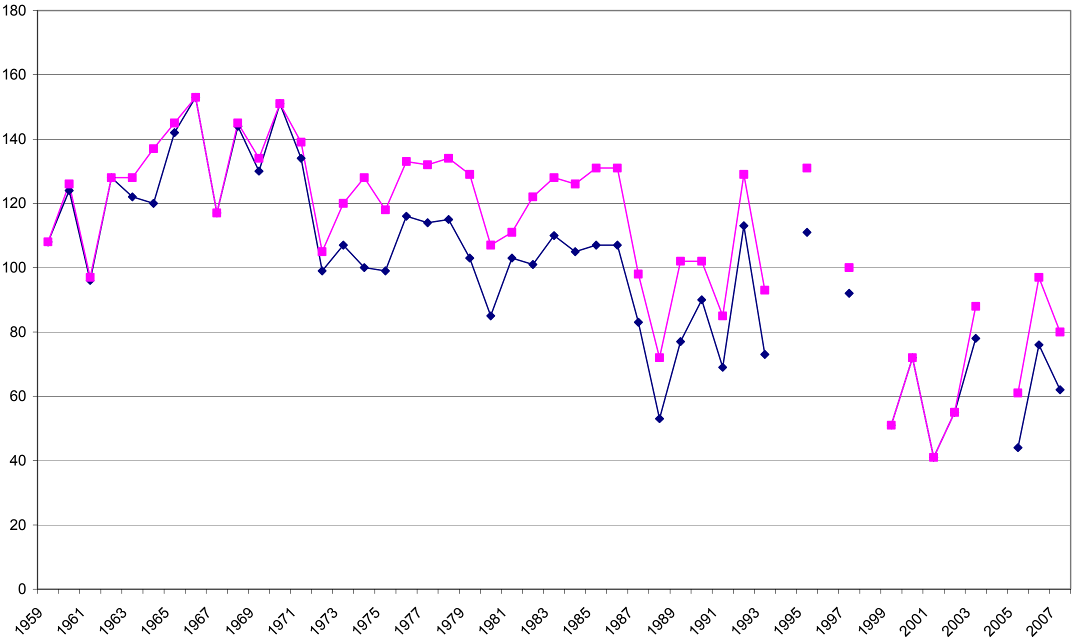


Fig. 21. Number of days per year with maximum temperature below 15.0°C at Eddystone Point (unadjusted data), for data reported to the nearest 0.1 degree (pink) and the nearest whole degree (blue).

This problem is less acute in homogenised data than it is in raw data, as the adjustment process involves a different adjustment for each of the 12 months, applying a level of randomisation to the final decimal digit of the adjusted data. However, it still has the potential to affect data to which no adjustments have been made (i.e. after the date of the last identified inhomogeneity, or at the very small number of locations with no identified inhomogeneities). To address this problem, a rounding-corrected version of the ACORN-SAT data set is to be prepared. Following the method of Zhang et al. (2009), it involves adding a random increment of between −0.5°C and +0.5°C to all data that have been rounded to the nearest degree (defined as location-months in which all values are in whole degrees).

8. POTENTIAL IMPACT OF NETWORK-WIDE CHANGES ON THE ACORN-SAT DATA SET

Changes which affect large parts of the network simultaneously provide a particular challenge for homogenisation, since they can affect reference series as well as a candidate station, causing assumptions of a locally homogeneous reference series around the time under examination to break down.

Three such changes have been identified in the period covered by the ACORN-SAT data set:

- Changes in observation time, principally the shift to an 0900-0900 observation day for maximum and minimum temperature in 1964, as well as the use of an 0000/1200 UTC observation day at some automatic weather stations in the 1990s and early 2000s, and the effective shift of observation time by one hour with the introduction of daylight saving time in some states from the early 1970s onwards.
- The change from imperial to metric measurements which took place on 1 September 1972.
- The introduction of automatic weather stations across large parts of the network from the early 1990s onwards.

8.1 Effect of observation time changes

The effect of observation time on temperatures has been the subject of considerable research attention in the United States of America. The United States is unusual in that historically there has been no fixed observation time at “co-operative” stations (the bulk of the network), with stations selecting the time which best suited them, and over time there has been a systematic shift from afternoon/evening to morning observations. This is one of the major sources of inhomogeneity in long-term American temperature records. A number of studies (e.g. Baker 1975, Karl et al. 1986, Vose et al. 2003) have found that a change from afternoon to early morning temperatures typically produces a shift in the order of −1°C in mean temperatures calculated using daily maximum and daily minimum temperature. This shift would be expected to be a function of inter-diurnal temperature variability, and thus would be smaller in climates where temperatures are less variable than is typical of much of the United States, especially in winter.

In some other countries observation time changes have been introduced on a national basis. Examples include:

- Canada (Vincent et al., 2009): a change from measuring minimum temperatures at 0600 UTC to 0000 UTC in 1961. This introduced a cold bias in minimum temperatures ranging from −0.2°C in western Canada to −0.8°C in eastern Canada.
- Norway (Nordli, 1997): a change from measuring minimum temperatures from 0800 to 1900 local time in 1938. This produced shifts of up to 1.5°C in some regions and seasons.

Previous Australian work in this field has been limited by a lack of high-resolution data. Shepherd (2000), using one-minute data for the 1995-1999 period at Hobart, that minimum temperatures for a 0900-0900 day averaged 0.26°C warmer than those for a midnight-midnight day, with the largest differences (0.3-0.5°C) in the warmer months, and little difference in the May-August period. Trewin (2001b) found differences in the vicinity of 0.3°C for minimum temperature for similar changes at Melbourne and Adelaide, but negligible impacts (less than

0.05°C) for a midnight-midnight day for maximum temperature, and for the use of daylight saving time for either maximum or minimum temperature.

To estimate the effect of observation time changes which have occurred over time, available one-minute data (mostly from the period 2003-2009) from 32 ACORN-SAT locations were used to calculate daily maximum and minimum temperatures for a range of time periods which have been used historically, as shown in Table 7. These were compared with temperatures measured using the current standard of 0900-0900 local time. These locations are reasonably well distributed across Australia, except for gaps in New South Wales (where most potential locations were ruled out because of excessive missing data). This study, and its results, are described in more detail in Trewin (2012).

The one-minute data indicate that the only historical observation practice which shows substantial systematic differences from the current standard is the measurement of minimum temperatures using a 0000-0000 day (i.e., midnight to midnight). Averaged across the 32 stations, this gives mean minimum temperatures 0.25°C cooler than the current standard, whilst the impact on extremes is stronger, with the mean value of the highest minimum temperature of each month being 0.58°C cooler on average. All 32 stations show cooler minimum temperatures for a 0000-0000 day than a 0900-0900 day, but the differences were smallest (typically near 0.1°C) in the tropics. They were largest (0.4-0.6°C) at some southern coastal stations (Fig. 22). As about 30% of the network was using the 0000-0000 day in some form prior to 1964, these results would suggest a potential inhomogeneity in Australian mean minimum temperatures of approximately +0.08°C in 1964.

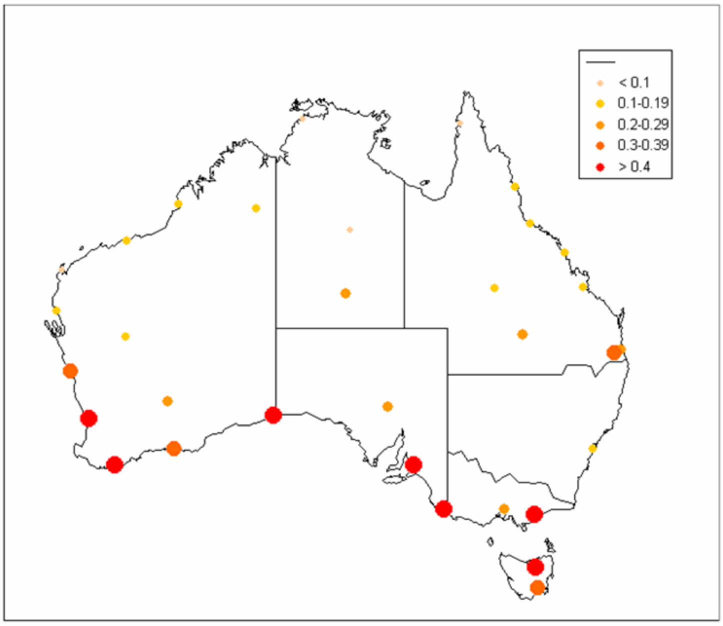


Fig. 22. Differences (°C) between mean minimum temperatures measured with an 0900-0900 day and those with a midnight-midnight day.

The effect on maximum temperatures of a midnight-midnight day is negligible on a national mean basis: the only location which shows a difference greater than 0.1°C is Eucla, which is unusually susceptible to 'reset' maxima for two reasons: its location on the Nullarbor coast means that it is susceptible to rapid temperature rises at any time of the day or night, especially in the warmer months, if winds shift to coming from the north and continental air displaces the marine layer, and its position in the far east of Western Australia means that 0900 standard time equates to about 0935 solar time on average, increasing the opportunity for temperatures to rise by the time of observation on the next day.

The impact of other changes in observation time practice, such as the introduction of daylight saving time or reversion to the 0900-1500 maximum if the 24-hour maximum is near the 0900 reset value, is negligible on an annual mean basis (although the latter practice does affect a small number of significant low maximum temperature extremes). It is also interesting to note that on a national mean basis the difference between 0900-0900 and 1500-0900 minimum temperatures is minimal, although it increases at the most southern locations, reaching around 0.15°C at Melbourne and Hobart.

The largest effects of observation time changes, as indicated by the results in Table 7, are likely to be on the frequencies of extreme high minimum temperatures and, to a lesser extent, extreme low maximum temperatures. This time-shift effect is particularly noticeable for extreme high minimum temperatures in southern coastal areas, where the hottest overnight temperatures typically occur on the last night before a cool change, which normally brings much lower temperatures by midnight than had occurred the previous night. This effect is illustrated by Fig. 23, which shows temperatures at Melbourne on 11-12 January 2010. The 'overnight' (1500-0900) minimum for 12 January was 30.6°C (the equal highest on record); the 0900-0900 minimum was 24.1°C (the reset temperature at 0900 on 11 January), the 1500-1500 minimum was 29.5°C, and the 0000-0000 minimum was 17.5°C (set at 2230 on 12 January).

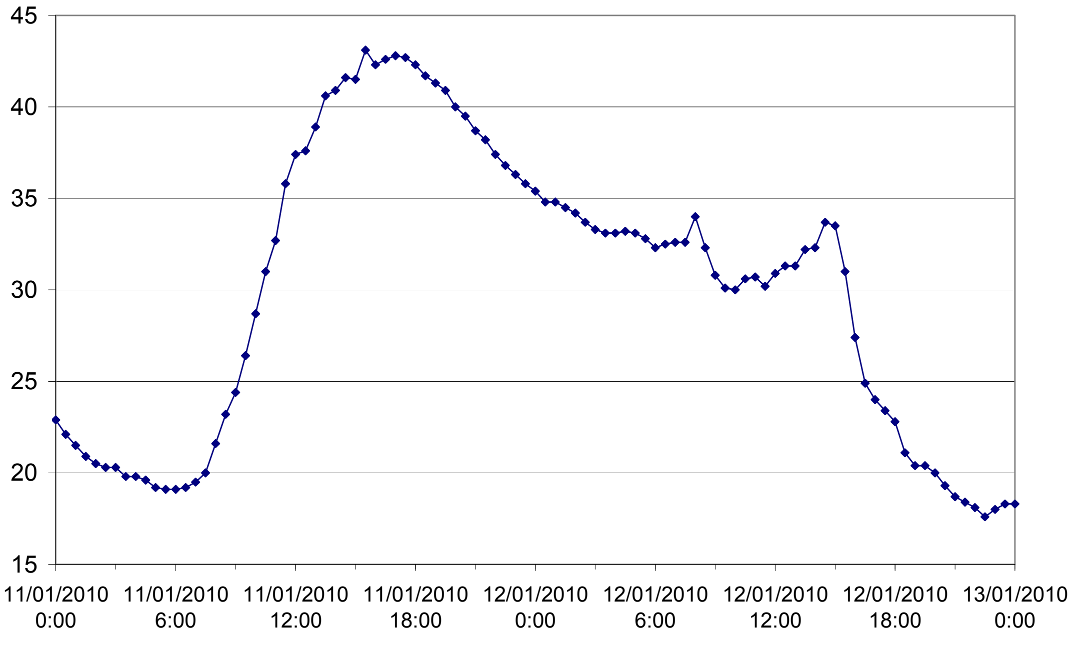


Fig. 23. Melbourne temperatures (°C), 11-12 January 2010

To illustrate the consequence of this effect for long-term time series of extremes, Melbourne has an average of 0.4 days per year with minimum temperatures of 25°C or above over the full 1910-2009 period, but had only one in the 32 years from 1932 to 1963 when the 0000-0000 observation was in use.

As a result of these findings, it was decided to define 1 January 1964 as a metadata-indicated potential inhomogeneity at all locations which were believed to use a 0000-0000 day prior to 1964, were outside the tropics and had not already had an adjustment made in the 1962-1966 period, and to adjust for that inhomogeneity, excluding stations subject to the 0000-0000 day as reference series, whether or not it reached the 0.3°C threshold defined as a minimum for adjustment in earlier sections (this takes into account the likelihood of extremes being affected more significantly than means).

This procedure was carried out at a total of 19 locations. However, at many of these locations, the inhomogeneity was much smaller than expected from the 2003-09 one-minute data (or even non-existent), suggesting that in practice the 0000-0000 observation day may not have been fully implemented at these locations in the earlier years. At the capital city sites, which had 24-hour staffing and where the fullest compliance with standards might be expected, differences were generally close to those expected. The results also suggest that, at the majority of the locations where the 0000-0000 day had been fully implemented, the resultant inhomogeneity had already been detected and adjusted for as part of the main ACORN-SAT procedures.

Table 7 Comparison of major observation times in previous use with current standard.

Method	In use	Temperature differences with current (0900-0900 local clock time) standard (°C)			
		Mean maximum	Mean minimum	Highest monthly minimum	Lowest monthly maximum
Maximum and minimum 0000-0000 (standard time)	1932-1963 at Bureau-staffed sites (and a few others)	−0.01	−0.25	−0.58	0.00
Maximum and minimum 0900-0900 standard time (i.e. no daylight saving)	1964-1972, later in states without daylight saving	0.00	−0.01	−0.05	−0.01
Maximum 1200-1200 UTC, minimum 0000-0000 UTC	Some AWSs from early 1990s to mid-2000s	0.02	−0.01 (−0.06 WA, 0.06 NSW/Vic/Tas)	−0.04 (−0.23 WA, 0.24 NSW/Vic/Tas)	0.08
Maximum 0900-0900 standard time (reverting to 0900-1500 if 0900 reset temperature within 0.5°C of maximum), minimum 1500-0900 standard time	Non-Bureau-staffed sites pre-1964	−0.03	0.03	0.13	−0.12

8.2 Effect of metrication

Metric measurements for temperature were introduced across the Australian network on 1 September 1972, with new instruments being issued to all stations. Previous research (Nicholls, 2004) had found no discernable impact of this change on temperatures. These conclusions, however, were partly based on the results of instrument testing for which no documentation could be located at the time of the current project. It was therefore decided to carry out a number of further tests using various data sets taken with different instruments to detect possible inhomogeneities with the change to the new instruments, these being:

- Comparison of Australian mean temperatures with sea surface temperatures in the Australian region.
- Comparison of mean temperatures at ACORN-SAT locations where upper-air observations were taken with the mean 850 hPa temperatures at those locations.
- Comparison of mean maximum and minimum temperatures at ACORN-SAT locations with the temperatures at 1500 and 0600 respectively (excluding locations/months where daylight saving time was in use).

Whilst the comparison data sets are not independent in the sense that they also changed to metric measurements at or near the same time, they used different instruments and hence would not necessarily be affected by any potential inhomogeneity which affected the maximum and minimum temperature observations..

All three comparisons showed mean Australian temperatures in the 1973-77 period were from 0.07 to 0.13°C warmer, relative to the reference series, than those in 1967-71. However, interpretation of these results is complicated by the fact that the temperature relationships involved (especially those between land and sea surface temperatures) are influenced by the El Niño-Southern Oscillation (ENSO), and the 1973-77 period was one of highly anomalous ENSO behaviour, with major La Niña events in 1973-74 and 1975-76. It was also the wettest five-year period on record for Australia, and 1973, 1974 and1975 were the three cloudiest years on record for Australia between 1957 and 2008 (Jovanovic et al., 2011).

The broad conclusion is that a breakpoint in the order of 0.1°C in Australian mean temperatures appears to exist in 1972, but that it cannot be determined with any certainty the extent to which this is attributable to metrication, as opposed to broader anomalies in the climate system in the years following the change. As a result no adjustment was carried out for this change.

8.3 Introduction of automatic weather stations

Automatic weather stations (AWSs) were introduced widely across the network from the early 1990s onwards. In many cases, their introduction coincided with a station move (often with a period of parallel observations). In other cases they were introduced without any site move. Often, once the automatic sensor became the primary instrument, the manual thermometers continued to be read and observations from them recorded, but in almost all cases those observations have not been digitised and are thus not available for a comparison study at this time.

A check was undertaken at those ACORN-SAT locations which had changed from manual to automatic observations without a documented site move (those where moves had taken place had already been dealt with in earlier parts of the process), using the adjustment procedure described earlier with only manual stations considered as reference stations. This showed a mean inhomogeneity with AWS introduction of −0.04°C for maximum temperature (positive at 8 locations, negative at 14) and −0.03°C for minimum temperature (positive at 13 locations, negative at 13). These results do not suggest any substantial inhomogeneity arising from the replacement of manual observations by automated observations *in situ*.

This conclusion is reinforced by results from two locations where parallel AWS and manual observations are available, Cape Byron (28.64°S, 153.64°E) (where the instruments were in the same screen) and Point Perpendicular (where the manual and automatic instruments were in different screens about 3 metres apart). In both cases the difference between manual and automatic temperatures during the overlap period matches, to within a few hundredths of a degree, the outcome of tolerance checks on the AWS temperature probe during the overlap period.

9. URBANISATION AND OTHER LOCAL LAND-USE CHANGE

It has been known for many years (e.g. Parker, 2010 and references therein) that urbanisation can have an impact on temperature, most commonly through an increase in overnight minimum temperatures. These impacts arise from a number of causes, including changes of land surface properties in the vicinity of observing sites (changing the partitioning between latent, sensible and radiative heat), emission of heat from buildings (especially overnight), and changes to local winds and radiation as a result of canyon effects from large buildings, or (in the case of winds) local mixing of air by moving vehicles or similar.

Some urbanisation effects on temperature will be manifested as a step change that is likely to be detected during the homogenisation process – for example, the construction of a building near the observation site is an event that occurs over a short period of time. However, the potential still exists for urbanisation to induce artificial warming trends relative to the surrounding region, and it is therefore necessary to identify such locations to prevent them from unduly influencing assessments of background climate change. It is conceivable that artificial trends, either positive or negative, could exist as a result of gradual changes in the surrounding environment unrelated to urbanisation, e.g. vegetation growth, but this is not considered explicitly in the ACORN-SAT data set.

Sites were initially classified into three categories:

- *Urban*: sites within the built-up area of a city or town with a population of more than 10,000.
- *Urban fringe*: sites associated with a city or town with a population of more than 10,000, that are either outside, but within 2 km of, the urban boundary, or are in a substantial parkland area or similar within the urban area, at least 50 metres from any buildings other than those directly associated with the observation site.
- *Non-urban*: sites not associated with a city or town with a population of more than 10,000, or associated with such a centre but more than 2 km from the urban boundary.

Of the 112 ACORN-SAT locations:

- Four are urban in the initial classification (all of them in large cities – Sydney, Melbourne, Adelaide and Hobart),

- 18 are currently urban fringe sites, five of which had been urban at some point in their history,
- 90 are currently non-urban sites, five of which had been urban at some point in their history.

The four locations that were classified as urban in the initial classification were included in the final list of urban locations. The 23 locations that were classified as urban fringe, or had formerly been classified as urban, were subjected to a statistical test of their minimum temperature trends, relative to the trends at surrounding non-urban locations (as per Della-Marta et al., 2004), over the period for which they were associated with a population centre of more than 10,000 (i.e., starting in the year when the population reached 10,000, and finishing either at the end of the data set, or, if applicable, when the site moved out of the urban area). If the location has a warming trend in minimum temperature that was significantly greater than the non-urban background trend then the location was included in the final list of locations classified as urban.

As a result of this analysis, four additional locations (Townsville, Rockhampton, Richmond (NSW) and Laverton) were found to have anomalously large minimum temperature trends and were classified as urban, bringing the total number of urban stations to eight. All of these locations, except Rockhampton, are at airport sites where major urban growth corridors are located outside the airport boundary.

Whilst it did not affect their assessment as urban sites, it is interesting to note that of the four city-centre sites (Sydney, Melbourne, Adelaide, Hobart), only Adelaide showed evidence of an anomalous minimum temperature trend over the 1910-2009 period. This suggests that either any urban influence on temperatures at these locations was already fully developed by 1910 or that anomalous urban warming was manifested as step changes, possibly associated with specific buildings or other developments in the vicinity, which were successfully removed in the homogenisation process. It is likely that the first explanation is the dominant influence at Sydney, and the second at Melbourne.

A full list of locations and their urban classification is shown in Table 8.

Some other land-use changes have also been found to affect local temperature observations (Trewin, 2010). One possible example of this in the ACORN-SAT data set is at Mildura, where there is no trend in maximum temperatures over the 1910-2009 period, compared with trends near 0.1°C/decade (or near 1°C in total) at other ACORN-SAT locations in the general region. The Mildura area saw rapid development in intensive irrigated agriculture in the period between the two World Wars (1918-1939). A closer analysis of trends at Mildura found that anomalous maximum temperature trends at Mildura were largely confined to the 1920-1949 period, over which maximum temperatures showed a trend of −0.23°C/decade, compared with a regional average of about −0.1°C/decade. A test of 1920-1949 data from Griffith (34.32°S, 146.07°E) (which was in the data set of Torok and Nicholls (1996), but excluded from ACORN-SAT and from Della-Marta et al. (2004) because of poor data quality post-1990), which has a similar history of irrigation development to Mildura, also found an anomalously large downward trend in maximum temperature at that location (−0.20°C/decade). This indicates a strong possibility that land-use change is a major contribution to pre-1950 maximum temperature trends at Mildura, with the magnitude of the difference (0.3-0.4°C over 30 years) comparable to that found in irrigated regions of India (Roy et al., 2007) and the north-central United States (Mahmood et al., 2006) . There is no evidence of anomalous trends at Mildura after 1950, by which time irrigated agriculture had reached approximately its current extent in the region.

Stations determined as being influenced by urbanisation or other local land-use change remain in the ACORN-SAT data set, but their designation allows them to be excluded from downstream products such as the calculation of national and regional temperature anomalies for the analysis of large scale climate change.

Table 8 Urban classification of ACORN-SAT locations.

Location	Current status	Former status (if more urbanised)	Years tested for anomalous trend	Final status	Other comments
Kalumburu	Non-urban			Non-urban	
Halls Creek	Non-urban			Non-urban	
Broome	Fringe		1996-2009	Non-urban	
Port Hedland	Fringe		1976-2009	Non-urban	
Marble Bar	Non-urban			Non-urban	
Learmonth	Non-urban			Non-urban	
Wittenoom	Non-urban			Non-urban	
Carnarvon	Non-urban			Non-urban	
Meekatharra	Non-urban			Non-urban	
Dalwallinu	Non-urban			Non-urban	
Geraldton	Fringe		1959-2009	Non-urban	Site moved outside town before pop reached 10K
Morawa	Non-urban			Non-urban	
Perth	Fringe	Urban (1910-62)	1910-2009	Non-urban	
Bridgetown	Non-urban			Non-urban	
Cape Leeuwin	Non-urban			Non-urban	Post-1965 site well outside town area
Albany	Non-urban	Urban (1960-65)	1950-1965	Non-urban	
Esperance	Non-urban			Non-urban	
Merredin	Non-urban			Non-urban	
Cunderdin	Non-urban			Non-urban	
Katanning	Non-urban			Non-urban	
Wandering	Non-urban			Non-urban	
Eucla	Non-urban			Non-urban	
Forrest	Non-urban			Non-urban	
Kalgoorlie	Fringe	Urban (1938-53)	1938-2009	Non-urban	
Giles	Non-urban			Non-urban	Population did not reach 10K until 1956
Darwin	Fringe		1956-2009	Non-urban	
Victoria River Downs	Non-urban			Non-urban	
Tennant Creek					
Alice Springs	Non-urban			Non-urban	Site well outside town
	Non-urban			Non-urban	
Rabbit Flat					
Woomera	Non-urban			Non-urban	
Tarcoola	Non-urban			Non-urban	
Marree	Non-urban			Non-urban	
Oodnadatta	Non-urban			Non-urban	
Ceduna	Non-urban			Non-urban	
Kyancutta	Non-urban			Non-urban	Post-1999 site well outside town
Port Lincoln	Non-urban			Non-urban	
	Non-urban	Urban (1975-99)	1975-1999	Non-urban	
Snowtown					
Cape Borda	Non-urban			Non-urban	
Adelaide	Non-urban			Non-urban	
	Urban		1910-2009	Urban	

Table 8 (cont.) Urban classification of ACORN-SAT locations.

Location	Current status	Former status (if more urbanised)	Years tested for anomalous trend	Final status	Other comments
Nuriootpa	Non-urban			Non-urban	
Mount Gambier	Non-urban			Non-urban	Site moved outside town before pop reached 10K
Robe	Non-urban			Non-urban	
Weipa	Non-urban			Non-urban	
Horn Island	Non-urban			Non-urban	
Palmerville	Non-urban			Non-urban	
Normanton	Non-urban			Non-urban	
Burketown	Non-urban			Non-urban	
Richmond (Qld)	Non-urban			Non-urban	
Georgetown	Non-urban			Non-urban	
Cairns	Fringe	Urban (1928-47)	1928-2009	Non-urban	
Townsville	Fringe		1942-2009	Urban	
Mackay	Fringe	Urban (1932-50)	1932-2009	Non-urban	
Charters Towers	Non-urban			Non-urban	
Barcaldine	Non-urban			Non-urban	
Longreach	Non-urban			Non-urban	
Camooweal	Non-urban			Non-urban	
Boulia	Non-urban			Non-urban	
Birdsville	Non-urban			Non-urban	
Gayndah	Non-urban			Non-urban	
Rockhampton	Fringe		1939-2009	Urban	
Bundaberg	Fringe	Urban (1924-90)	1924-2009	Non-urban	
Amberley	Non-urban			Non-urban	Site well clear of major towns
Cape Moreton	Non-urban			Non-urban	
Brisbane Airport	Fringe		1949-2009	Non-urban	
Miles	Non-urban			Non-urban	
St. George	Non-urban			Non-urban	
Charleville	Non-urban			Non-urban	
Thargomindah	Non-urban			Non-urban	
Tibooburra	Non-urban			Non-urban	
Wilcannia	Non-urban			Non-urban	
Cobar	Non-urban			Non-urban	
Bourke	Non-urban			Non-urban	
Walgett	Non-urban			Non-urban	
Moree	Non-urban			Non-urban	
Gunnedah	Non-urban			Non-urban	
Inverell	Non-urban			Non-urban	
Yamba	Non-urban			Non-urban	In fringe site pre-1983
Coffs Harbour	Fringe		1971-2009	Non-urban	
Port Macquarie	Fringe	Urban (1983-2003)	1972-2009	Non-urban	Site well clear of major towns
Williamtown	Non-urban			Non-urban	
Scone	Non-urban			Non-urban	

Table 8 (cont.) Urban classification of ACORN-SAT locations.

Location	Current status	Former status (if more urbanised)	Years tested for anomalous trend	Final status	Other comments
Bathurst	Fringe		1928-2009	Non-urban	
Dubbo	Non-urban	Urban (1948-99)	1948-1999	Non-urban	Post-1999 site well outside town
Sydney	Urban		1910-2009	Urban	
Richmond (NSW)	Fringe		1939-2009	Urban	
Nowra	Non-urban			Non-urban	Site well outside town
Point Perpendicular	Non-urban			Non-urban	
Moruya Heads	Non-urban			Non-urban	
Canberra	Fringe		1939-2009	Non-urban	
Wagga Wagga	Fringe	Urban (1928-50)	1928-2009	Non-urban	
Cabramurra	Non-urban			Non-urban	
Wyalong	Non-urban			Non-urban	
Deniliquin	Non-urban			Non-urban	Site moved outside town before pop reached 10K
Mildura	Non-urban			Non-urban	
Nhill	Non-urban			Non-urban	
Kerang	Non-urban			Non-urban	
Rutherglen	Non-urban			Non-urban	
Gabo Island	Non-urban			Non-urban	Site well outside town
Orbost	Non-urban			Non-urban	
Sale	Non-urban			Non-urban	
Wilsons Promontory	Non-urban			Non-urban	
Melbourne	Urban		1910-2009	Urban	
Laverton	Fringe		1945-2009	Urban	Post-1946 site well outside town
Cape Otway	Non-urban			Non-urban	
Low Head	Non-urban			Non-urban	
Launceston	Non-urban	Urban (1910-46)	1910-1946	Non-urban	
Eddystone Point	Non-urban			Non-urban	
Cape Bruny	Non-urban			Non-urban	
Hobart	Urban		1918-2009	Urban	
Grove	Non-urban			Non-urban	
Butlers Gorge	Non-urban			Non-urban	

10. CASE STUDIES OF SOME SPECIFIC INHOMOGENEITIES

In this section case studies of some inhomogeneities are presented. Most of the examples in this section are towards the large end of the range of adjustments made in the ACORN-SAT data set but are illustrative of some typical situations encountered.

10.1 Move from enclosed to open site within a town – Inverell, 1967

The observing site at Inverell moved about 100 metres on 20 September 1967, from the grounds of the post office to the grounds of the local library. Both sites were on flat ground within the central town area, but the post office site was very built-up with several buildings within a 10-metre radius; the library site was much more open.

The move from an enclosed to an open site resulted in large decreases in both maximum and minimum temperature in all seasons, with an estimated mean annual adjustment of −1.0°C for maximum temperature and −1.2°C for minimum temperature, with fairly similar adjustments in all seasons (Fig. 24(a)). There is a weak tendency towards larger differences in colder conditions, both for minimum and maximum temperatures, but a substantial difference remains even on warm days (Fig. 24(b)).

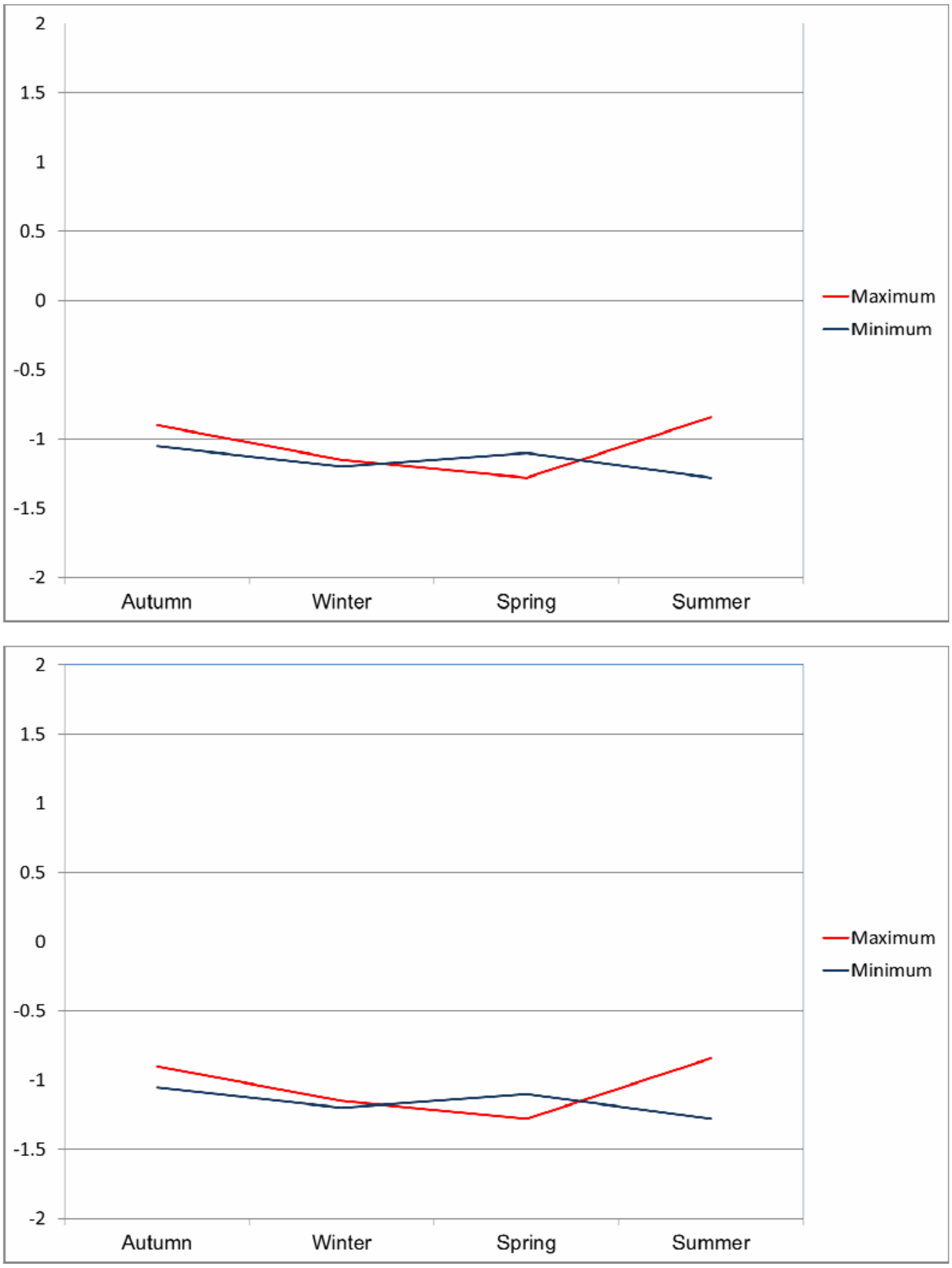


Fig. 24. (a, top) Size of adjustments (°C) for 1967 site move at Inverell, by season (b, bottom) Transfer function for 1967 site move at Inverell, for July maximum (red) and minimum (blue) temperatures.

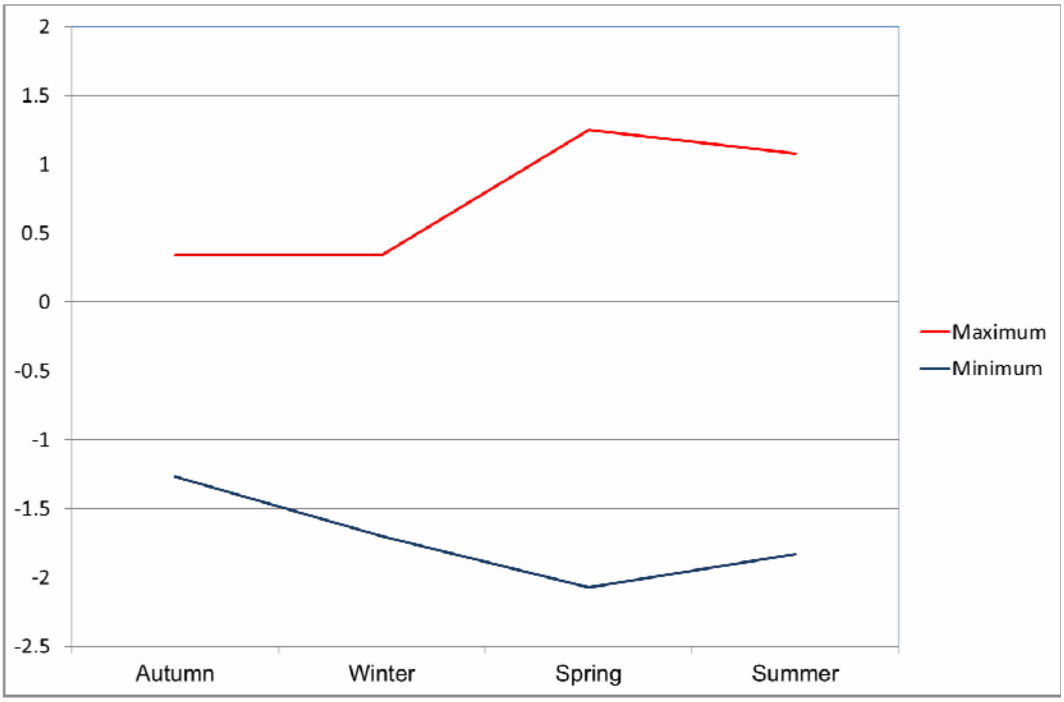
10.2 Move from near-coastal to inland site – Port Macquarie, 2000

The Port Macquarie site relocated from a location within the town to the airport in the late 1990s/early 2000s. The airport site opened in 1995 (although it was not used in the ACORN-SAT data set prior to 2000 because of issues with the data in its early years), while the town site continued to operate until 2003.

The town site was in the grounds of a suburban house, about 1.5 kilometres south of the town centre and 700 metres from the ocean (although it was arguably less exposed to the ocean than that distance suggests because of the presence of a ridge between the site and the shoreline). The airport site is on level ground about 5 kilometres to the west, and 5 kilometres from the ocean.

As indicated earlier in this report (Fig. 13), maximum temperatures at the airport are generally warmer than at the town site but with a marked seasonal cycle, with the difference small in winter and much larger during the summer months. Adjustments made for maximum temperature (Fig. 25(a)) reflect this seasonal cycle. The airport site is much cooler for minimum temperatures all year, with the largest differences in spring.

The transfer functions used (Fig. 25(b)) show a marked dependence of the adjustment on the location in the frequency distribution, both for maximum and minimum temperatures. For January maximum temperatures the adjustments are modest in the bottom half of the frequency distribution (below 26-27°C), but become much larger on hot days with an estimated adjustment of 3.5°C for days above the 95th percentile (approximately 30°C). It is worth noting here that, as discussed in section 7.8, extreme high maximum temperatures at Port Macquarie were found to be not homogenisable, suggesting that the difference between the two sites becomes erratic above the 95th percentile (most likely because the difference collapses on the very hottest days as offshore winds override the seabreeze, as at Albany). For July minimum temperatures the differences between the sites are modest on warm nights and progressively increase as nights get colder, a typical pattern for a coastal-inland site pair.



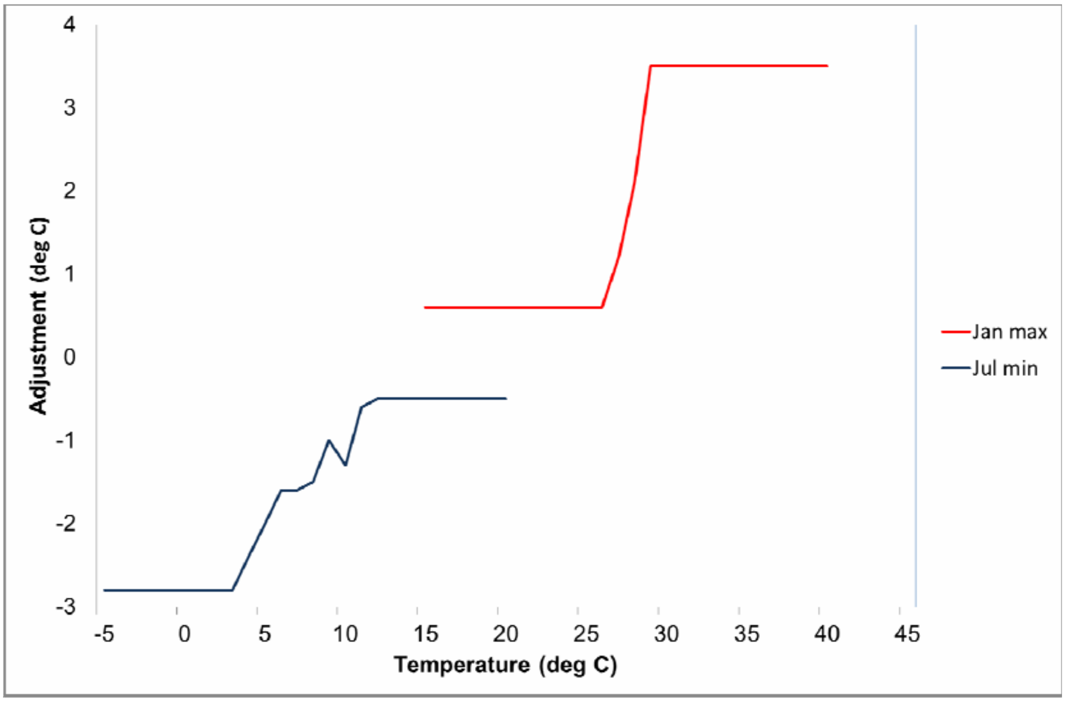


Fig. 25. (a, top) Size of adjustments (°C) for 2000 site move at Port Macquarie, by season (b, bottom) Transfer function for 2000 site move at Port Macquarie, for January maximum (red) and July minimum (blue) temperatures.

10.3 Change of land surface – Burketown, 2002

The principal observation site at Burketown moved from the town to the airport in late 2001, with the ACORN-SAT data set switching over at the start of 2002. The airport is approximately 2 kilometres southwest of the town. The town site was well clear of buildings. The principal difference between the two sites is that the town site was over a watered lawn with green grass all year, whereas the airport site is over natural vegetation which would normally be green during the wet season (summer and early autumn) but brown/yellow at other times of year, particularly in winter and spring.

The difference in maximum temperatures between the two sites was estimated at 0.6°C on an annual mean basis, but with a marked seasonal cycle, being near 1°C in winter and spring but only 0.1°C in summer (Fig. 26(a)). In winter the difference between the sites is fairly uniform except on the coolest days (Fig. 26(b)); in summer the airport site tends to be warmer on days with below-normal temperatures (which at that time of year would normally be cloudy) and the town site is generally warmer on days with above-normal temperatures (which would normally be sunny).

Victoria River Downs, where there was no site change but the grass around the site stopped being watered in 2007, shows a very similar pattern (differences near 1°C in the dry season but less than 0.2°C in the wet season), although no adjustment was implemented here in the 2009 version of the ACORN-SAT data set because the change was too close to the end of the data set. On the other hand, another tropical location where a site moved from a watered to an unwatered site (Boulia, 1999) shows a similar shift in mean annual maximum temperature but no seasonal cycle in the adjustments, possibly because Boulia is a semi-arid location and there is no time of year when the natural vegetation is reliably green.

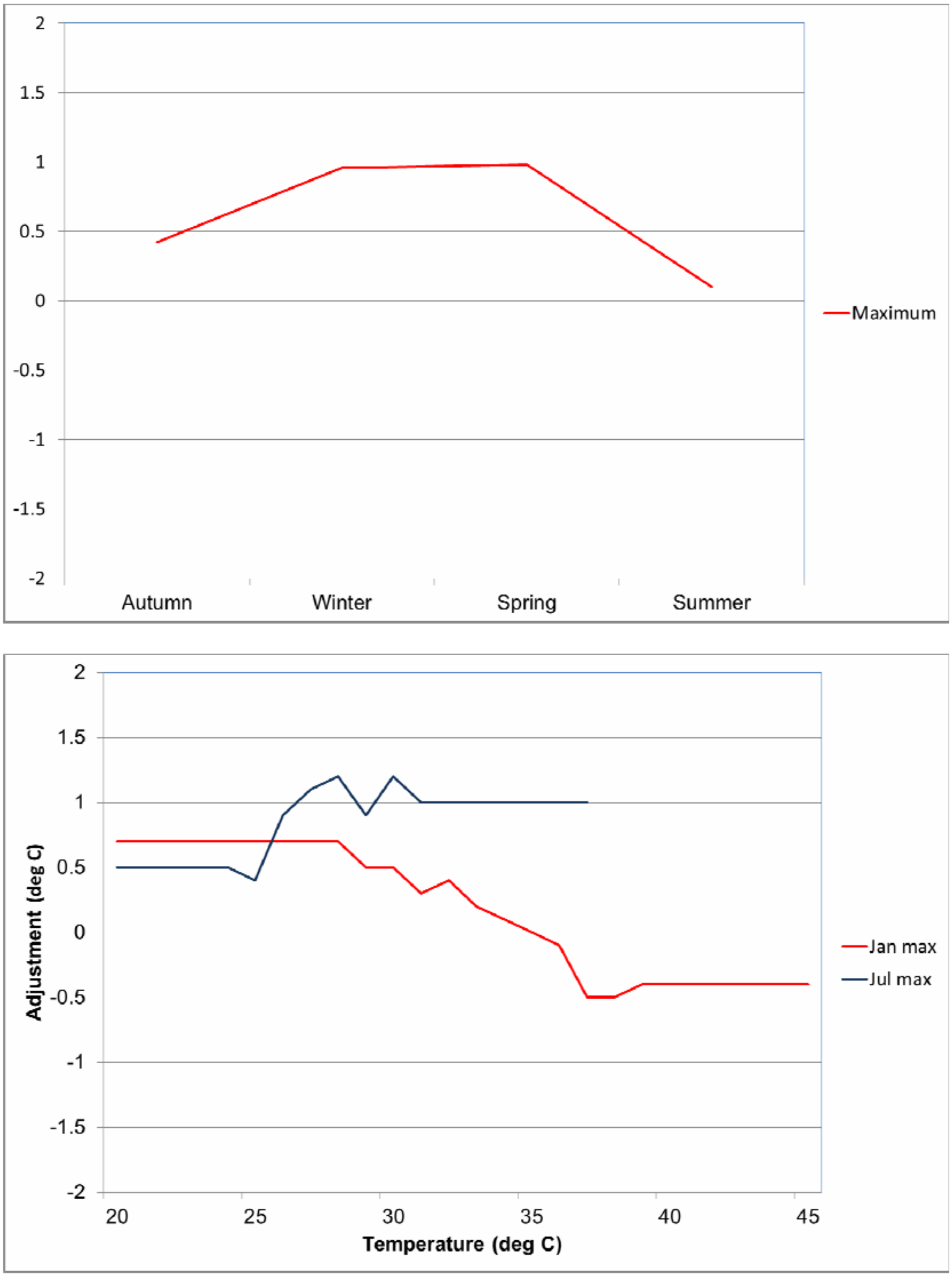


Fig. 26. (a, top) Size of adjustments (°C) for 2002 site move at Burketown, by season (b, bottom) Transfer function for 2002 site move at Burketown, for January (red) and July (blue) maximum temperatures.

10.4 Small town to rural move and land surface change – Snowtown, 1998

The observing site at Snowtown, a small town, moved from the town centre to a location about 2 kilometres north of the town in the late 1990s, with the new site opening in 1998 and the old site continuing until 2001. The old site was a typical small-town location with buildings nearby and a ground surface (lawn) which remained similar all year, whereas the new site is in cropland, with generally bare soil nearby in summer and early autumn and green crops during the winter/spring growing season.

The difference in annual mean maximum temperatures between the two sites is minimal (the new site is 0.1°C cooler), but this conceals a very marked seasonal cycle (Fig. 27(a)). In summer, when the new site is over bare soil, its maximum temperatures are substantially higher than those in the town, whereas in winter, when it is surrounded by green vegetation (and in a much less enclosed position than the old site), its maximum temperatures are substantially cooler. These differences apply throughout the frequency distribution (Fig. 27(b)), although there is some indication of a larger temperature difference on the hottest summer days. Minimum temperatures at the rural site are substantially lower than those at the town site all year, with a weak tendency towards larger differences on cooler nights.

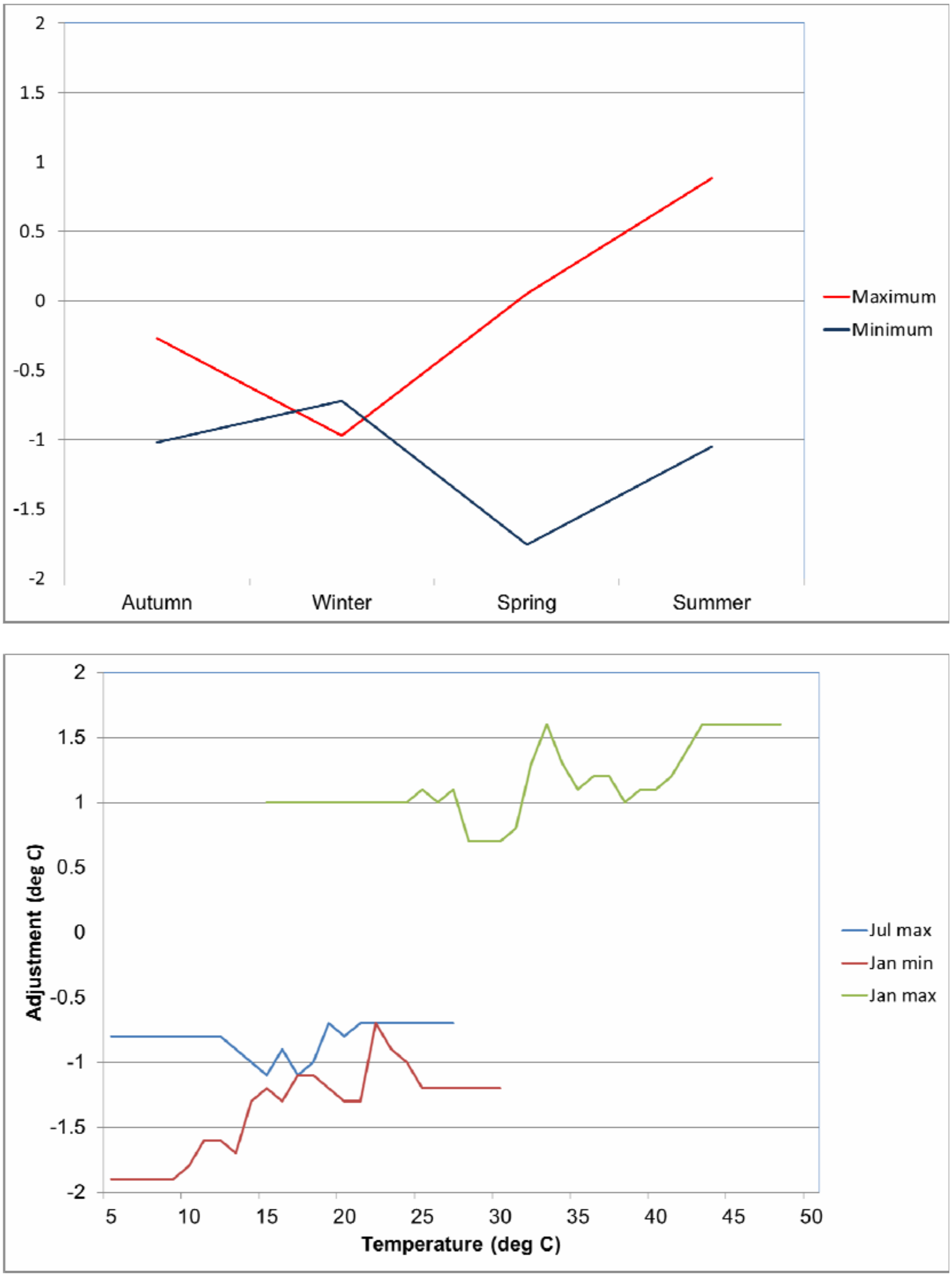


Fig. 27. (a, top) Size of adjustments (°C) for 1998 site move at Snowtown, by season (b, bottom) Transfer function for 1998 site move at Snowtown, for January maximum (green) and January (red) and July (blue) minimum temperatures.

10.5 New development nearby – Canberra, 2006

A bitumen car park was built within 30 metres of the observation site at Canberra Airport in May 2006. This resulted in an increase in minimum temperatures in all seasons, slightly more pronounced in the warmer months (Fig. 28(a)), with a mean annual adjustment of approximately 0.5°C.

There were seasonal differences in the way that different parts of the frequency distribution were affected (Fig. 28(b)); in winter the differences were greatest on colder nights, whereas in summer they tended to be greater on warm nights. A possible explanation for this behaviour could be the release of absorbed heat from the car park after a sunny day (which in Canberra's climate is typically associated with cold nights in winter, but not necessarily in summer), but the difference between clear and cloudy conditions has not been assessed.

The site moved away from the car park in 2008 (with overlapping data extending until 2010), essentially reversing the effect of the 2006 developments on minimum temperatures.

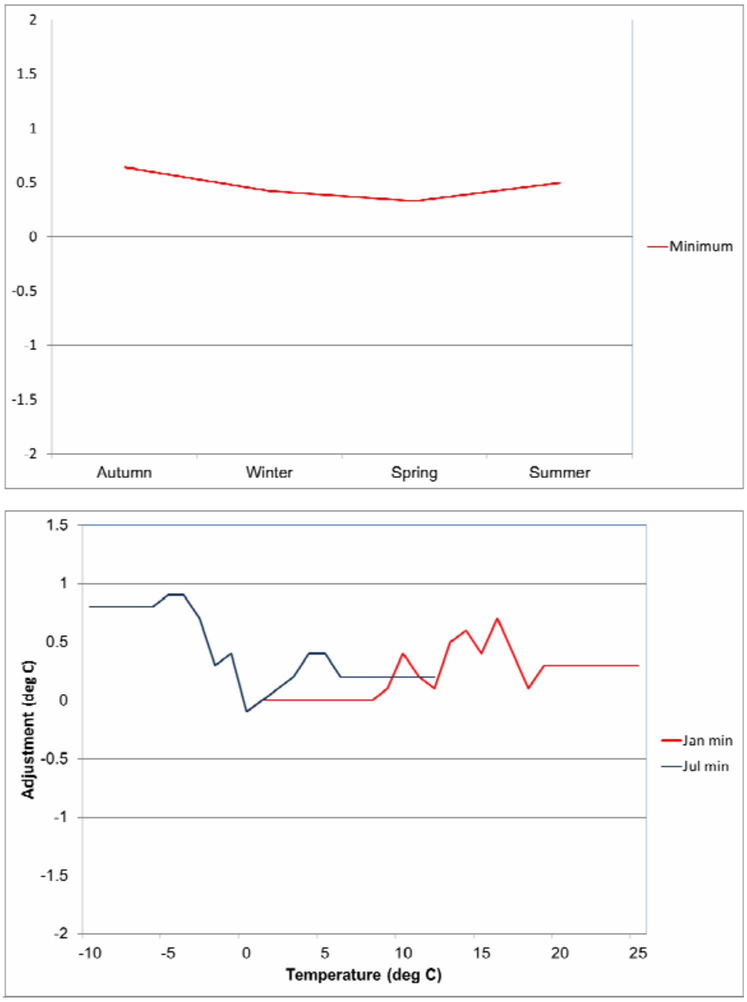


Fig. 28. (a, top) Size of adjustments (°C) for 2006 inhomogeneity at Canberra, by season (b, bottom) Transfer function for 2006 inhomogeneity at Canberra, for January (red) and July (blue) minimum temperatures.

10.6 Changes in surrounding vegetation – Alice Springs

The homogenisation procedure detected a large number of relatively small shifts, in both directions, in minimum temperature at Alice Springs in the post-1974 period (Fig. 29). There was no significant site move at Alice Springs after 1974, but the shifts matched multi-year rainfall variations quite closely, with a tendency for higher minimum temperatures in the years following very dry periods and lower minimum temperatures in the years following very wet periods.

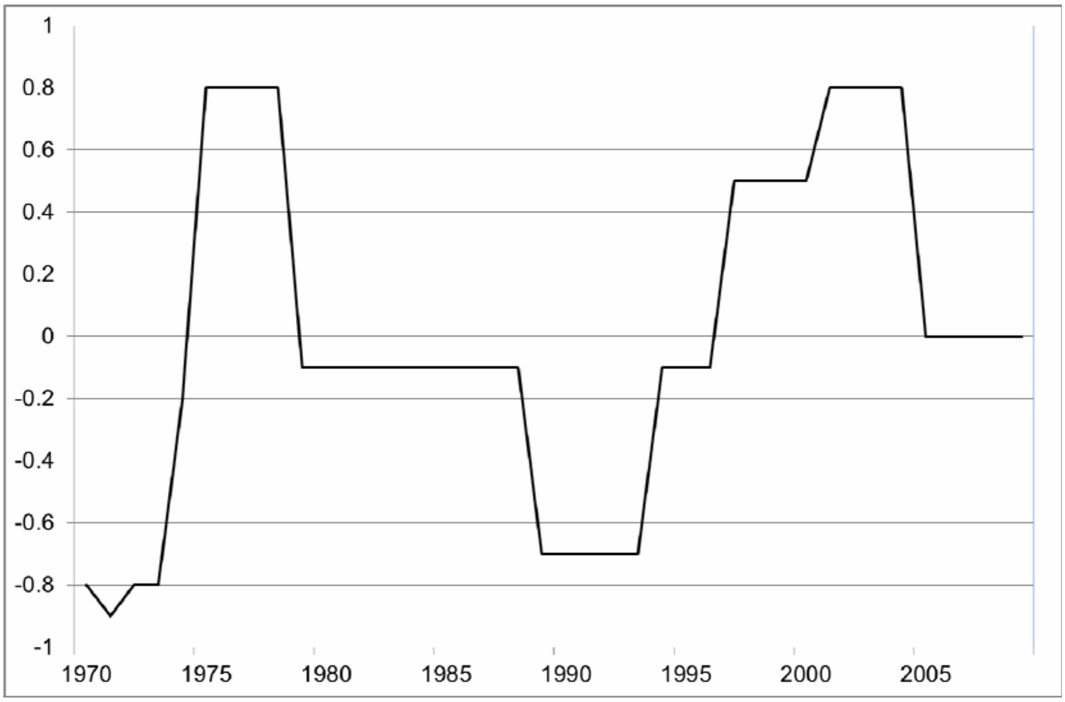


Fig. 29. Accumulated annual mean adjustments (°C) for minimum temperature at Alice Springs, relative to 2009.

Station photographs over the post-2000 period suggest that the height of vegetation (mostly grass) in the 100-200 metres around the immediate instrument enclosure varies considerably, being quite high during wet periods (e.g. immediately after the very wet 2000-2001 period) but almost non-existent after very dry years as occurred in the mid-2000s. A possible mechanism for this to affect minimum temperatures in the way observed would be for high grass to reduce screen-level wind speeds at the observation site, and therefore reduce mixing of the near-surface layer, reducing observed minimum temperatures (conversely, the absence of nearby vegetation would increase screen-level wind speeds and increase mixing of the near-surface layer).

It is interesting to note in this context that the treatment of the 1974 site move at Alice Springs is one of the more important systematic differences between the ACORN-SAT data set and that of Della-Marta et al. (2004). The Della-Marta et al. (2004) data set applies an adjustment of −1.5°C for the 1974 move and makes no subsequent adjustments, but the ACORN-SAT analysis suggests that only about half this adjustment is attributable to the site move, with the remainder being most likely attributable to vegetation growth arising from the extremely wet conditions which prevailed in central Australia in the 1973-1976 period, and subsequently being reversed when rainfall returned to more normal levels in the late 1970s.

11. CONCLUDING REMARKS

The ACORN-SAT data set provides national coverage of daily homogenised data for Australia for the period from 1910 to the present, although with a lower station density in the first half of the 20th century. This, in turn, will support century-scale analyses of changes in mean temperatures (which have been carried out in the past using existing data sets), as well as of extremes (which has not previously been possible for the period prior to about 1957).

The use of the PM algorithm for adjustment does not greatly alter the effectiveness of homogenisation for mean temperatures compared with methods based on uniform monthly or annual adjustment, but does considerably improve the representation of extremes.

The techniques used in the development of the ACORN-SAT data set are portable, in principle, to any region with a sufficient density of reference series. In cases where no reference series is available (e.g. remote islands), techniques such as RHtestsV3, which do not use reference series, are also available. The use of alternate elements, such as local sea surface temperatures for island locations (e.g. Jovanovic et al., 2010), as reference series is also a possibility at some locations.

The methods implemented for the ACORN-SAT data set were very labour-intensive, particularly for quality control. This presents an obstacle to such methods being scaled up to be applied to multi-national or global-scale data sets, along with the limited international availability of supporting data (for reference series) and metadata. In principle, however, it should be feasible to automate the homogenisation process, at least up to the point of producing an initial set of homogenised data.

12. ACKNOWLEDGMENTS

The assistance of Robert Fawcett, Branislava Jovanovic, Karl Braganza and Robert Smalley in carrying out analyses related to the data set or other supporting activities is gratefully acknowledged, as is the contribution of many other Bureau of Meteorology staff (particularly in the Observations Program) in supplying relevant information. Comments from David Jones, Michael Coughlan and Brad Murphy have been of great assistance in improving the original manuscript. Information on the irrigation history of Mildura was supplied by Neil Barr of the Victorian Department of Primary Industries. The Western Australian component of this research has been supported by the Western Australian Department of Environment and Conservation under the Indian Ocean Climate Initiative Stage 3.

13. REFERENCES

Alexandersson, H. 1986. A homogeneity test applied to precipitation data. *J. Climatol.*, 6, 661-675.

Allen, R.J. and DeGaetano, A.T. 2000. A method to adjust long-term temperature extreme series for nonclimatic inhomogeneities. *J. Clim.*, 13, 3680-3695.

Ashcroft, L., Gergis, J. and Karoly, D.J. 2011. Early meteorological observations in New South Wales. *Extreme Weather 2011: Joint Conference of the New Zealand Meteorological Society and the Australian Meteorological and Oceanographic Society, Wellington, New Zealand, 9-11 February 2011.*

Baker, C.B. and Helfert, M. 2008. U.S. Climate Reference Network (USCRN): a unique national long-term climate monitoring network. *12th Conference on IOAS-OALS, 88th AMS Annual Meeting, New Orleans, 20-24 January 2008.*

Baker, D.G. 1975. Effect of observation time on mean temperature estimation. *J. Appl. Met.*, 14, 471-476.

Begert, M., Schlegel, T. and Kirchhofer, W. 2005. Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *Int. J. Climatol.*, 25, 65-80.

Brandsma, T. and Können, G.P. 2006. Application of nearest-neighbour resampling for homogenizing temperature records on a daily to sub-daily level. *Int. J. Climatol.*, 26, 75-89.

Borhan, P., Kennedy, J.J., Harris, I., Tett, S.F.B. and Jones, P.D. 2006. Uncertainty estimates in regional and global observed temperatures changes: a new data set from 1850. *J. Geophys. Res.*, 111:D12106.

Brunet, M., Saladié, O., Jones, P., Sigró, J., Aguilar, E., Moberg, A., Walther, A., Lister, D., Lopez, D. and Almarza, C. 2006. The development of a new daily adjusted temperature dataset for Spain (SDATS) (1850-2003). *Int. J. Climatol.*, 26, 1777-1802.

Bureau of Meteorology. 1925. *Australian Meteorological Observer's Handbook*, Bureau of Meteorology, Melbourne, 172 pp.

Bureau of Meteorology. 1954. *Australian Meteorological Observer's Handbook*, Bureau of Meteorology, Melbourne, 148 pp.

Bureau of Meteorology. 1984. *Observing the weather: the Australian co-operative observers guide*. Australian Government Publishing Service, Canberra, 49 pp.

Bureau of Meteorology. 1995. *Australia's Reference Climate Stations*, Bureau of Meteorology, Melbourne, 4 pp.

Bureau of Meteorology. 1997. Guidelines for the siting and exposure of meteorological instruments and observing facilities. *Observations Specification No. 2013.1*, Bureau of Meteorology. Available online at http://www.bom.gov.au/inside/oeb/networks/20131.pdf.

Bureau of Meteorology. 2011. *Annual climate summary 2010.* Bureau of Meteorology, Melbourne, 20 pp. Available online at http://www.bom.gov.au/climate/annual_sum/2010/index.shtml.

Clarkson, N.M., Trewin, B.C., Jones, D.A., Plummer, N., Hutchinson, R.L. and Wong, K. 2001. Extending the computerised Australian climate archives to unlock our climate history – the CLIMARC project. *14th Australia-New Zealand Climate Forum, Darwin, 18-21 September 2001.*

Conrad, V. and Pollak, L.W. 1950. *Methods in climatology: 2nd edition*. Harvard University Press, Cambridge, 459 pp.

COST. 2009. *Action ES0601: advances in homogenisation methods of climate series: an integrated approach (HOME)*. http://www.homogenisation.org, viewed 1 October 2009.

DeGaetano, A.T. 2006. Attributes of several methods for detecting discontinuities in mean temperature series. *J. Clim.*, 19, 838-853.

Della-Marta, P., Collins, D. and Braganza, K. 2004. Updating Australia's high-quality annual temperature dataset. *Aust. Met. Mag.*, 53, 75-93.

Della-Marta, P.M. and Wanner, H. 2006. A method of homogenizing the extremes and mean of daily temperature measurements. *J. Clim.*, 19, 4179-4197.

Easterling, D.R. and Peterson, T.C. 1995. A new method for detecting and adjusting for undocumented discontinuities in climatological time series. *Int. J. Climatol.*, 15, 369-377.

Folland, C.K. and Salinger, M.J. 1995. Surface temperature trends and variations in New Zealand and the surrounding ocean, 1871-1993. *Int. J. Climatol.*, 15, 1195-1218.

Gergis, J., Karoly, D.J. and Allan, R.J. 2009. A climate reconstruction of Sydney Cove, New South Wales, using weather journal and documentary data, 1788-1791. *Aust. Met. Oceanogr. J.*, 58, 83-98.

Grose, M.R., Barnes-Keoghan, I., Corney, S.P., White, C.J., Holz, G.K., Bennett, J.B., Gaynor, S.M. and Bindoff, N.L. 2010. *Climate Futures for Tasmania: technical report, General Climate Impacts*. Antarctic Climate and Ecosystem Cooperative Research Centre, Hobart, 72 pp.

Guttman, N.B. and Baker, C.B. 1996. Exploratory analysis of the difference between temperature observations recorded by ASOS and conventional methods. *Bull. Amer. Met. Soc.*, 77, 2865-2873.

Hansen, J., Ruedy, R., Glascoe, J. and Sato, M. 1999. GISS analysis of surface temperature change. *J. Geophys. Res.*, 104, 30997-31022.

Hansen, J., Ruedy, R., Sato, M., Imhoff, M., Lawrence, W., Easterling, D., Peterson, T. and Karl, T. 2001. A closer look at U.S. and global surface temperature change. *J. Geophys. Res.*, 106, 23947-23963.

Hanssen-Bauer, I. and Nordli, P.Ø. 1998. Annual and seasonal temperature variations in Norway, 1876-1997. *KLIMA Report 25/98*, Norwegian Meteorological Institute, Oslo.

Jones, D.A. 1999. Characteristics of Australian land surface temperature variability. *Theor. Appl. Climatol.*, 63, 11-31.

Jones, D.A. and Trewin, B.C. 2002. On the adequacy of digitised historical Australian daily temperature data for climate monitoring. *Aust. Met. Mag.*, 51, 237-250.

Jones, D.A., Wang, W. and Fawcett, R. 2009. High-quality spatial climate data sets for Australia. *Aust. Met. Oceanogr. J.*, 58, 233-248.

Jones, P.D., Raper, S.C.B., Bradley, R.S., Diaz, H.F., Kelly, P.M. and Wigley, T.M.L. 1986. Northern Hemisphere surface air temperature variations: 1851-1984. *J. Climate Appl. Meteor.*, 25, 161-179.

Jones, P.D. and Wigley, T.M.L. 2010. Estimation of global temperature trends: what's important and what isn't. *Climatic Change*, 100, 59-69.

Jovanovic, B., Jones, D.A. and Collins, D. 2008. A high quality monthly pan evaporation dataset for Australia. *Clim. Change*, 87, 517-535.

Jovanovic, B., Braganza, K., Collins, D. and Jones, D. 2010. High-quality monthly climate data for Australia's Antarctic and remote island weather stations. *Australia-New Zealand Climate Forum 2010, Hobart, 13-15 October 2010.*

Jovanovic, B., Collins, D., Braganza, K., Collins, D. and Jones, D.A. 2011. A high quality monthly total cloud amount dataset for Australia. *Clim. Change*, 108, 485-517.

Karl, T.R., Williams, C.N., Young, P.J. and Wendland, W.M. 1986. A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States. *J. Clim. Appl. Met.*, 25, 145-160.

Klein Tank, A.M.G. and coauthors. 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.*, 22, 1441-1453.

Mahmood, R., Foster, S.A., Keeling, T., Carlson, C. and Leeper, R. 2006. Impacts of irrigation on 20th-century temperatures in the Northern Great Plains. *Glob. Planet. Change*, 54, 1-18.

Menne, M.J. and Williams, C.N. 2009. Homogenization of temperature series via pairwise comparisons. *J. Clim.*, 22, 1700-1717.

Nicholls, N., Tapp, R., Burrows, K. and Richards, D. 1996. Historical thermometer exposures in Australia. *Int. J. Climatol.*, 16, 705-710.

Nicholls, N. 2004. Did metrication bias trends in Australian temperature? *Bull. Aust. Met. Ocean. Soc.*, 17, 102-103.

Nordli, P.Ø. 1997. Adjustments of Norwegian monthly means of daily minimum temperature. *KLIMA Report 6/97*, Norwegian Meteorological Institute, Oslo.

Parker, D.E. 1994. Effects of changing exposure of thermometers at land stations. *Int. J. Climatol.*, 14, 1-31.

Parker, D.E. 2010. Urban heat island effects on estimates of observed climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 1, 123-133.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E.J., Hanssen-Bauer, I., Alexandersson, H., Jones, P. and Parker, D. 1998. Homogeneity adjustments of *in situ* atmospheric climate data: a review. *Int. J. Climatol.*, 18, 1493-1517.

Reeves, J., Chen, J., Wang, X.L., Lund, R. and Lu, Q. 2007. A review and comparison of changepoint detection techniques for climate data. *J. Appl. Met. Climatol.*, 46, 900-915.

Roy, S.S., Mahmood, R., Nivogi, D., Lei, M., Foster, S.A., Hubbard, K.G., Douglas, E. and Pielke, R.A. snr. 2007. Impacts of the agricultural Green Revolution-induced land use changes on air temperatures in India. *J. Geophys. Res.*, 112, D21108.

Shepherd, D.J. 2000. Minimum temperature time. *Bull. Aust. Met. Oceanogr. Soc.*, 13, 104-109.

Smith, T.M., Reynolds, R.W., Peterson, T.C. and Lawrimore, J. 2008. Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *J. Clim.* 21, 2283-2286.

Thorne, P.W., Willett, K.M., Allan, R.J., Bojinski, S., Christy, J.R., Fox, N., Gilbert, S., Jolliffe, I., Kennedy, J.J., Kent, E., Klein Tank, A., Lawrimore, J., Parker, D.E., Rayner, N., Simmons, A., Song, L., Stott, P.A. and Trewin, B.C. 2011. Guiding the creation of a comprehensive surface temperature resource for 21st century climate science. *Bull, Amer. Met. Soc.*, published online 11 July 2011.

Torok, S.J. 1996. *The development of a high quality historical temperature data base for Australia*. Ph.D. thesis, School of Earth Sciences, University of Melbourne.

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

Trewin, B.C. and Trevitt, A.C.F. 1996. The development of composite temperature records. *Int. J. Climatol.*, 16, 1227-1242.

Trewin, B.C. 2001a. The development of a high-quality daily temperature data set for Australia. *11th Symposium on Meteorological Observations and Instrumentation, Albuquerque, 14-18 January 2001.*

Trewin, B.C. 2001b. *Extreme temperature events in Australia.* Ph.D. thesis, School of Earth Sciences, University of Melbourne.

Trewin, B.C. 2010. Exposure, instrumentation and observing practice effects on land temperature measurements. *Wiley Interdisciplinary Reviews: Climate Change*, 1, 490-506.

Trewin, B.C., Wang, X.L., Jones, P.D., Menne, M., Lund, R. and Machin, G. 2010. Creation of quality controlled homogenised datasets from the databank. *Surface Temperatures Workshop, Exeter, United Kingdom, 7-9 September 2010.*

Trewin, B.C. 2012. Influence of time of observation on temperatures in Australia. *2012 AMOS National Conference, Sydney, 31 January-3 February 2011.*

Venema, V.K.C. and coauthors. 2012. Benchmarking homogenization algorithms for monthly data. *Clim. Past*, 8, 89-115.

Vincent, L.A., Zhang, X., Bonsal, B.R. and Hogg, W.D. 2002. Homogenization of daily temperatures over Canada. *J. Clim.*, 15, 1322-1334.

Vincent, L.A., Milewska, E.J., Hopkinson, R. and Malone, L. 2009. Bias in minimum temperature introduced by a redefinition of the climatological day at the Canadian synoptic stations. *J. Appl. Met. Climatol.*, 48, 2160-2168.

Vose, R.S., Williams, C.N., Peterson, T.C., Karl, T.R. and Easterling, D.R. 2003. An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophys. Res. Lett.*, 30, 2046.

Vose, R.S. and Menne, M.J. 2004. A method to determine station density requirements for climate observing networks. *J. Clim.*, 17, 2961-2971.

Wang, X.L., Chen, H., Wu, Y., Feng, Y. and Pu, Q. 2010. New techniques for detection and adjustment of shifts in daily precipitation time series. *J. Appl. Met. Climatol.*, 49, 2416-2436.

Wijngaard, J.B., Klein Tank, A.M.G. and Können, G.P. 2003. Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.*, 23, 679-692.

World Meteorological Organization. 2011. *Guide to Climatological Practices, 2011 edition*. WMO No. 100, World Meteorological Organization, Geneva.

Zhang, X., Zwiers, F.W. and Hegerl, G. 2009. The influences of data precision on the calculation of temperature percentile indices. *Int. J. Climatol.*, 29, 321-327.

The Centre for Australian Weather and
Climate Research is a partnership between
CSIRO and the Bureau of Meteorology.