Aust. Met. Mag. 53 (2004) 75-93

Updating Australia's high-quality annual temperature dataset

Paul Della-Marta and Dean Collins

National Climate Centre, Bureau of Meteorology, Australia
and

Karl Braganza

CSIRO Atmospheric Research, Aspendale, Australia

(Manuscript received June 2003; revised September 2003)

An updated and improved version of the Australian high-quality annual mean temperature dataset of Torok and Nicholls (1996) has been produced. This was achieved by undertaking a thorough post-1993 homogeneity assessment using a number of objective and semi-objective techniques, by matching closed records onto continuing records, and by adding some shorter duration records in data-sparse regions. Each record has been re-assessed for quality on the basis of recent metadata, resulting in many records being rejected from the dataset. In addition, records have been re-examined for possible urban contamination using some new approaches. This update has highlighted the need for accurate and complete station metadata. It has also demonstrated the value of at least two years of overlapping observations for major site changes to ensure the homogeneity of the climate record. A total of 133 good-quality, homogenised records have been produced. A non-urban subset of 99 stations provides reliable calculations of Australia's annual mean temperature anomalies with observation error variances between 15 and 25 per cent of the total variance and decorrelation length scales greater than the average inter-station separation.

Introduction

Torok and Nicholls (1996) produced a homogenised or 'high-quality' dataset of annual mean maximum and minimum temperature series for Australia. The primary purpose of this dataset was to enable the reliable monitoring of climate trends and variability at annual and decadal time-scales. Consequently, each station record was corrected for discontinuities caused by changes in site location and exposure, and other known data problems (Peterson et al. 1998). Such discontinuities can be as large, or larger than,

---

Corresponding author address: D. Collins, National Climate Centre, Bureau of Meteorology, GPO Box 1289K, Melbourne, Vic. 3001, Australia.
Email: D.Collins@bom.gov.au

real temperature changes (e.g. the approximate 1°C drop in maximum temperatures associated with the switch to Stevenson screen exposure (Nicholls et al. 1996)) and consequently confound the true long-term trend. Temperature records for 224 observation locations were homogenised, 170 of which were classified as non-urban i.e. free from the influence of urban warming. Generally the high-quality records were homogenised from 1910, by which time most stations are believed to have been equipped with the current standard instrument shelter. The dataset ended in 1993.

At the end of each year Australia's annual mean temperature anomaly is calculated using all available high-quality, non-urban temperature records. This is routinely provided to the public via an annual climate summary, a media statement and the Bureau of Meteorology's internet site (e.g. http://www.bom.gov.au/announcements/media_releases/climate/change/20030106.shtml). The dataset has also been used extensively in the calculation of atmospheric indicators for Australia's 2001 *State of the Environment* report (Collins and Della-Marta 2002). In order for such analyses to remain credible, the climate records used need to be routinely checked for homogeneity.

Since 1993, the high-quality dataset has been updated using data that have undergone basic quality assurance checks but no checks for homogeneity. Many stations are likely to have undergone changes requiring data adjustment in this time. Also, the station exposure and quality of observations may have changed since the initial assessment of Torok and Nicholls (1996). Until this update, the number of high-quality stations still operating has gradually declined due to site moves and closures, resulting in gaps in the network. For the 2002 annual summary only 138 stations had a full year of data available; 99 of these classified as non-urban. The intention of this update was to assess the quality and homogeneity of open stations, to re-open 'closed' records using nearby open stations and to fill data voids if possible.

Most of the homogeneity techniques employed by Torok and Nicholls (1996) have been replicated. However, a number of new methodologies have also been used to assess whether a record is likely to have been contaminated by urban warming, and to utilise available comparison observation data in cases where a station has moved. The development of web-based technology and greater digitisation of metadata since the production of the original dataset have allowed summary information, relevant analyses and metadata to be conveniently combined into a webpage for each individual station. This has enabled the efficient assessment of each record.

Assessment of station quality

Torok and Nicholls (1996) subjectively ranked the quality of records on a scale of 1 to 6, with rank 1 being very good, 5 being very poor and 6 being urban. This update re-assessed each station, mainly on the basis of post-1993 metadata, but also using metadata compiled by Torok (1996). All stations were ranked on the original 1 to 5 scale, but instead of a rank 6, all stations were noted as either urban or non-urban, enabling urban stations to also be given a quality rank. Ranks were subjectively assigned, taking into account the current and historical standard of site exposure, number of site moves, number of instrument changes, recent standard of observations and amount of missing data.

Since the development of the initial high-quality dataset there have been changes in the way station history information is recorded by the Australian Bureau of Meteorology. Traditionally station metadata have been recorded on paper history files. With the development of SitesDB (Lane 1995), an electronic metadata database, the process of assessing station quality has been streamlined. However, at present, the information in the database is only reliable since the mid to late 1990s. Also, there is currently great variation in the quality of information held in SitesDB, reflecting different levels of use and appreciation of needs throughout the Bureau.

The introduction of SitesDB has resulted in paper history files generally not being used routinely. If SitesDB is not supported adequately this will result (and possibly already has resulted) in the loss of important information. Also, most communication about observation stations is now via email, rather than formal inter-office memoranda. Such communications often contain useful information for future assessments of the climate record, but are generally not placed on file. A concerted effort is currently underway within the Bureau of Meteorology to populate SitesDB with historical metadata from paper history files and to archive station-specific email communication.

After this latest assessment, sixteen of the original non-urban stations were considered to have better quality than the assessment of Torok and Nicholls (1996). This was usually due to the station having moved to a location with better exposure. However, 99 stations were ranked with lower quality. Predominantly this has occurred at manual stations, possibly due to diminishing resources being used to maintain the cooperative network as reliance on the automatic weather station network increases. This has resulted in documented, but uncorrected, exposure problems, less interest from cooperative observers and the downgrading of manual observation programs, such as the cessation of weekend observations.

Some of the apparent decline in quality is simply due to more stringent assessment criteria during the latest assessment. Some records contain many years in which annual mean values were estimated by Torok and Nicholls (1996) using a reference series based on highly correlated neighbour series. Often this amounted to more than twenty contiguous years of estimated data. The amount of in-filling appears to have had no bearing on the original station rank. However, for this assessment it was decided that substantial periods of estimated data could be misleading to potential users. Generally, any records with around twenty or more years of estimated annual mean values were ranked as poor or very poor in the updated dataset (see Table 1).

Stations assessed as poor or very poor were not removed from the original dataset in order to maximise the number of long-term records available (Torok 1996). Jones and Trewin (2000, 2002) have since shown that significant interstation correlations of monthly (daily) temperature anomalies can occur for separations up to at least 700 (360) km, and up to 1200 (570) km in some months. It would be reasonable to expect that significant interstation correlations on the annual time-scale would be equivalent or greater than those for monthly temperature anomalies since the spatial variability of the annual time-scale is lower.

The monthly decorrelation length scales calculated by Jones and Trewin (2000) suggest that, in most of Australia, the number of high-quality records available is more than adequate for calculating a reliable annual mean temperature anomaly. This redundancy allowed the removal of all stations ranked 4 (poor) or 5 (very poor) from the updated dataset. The adequacy of the updated network density is examined later.

Most of the removed stations were located in southeast Australia and were ranked poorly due to large data gaps in their record or poor site exposure. The reasons for removing each poor quality record are described in Table 1. Four stations in New South Wales (Condobolin, Coonamble, Trangie and White Cliffs) were removed from the dataset due to high maximum temperatures in the early part of their record. These data were considered suspect because some metadata indicated that the instrument shelters used were not standard Stevenson screens until the 1940s, which would bias maximum temperatures higher (Torok 1996). Despite intending to resurrect as many high-quality records as possible, a total of 81 poor-quality records were rejected from the updated dataset in order to produce higher quality overall.

Assessment of recent homogeneity

Torok (1996) provides a brief metadata history up to 1993 for each high-quality station. These histories were updated for this assessment using the metadata contained in SitesDB and hardcopy station files to identify potential dates of new discontinuities. Torok and Nicholls (1996) and Torok (1996) also used a number of objective statistical tests to help determine dates of discontinuity. Using the metadata where possible, decisions were made as to which discontinuities would be corrected to homogenise a record. Torok (1996) provides a record of the dates at which homogeneity adjustments were made, the magnitude and sign of adjustments, the reasons for adjustments and which tests were used to identify them.

In an attempt to validate previous adjustments it was decided to reconstruct the unhomogenised composite candidate series using raw station data for each high-quality station. The raw dataset was created by merging data from the Australian Data Archive for Meteorology (ADAM) and data compiled and digitised by Torok (1996). These data were occasionally slightly different from the unhomogenised dataset used by Torok and Nicholls (1996), which was sourced from an early version of ADAM. In most cases the candidate records were a composite of two or more records with different station numbers, the combination being chosen to achieve the highest number of contiguous years of record.

Objective statistical test

The objective test of Torok and Nicholls (1996), based on a two-phase regression model (Easterling and Peterson 1995), was replicated as closely as possible for the update assessment. The objective test relies on the creation of a homogeneous reference series to create a difference series (candidate series minus reference series). Two methods of creating a reference series were compared to assess the reliability of the objective test results. One used the median of interannual temperature difference of highly correlated neighbouring stations (hereafter known as the median reference series) (Torok and Nicholls 1996; Torok 1996) and the other was created using a distance weighted mean of highly correlated neighbouring station anomalies (hereafter known as the distance weighted reference series) (Trewin 2001).

For each method an initial search of stations within six degrees of latitude and longitude and an elevation difference of no more than 300 m was performed to find suitable reference stations. These limits were relaxed to included stations within eight degrees latitude and longitude and 500 m elevation difference in data-sparse areas if the initial search yielded fewer than

Table 1. Stations included in the updated annual temperature dataset. Quality rank and whether or not a station is considered urban are also shown. (Rank 1 = very good, Rank 2 = good, Rank 3 = fair, Rank 4 = poor, Rank 5 = very poor.) Note that stations ranked 4 or 5 have been removed from the high-quality dataset and are not used for the calculation of Australian mean temperatures. The reason(s) for the removal of these stations is shown. MD = missing data (e.g. more than 20 contiguous missing years, or only weekday or part-year observations during parts of the record); PE = poor exposure (e.g. screen in poor condition, partly shaded or set in bitumen during parts of the record); PO = poor quality or no overlapping data for a recent site move; PD = poor quality or suspect observations during parts of the record, and; NS = suspected to have a non-standard screen until the 1940s.

Station name	*Station no.*	*Latitude (°S)*	*Longitude (°E)*	*Elevation (m)*	*Start year*	*Quality Rank*	*Reason*	*Urban (Y/N)*
ALBANY AIRPORT	009741	-34.94	117.80	0068.0	1910	3		N
ALICE SPRINGS AIRPORT	015590	-23.80	133.89	0546.0	1910	3		N
AMBERLEY	040004	-27.63	152.71	0027.0	1913	2		N
ARARAT PRISON	089085	-37.28	142.98	0295.0	1910	3		N
ARMIDALE	056037	-30.53	151.67	0987.0	1910	4	PE	Y
ATHERTON SHIRE COUNCIL	031193	-17.26	145.48	0752.0	1910	3		N
AYR DPI RESEARCH	033002	-19.62	147.38	0012.0	1910	3		N
BAIRNSDALE AIRPORT	085279	-37.88	147.57	0049.4	1910	3		Y
BALLADONIA	011017	-32.46	123.86	0148.0	1910	5	MD	N
BALLARAT AIRPORT	089002	-37.51	143.79	0435.2	1910	2		Y
BALRANALD	049002	-34.64	143.56	0061.0	1910	5	MD, PE	N
BARCALDINE POST OFFICE	036007	-23.55	145.29	0266.9	1913	2		N
BAROON POCKET DAM	040850	-26.72	152.87	0230.0	1913	5	PO, MD	N
BATHURST AG. STATION	063005	-33.43	149.56	0713.0	1910	2		Y
BEECHWORTH WOOLSHED	082137	-36.32	146.68	0330.0	1910	5	MD	N
BEGA	069139	-36.67	149.82	0041.0	1910	4	PE, MD	N
BENALLA	082002	-36.55	145.97	0169.5	1910	3		Y
BOLLON	044010	-28.03	147.48	0182.9	1910	3		N
BOMBALA POST OFFICE	070005	-36.91	149.24	0705.0	1912	5	MD	N
BOULIA AIRPORT	038003	-22.91	139.90	0161.8	1910	3		N
BOURKE AIRPORT	048245	-30.04	145.95	0107.3	1910	2		N
BOWEN AIRPORT	033257	-20.02	148.22	0005.0	1910	3		N
BOWRAL	068102	-34.49	150.40	0690.0	1910	3		Y
BREWARRINA HOSPITAL	048015	-29.96	146.87	0115.0	1911	5	MD	N
BRIDGETOWN POST OFFICE	009510	-33.96	116.14	0149.9	1910	2		N
BRISBANE AIRPORT	040842	-27.39	153.13	0004.0	1910	3		Y
BROKEN HILL	047007	-31.98	141.47	0315.0	1910	3		Y
BROOME AIRPORT	003003	-17.95	122.23	0007.0	1910	2		N
BUNDABERG AIRPORT	039128	-24.89	152.32	0027.0	1910	5	MD	N
BURKETOWN POST OFFICE	029004	-17.74	139.55	0005.5	1910	5	MD, PE	N
BURRINJUCK DAM	073007	-35.00	148.60	0390.0	1912	4	MD, PE	N
CAIRNS AIRPORT	031011	-16.87	145.75	0003.0	1910	2		N
CAPE BORDA	022801	-35.75	136.59	0143.0	1910	3		N
CAPE BRUNY	094198	-43.49	147.14	0059.7	1924	2		N
CAPE LEEUWIN	009518	-34.37	115.13	0013.0	1910	1		N
CAPE MORETON LIGHTHOUSE	040043	-27.03	153.47	0099.9	1913	2		N
CAPE NATURALISTE	009519	-33.54	115.02	0109.0	1910	1		N
CAPE NORTHUMBERLAND	026005	-38.06	140.67	0005.0	1910	2		N
CAPE OTWAY LIGHTHOUSE	090015	-38.86	143.51	0082.0	1910	3		N
CARDWELL	032004	-18.26	146.02	0005.7	1910	3		N
CARNARVON AIRPORT	006011	-24.89	113.67	0004.0	1910	2		N
CASHMERE DOWNS	012022	-28.97	119.57	0450.0	1910	4	MD	N
CASHMORE AIRPORT	090171	-38.32	141.47	0080.9	1910	2		N
CASINO AIRPORT	058063	-28.88	153.05	0026.0	1910	5	MD	N
CASTLEMAINE PRISON	088110	-37.08	144.24	0330.0	1910	5	MD	N
CEDUNA	018012	-32.13	133.71	0015.3	1910	3		N
CHARLEVILLE AIRPORT	044021	-26.41	146.26	0302.6	1910	2		N
CHARTERS TOWERS AIRPORT	034084	-20.04	146.27	0290.0	1910	2		N
CLARE HIGH SCHOOL	021131	-33.82	138.59	0395.0	1910	5	PD	N
CLERMONT POST OFFICE	035019	-22.83	147.64	0267.0	1910	5	PE	N

Table 1. Continued.

Station name	*Station no.*	*Latitude (°S)*	*Longitude (°E)*	*Elevation (m)*	*Start year*	*Quality Rank*	*Reason*	*Urban (Y/N)*
CLONCURRY AIRPORT	029141	-20.67	140.51	0186.0	1910	5	MD	N
COBAR	048027	-31.49	145.83	0260.0	1910	3		N
COEN AIRPORT	027006	-13.76	143.11	0160.5	1910	4	PD	N
COLLARENEBRI	048031	-29.55	148.59	0145.0	1910	5	MD	N
CONDOBOLIN RESEARCH	050052	-33.07	147.23	0195.0	1910	4	NS	N
COOKTOWN AIRPORT	031017	-15.45	145.19	0005.5	1910	2		N
COOMA VISITORS CENTRE	070278	-36.23	149.12	0778.0	1910	4	PD	N
COONABARABRAN	064008	-31.27	149.27	0505.0	1910	4	PE	N
COONAMBLE COMPARISON	051010	-30.98	148.38	0180.0	1910	5	PD, NS	N
COOTAMUNDRA AIRPORT	073142	-34.63	148.04	0335.0	1910	5	PE	N
COROWA AIRPORT	074034	-35.99	146.36	0143.0	1910	5	MD	N
COWEAL TOWNSHIP	037010	-19.92	138.12	0231.2	1910	3		N
COWRA AIRPORT	065091	-33.85	148.65	0300.0	1910	3		N
CROYDON TOWNSHIP	029012	-18.20	142.24	0116.3	1912	5	MD	N
CUNNAMULLA POST OFFICE	044026	-28.07	145.68	0188.7	1910	3		N
DALBY AIRPORT	041522	-27.17	151.27	0345.0	1910	3		Y
DARWIN AIRPORT	014015	-12.42	130.89	0030.4	1910	3		Y
DENILIQUIN POST OFFICE	074128	-35.55	144.95	0093.0	1910	2		N
DERBY AIRPORT	003032	-17.37	123.66	0006.2	1910	3		N
DONNYBROOK	009534	-33.57	115.83	0063.0	1910	4	PE	N
DUBBO AIRPORT	065070	-32.22	148.58	0284.0	1910	3		Y
EAST SALE AIRPORT	085072	-38.11	147.13	0004.6	1910	2		N
ECHUCA AIRPORT	080015	-36.17	144.76	0096.0	1910	3		Y
EMERALD AIRPORT	035264	-23.57	148.18	0187.9	1910	5	PE	N
ESPERANCE	009789	-33.83	121.89	0025.0	1910	1		N
EYRE	011019	-32.25	126.30	0005.6	1910	5	MD	N
FORBES AIRPORT	065103	-33.36	147.92	0230.4	1910	4	PE	Y
GABO ISLAND LIGHTHOUSE	084016	-37.57	149.91	0015.2	1910	2		N
GATTON	040082	-27.55	152.34	0094.0	1913	4	MD	N
GEELONG AIRPORT	087163	-38.23	144.33	0033.4	1910	3		Y
GEORGETOWN POST OFFICE	030018	-18.29	143.55	0291.7	1910	3		N
GERALDTON AIRPORT	008051	-28.80	114.70	0033.0	1910	2		N
GILES METEOROLOGICAL OFFICE	013017	-25.04	128.29	0598.0	1957	1		N
GLADSTONE RADAR	039123	-23.86	151.26	0074.5	1910	2		N
GLEN INNES AIRPORT	056243	-29.68	151.69	1044.3	1910	3		N
GOONDIWINDI AIRPORT	041521	-28.52	150.33	0217.6	1910	3		N
GOOSEBERRY HILL	009204	-31.94	116.05	0220.0	1910	5	PO, MD	Y
GOULBURN	070263	-34.72	149.74	0650.0	1910	4	MD, PD	Y
GRAFTON OLYMPIC POOL	058130	-29.68	152.93	0009.0	1910	3		Y
GRENFELL	073014	-33.90	148.17	0410.0	1910	5	MD	N
GRIFFITH AIRPORT	075041	-34.25	146.07	0134.0	1916	5	PD, PO	Y
GUNNEDAH POOL	055023	-30.98	150.25	0306.0	1910	3		Y
GUYRA HOSPITAL	056229	-30.21	151.68	1332.0	1911	5	MD	N
GYMPIE	040093	-26.18	152.64	0064.5	1910	4	MD	Y
HALLS CREEK AIRPORT	002012	-18.23	127.66	0422.0	1910	2		N
HAMILTON AIRPORT	090173	-37.65	142.06	0241.1	1910	3		Y
HAY	075031	-34.52	144.85	0093.3	1910	3		N
HILLSTON AIRPORT	075032	-33.49	145.52	0122.0	1910	5	MD	N
HOBART	094029	-42.89	147.33	0050.5	1910	3		Y
HUGHENDEN AIRPORT	030022	-20.82	144.23	0315.5	1910	3		N
INNISFAIL	032025	-17.53	146.03	0008.0	1910	3		Y
INVERELL	056242	-29.78	151.11	0582.0	1910	2		Y
ISISFORD POST OFFICE	036026	-24.26	144.44	0205.4	1913	4	MD	N
JARRAHWOOD	009842	-33.80	115.66	0130.0	1910	2		N
JERRYS PLAINS POST OFFICE	061086	-32.50	150.91	0090.0	1910	3		N
KALGOORLIE-BOULDER AIRPORT	012038	-30.78	121.45	0365.3	1910	2		N
KATANNING COMPARISON	010579	-33.69	117.56	0310.0	1910	3		N

Table 1. Continued.

Station name	*Station no.*	*Latitude (°S)*	*Longitude (°E)*	*Elevation (m)*	*Start year*	*Quality Rank*	*Reason*	*Urban (Y/N)*
KATOOMBA	063039	-33.71	150.30	1030.0	1910	4	PE	Y
KELLERBERRIN	010073	-31.62	117.72	0250.0	1910	2		N
KEMPSEY	059017	-31.08	152.82	0010.0	1910	4	MD, PE	Y
KENT TOWN	023090	-34.92	138.62	0048.0	1910	2		Y
KERANG	080023	-35.73	143.92	0077.7	1910	2		N
KING ISLAND AIRPORT	098017	-39.88	143.88	0036.2	1914	3		N
KINGSCOTE AIRPORT	022841	-35.71	137.52	0005.0	1914	4	PE	N
LAKE CARGELLIGO AIRPORT	075039	-33.28	146.37	0169.0	1910	5	PE, MD	N
LAMEROO	025509	-35.33	140.52	0099.0	1915	5	PE	N
LAUNCESTON AIRPORT	091104	-41.54	147.20	0170.0	1910	2		N
LAVERTON	087031	-37.86	144.75	0016.0	1910	2		N
LAVERTON AIRPORT	012305	-28.62	122.42	0464.8	1910	4	PD, MD	N
LEONORA	012046	-28.88	121.33	0376.0	1910	5	MD	N
LISMORE	058037	-28.81	153.26	0011.0	1910	2		Y
LITHGOW	063224	-33.49	150.15	0950.0	1913	4	PE, MD	Y
LONGERENONG	079028	-36.67	142.30	0091.0	1910	5	MD	N
LONGREACH AIRPORT	036031	-23.44	144.28	0192.2	1910	1		N
LOW HEAD	091293	-41.06	146.79	0003.0	1910	2		N
MACKAY	033119	-21.12	149.22	0030.2	1910	2		Y
MARBLE BAR COMPARISON	004020	-21.18	119.75	0182.3	1910	3		N
MARYBOROUGH QUEENSLAND	040126	-25.52	152.71	0011.0	1910	2		Y
MARYBOROUGH VICTORIA	088043	-37.06	143.73	0249.3	1910	3		Y
MEEKATHARRA AIRPORT	007045	-26.61	118.54	0517.0	1910	2		N
MELBOURNE REGIONAL OFFICE	086071	-37.81	144.97	0031.2	1910	3		Y
MENINDEE POST OFFICE	047019	-32.39	142.42	0061.0	1910	5	PE, PD	N
MERREDIN	010092	-31.48	118.28	0315.0	1913	2		N
MILDURA AIRPORT	076031	-34.23	142.08	0050.0	1910	2		N
MILES	042112	-26.66	150.18	0304.8	1910	2		N
MITCHELL POST OFFICE	043020	-26.49	147.98	0336.3	1910	4	MD, PE	N
MOORA	008091	-30.64	116.01	0203.0	1910	5	MD	N
MOREE AIRPORT	053115	-29.49	149.85	0213.0	1910	2		Y
MORUYA HEADS	069018	-35.91	150.15	0017.0	1910	1		N
MOUNT BARKER S.A.	023733	-35.06	138.85	0360.0	1910	4	PE, PD	Y
MOUNT BARKER W.A.	009581	-34.63	117.64	0300.0	1910	5	MD, PE	N
MOUNT GAMBIER AIRPORT	026021	-37.75	140.77	0063.0	1910	1		N
MT GELLIBRAND	090035	-38.24	143.79	0261.0	1910	4	PO	Y
MUDGEE AIRPORT	062101	-32.56	149.62	0471.0	1910	3		N
MUNGINDI POST OFFICE	052020	-28.98	148.99	0160.0	1915	4	MD	N
MURRURUNDI POST OFFICE	061051	-31.76	150.84	0466.0	1910	5	MD, PE	N
MUSGRAVE	028007	-14.78	143.50	0080.1	1910	5	MD	N
NARRABRI WEST POST OFFICE	053030	-30.34	149.76	0212.0	1910	5	PE	N
NARROGIN	010614	-32.93	117.18	0338.0	1913	5	MD	N
NEWCASTLE	061055	-32.92	151.80	0033.0	1910	2		Y
NEWMAN AIRPORT	007176	-23.42	119.80	0524.0	1916	2		N
NHILL	078031	-36.34	141.64	0133.0	1910	3		N
NORMANTON AIRPORT	029063	-17.69	141.07	0018.4	1910	1		N
NORTHAM	010111	-31.64	116.67	0155.0	1910	4	PD	N
NYANG STATION	006072	-23.03	115.04	0111.0	1910	5	MD	N
OMEO	083025	-37.10	147.60	0685.0	1910	2		N
ONSLOW	005016	-21.64	115.11	0004.0	1910	4	PE	N
ORANGE AIRPORT COMPARISON	063231	-33.38	149.12	0948.0	1910	3		Y
PADTHAWAY SOUTH	026100	-36.65	140.52	0035.0	1910	4	MD	N
PALMERVILLE	028004	-16.00	144.08	0203.8	1910	3		N
PARKES	065026	-33.14	148.16	0324.0	1910	4	PE	Y
PARRAMATTA NORTH	066124	-33.79	151.02	0055.0	1910	4	PE, MD	Y
PERTH AIRPORT	009021	-31.93	115.98	0015.4	1910	1		Y
POINT PERPENDICULAR LIGHT.	068034	-35.09	150.80	0085.0	1910	3		N

Table 1. Continued.

Station name	*Station no.*	*Latitude (°S)*	*Longitude (°E)*	*Elevation (m)*	*Start year*	*Quality Rank*	*Reason*	*Urban (Y/N)*
POLKEMMETT	079023	-36.65	142.10	0141.0	1910	4	PD	Y
PORT HEDLAND AIRPORT	004032	-20.37	118.63	0006.4	1913	2		N
PORT LINCOLN	018192	-34.60	135.88	0008.5	1910	3		Y
PORT MACQUARIE	060026	-31.44	152.92	0007.0	1910	2		Y
QUIRINDI POST OFFICE	055049	-31.51	150.68	0390.0	1910	4	PE, MD	N
RAYVILLE PARK	021133	-33.77	138.22	0109.1	1910	3		N
RICHMOND POST OFFICE QLD.	030045	-20.73	143.14	0211.1	1910	1		N
RICHMOND RAAF N.S.W.	067105	-33.60	150.78	0019.0	1910	2		N
ROBE	026026	-37.16	139.76	0003.3	1910	2		N
ROCKHAMPTON AIRPORT	039083	-23.38	150.48	0010.0	1910	2		N
ROEBOURNE	004035	-20.78	117.15	0012.0	1910	2		N
ROMA AIRPORT	043091	-26.55	148.78	0303.2	1910	3		N
ROSEWORTHY	023122	-34.51	138.68	0065.0	1910	5	MD	N
ROTTNEST ISLAND	009193	-32.01	115.50	0043.1	1910	3		N
RUTHERGLEN RESEARCH	082039	-36.11	146.51	0167.6	1910	1		N
SANDY CAPE LIGHTHOUSE	039085	-24.73	153.21	0099.1	1910	2		N
SCONE	061089	-32.06	150.93	0216.0	1910	4	MD	N
SOUTHERN CROSS	012074	-31.23	119.33	0355.0	1910	3		N
STAWELL AIRPORT	079105	-37.07	142.74	0227.9	1910	4	PD, PE	Y
STRAHAN AIRPORT	097072	-42.16	145.29	0020.0	1910	3		N
STRATHALBYN RACECOURSE	024580	-35.29	138.89	0062.0	1910	3		N
STREAKY BAY	018079	-32.80	134.21	0013.0	1926	5	PE	N
SWAN HILL AIRPORT	077094	-35.38	143.54	0071.0	1910	1		N
SYDNEY	066062	-33.86	151.21	0039.0	1910	3		Y
TAMWORTH AIRPORT	055325	-31.07	150.84	0394.9	1910	3		Y
TARCOOLA AIRPORT	016098	-30.71	134.58	0123.0	1922	3		N
TAREE	060030	-31.90	152.48	0005.0	1910	4	PE	Y
TE KOWAI	033047	-21.16	149.12	0013.7	1910	5	MD	N
TENNANT CREEK AIRPORT	015135	-19.64	134.18	0375.7	1911	2		N
TENTERFIELD	056032	-29.05	152.02	0838.0	1910	3		N
TIBOOBURRA POST OFFICE	046037	-29.44	142.01	0183.0	1910	3		N
TRANGIE RESEARCH STATION	051049	-31.99	147.95	0215.0	1913	5	MD, NS	N
WAGGA WAGGA	072150	-35.16	147.46	0212.0	1910	2		N
WALGETT AIRPORT	052088	-30.04	148.12	0133.0	1910	3		N
WANDERING	010917	-32.67	116.67	0275.0	1910	2		N
WANGARATTA AIRPORT	082138	-36.42	146.31	0152.6	1910	3		Y
WARRNAMBOOL AIRPORT	090186	-38.29	142.45	0070.8	1910	2		Y
WELLINGTON	065034	-32.56	148.95	0300.0	1910	4	PE, PD	N
WHITE CLIFFS POST OFFICE	046042	-30.85	143.09	0151.0	1910	4	PE, NS	N
WILCANNIA	046043	-31.56	143.37	0075.0	1910	3		N
WILSONS PROMONTORY LIGHT.	085096	-39.13	146.42	0088.7	1910	1		N
WILUNA	013012	-26.59	120.23	0521.0	1910	5	MD	N
WOLLONGONG UNIVERSITY	068188	-34.40	150.88	0030.0	1910	4	PE	Y
WOOMERA AIRPORT	016001	-31.16	136.80	0166.6	1950	1		N
WYNDHAM POST OFFICE	001013	-15.49	128.12	0011.0	1910	4	PE, PD	N
YAMBA PILOT STATION	058012	-29.43	153.36	0029.0	1910	1		N
YASS	070091	-34.83	148.91	0520.0	1910	5	MD	N
YONGALA	019062	-33.03	138.75	0515.0	1910	1		N
YORK	010311	-31.90	116.77	0179.0	1910	2		N
YOUNG AIRPORT	073138	-34.25	148.25	0379.6	1910	5	MD	N

ten potential stations to include in the reference series. After the initial station search, the reference station list was shortened by only accepting stations which had at least ten years of monthly data and a population less than 10 000 (here population values used by Torok (1996) are used as a proxy for determining a significant influence of urbanisation on the temperature record). Time series from reference stations also had to be highly correlated with the candidate station such that the one-sided significance (H_0: correlation > 0) was 95 per cent or greater (Press et al. 1996). The minimum number of reference stations required to create a reference series was one, however the reliability of such reference series was low and objective test results were treated with extreme caution if the number of reference stations was less than ten.

The median reference series was created by taking the median of the ranked interannual temperature differences from each suitable reference station series for each year. This series was then added to the annual mean temperature of the first year of the candidate series to create the final median reference series. Torok (1996) suggests that taking the median of the interannual differences is more robust to outliers caused by inhomogeneities in reference stations than taking the mean and that the median difference series will only be significantly biased if a systematic change in observing techniques, data processing or exposure occurs at the same time throughout the network (e.g. the systematic change to expose thermometers in Stevenson screens). Torok (1996) concedes that the median reference series overall trend and variability can be biased by small differences in the reference station selection process and that the objective test can detect trends in the difference series by identifying a series of small discontinuities. While the detection of a trend in the difference series is desirable if the trend is the result of a slow change in the exposure of a candidate site (e.g. vegetation growth), it is clearly not desirable if the trend in the difference series is the result of an artificial trend in the median reference series. Trewin (2001) comments that the median reference series used by Torok (1996) can have biases introduced when converting the median interannual differences back into an absolute reference series by the accumulation of rounding errors in the interannual differences (also well documented in Peterson and Easterling (1994)). This can introduce spurious trends in the reference series. Trewin (2001) uses a distance weighted mean of highly correlated reference station anomalies to create the reference series and so avoids the problems associated with the conversion of interannual differences. The weights, W_i, are given by Eqn 1 where d_i is the inter-station separation measured in degrees, and r_i is the interstation correlation and N is the number of stations.

$$W_i = \left\{ \begin{array}{ll} r_i(6-d_i)^2, & r > 0.6, d_i < 6 \\ 0 & , r < 0.6, d_i > 6 \end{array} \right\} i=1,...,N \qquad ...1$$

The trend in the median reference series and distance weighted reference series were compared spatially to assess the potential biases in the two methods. Both methods show a general overall warming trend in minimum temperature across Australia (not shown). However, the magnitudes of the warming trends in the median reference series are generally greater than the trends in the distance weighted reference series. For maximum temperature, both methods show general warming trends in the central and western parts of the continent and distinct cooling trends in New South Wales and parts of Queensland. Again, the magnitude of the positive and negative trends in the median reference series is greater in magnitude than the distance weighted reference series. This leads us to remain cautious in the use of the median reference series as a basis for the objective test. Where possible the results of the objective test based on the distance weighted reference series were used to help identify and correct inhomogeneities.

The objective test has been found to over-estimate the number of discontinuities in a difference series since the nominated distribution of the test statistic, $F_{3,n\text{-}4}$ incorrectly assumes that multiple discontinuities are independent (Lund and Reeves 2002). The critical values of the corrected underlying distribution are typically twice those of the $F_{3,n\text{-}4}$ critical values, indicating that if detected discontinuities in a difference series are applied without metadata support then the resulting series could include adjustments associated with natural climate variability or as we have just discussed artificial trends associated with biased reference series. This problem is not likely to have greatly affected this dataset, since only 22 per cent of adjustments made were based on the results of the objective test without supporting documentation. Some unsupported discontinuities must be accepted as real for it is known that many historical station changes have not been documented. The metadata summaries of Torok (1996) reveal large gaps in station history files. There can often be many years between file entries, with station photos or site plans before and after the gap revealing obvious, but undocumented, changes occurring sometime in the intervening period.

Reference series test

The reference series test was simply a visual examination of the difference series used by the objective statistical test to detect discontinuities. Suspect looking step jumps in the series were used as evidence of possible dates of discontinuity in a record. Analyses of the number of neighbour stations used in a reference series throughout time were undertaken, as well as which stations actually contributed data to each reference series. These provided an indication of whether an apparent discontinuity in a difference series could simply be due to a change in the number of reference stations, or a change in which stations were used for the reference series.

Diurnal temperature range test

Visual examination of the unhomogenised annual mean diurnal temperature range (DTR) time series at each station was also used to detect discontinuities. Occasionally the presence of an inhomogeneity was more obvious in the DTR series due to opposite sign responses to a non-climatic influence in the individual maximum and minimum temperature series. The visual examination procedure was subjective and hence a discontinuity in the DTR series was only accepted if supported by metadata.

Parallel observation data

Since 1993 most major observation site moves in Australia have involved a change in station number. Prior to this update, if a high-quality station moved in the post-1993 period, and assumed a new station number, its entire record was removed from the dataset to avoid the known bias associated with compositing a new record without adjustment. One of the main objectives of this update was to resurrect as many of these 'closed' stations as possible by matching the homogenised record of Torok and Nicholls (1996) onto a continuing record.

Since the development of the original dataset, 44 stations moved to a new location with a new station number. Usually this was due to the closure of a co-operative manual observation site after the installation of an automatic weather station (AWS) at a nearby airport. A few stations were moved with no change in site number and, theoretically, should not have been used in the calculation of Australian mean temperatures until corrected for the change.

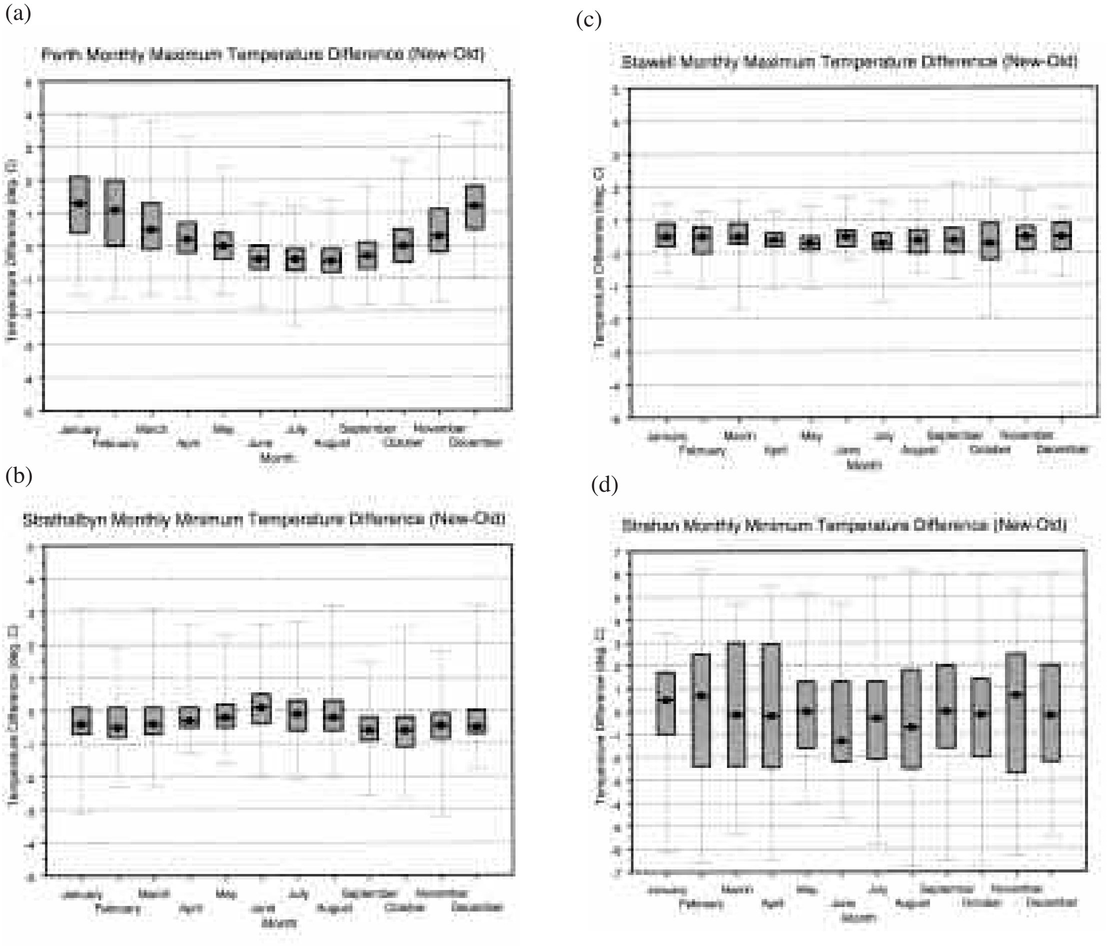
Karl et al. (1995) states that one of the basic requirements for sound climate monitoring is the undertaking of periods of parallel observation to enable the impact of changes in the observation regime to be assessed. This has been adopted as one of the standard practices for stations included in the Global Climate Observation System (GCOS) Surface Network which requires one, preferably two, years of overlap (Daan 2002). In Australia, an overlap period of preferably two years is requested for site moves at Reference Climate Stations (RCSs), stations specifically chosen to monitor long-term climate changes. Prior to the development of the RCS network in the early 1990s, parallel observations for site moves were not an official Bureau of Meteorology policy and were less common. Many of the high-quality temperature stations used here are classified as RCSs but not all, and consequently some were moved without a period of overlap observations being undertaken. However, some location moves at non-RCSs had overlap observations available due to chance circumstances, or the diligence of regional observers and managers who recognised the value of overlap observations for all long-record stations, not just RCSs. Of the 44 station moves that involved a change in station number, 30 had some parallel data between the old and new sites. These data were examined for quality and completeness.

Only 26 station moves were considered to have overlap data of sufficient quality to calculate a reliable adjustment between the old and new temperature records. The climatological difference between an old and new site can differ throughout the year, and the period of overlap rarely corresponds to calendar years. Consequently the annual mean adjustment cannot simply be calculated using the average of daily differences over the overlap period. Instead, the annual mean adjustment for a site move was calculated using the average of the mean monthly differences. To determine the mean monthly differences for each calendar month, all available monthly differences between the sites were averaged (Fig. 1).

Generally, comparison observations for longer than five years were found to provide excellent comparison statistics between the old and new sites, e.g. Figs 1(a) and 1(b). Comparisons longer than two years, and sometimes between one and two years, were also found to be useful if complete and of good quality, e.g. Fig. 1(c). Poor quality comparisons lasting less than two years were generally found to be of limited use, e.g. Fig. 1(d). Also, the deterioration of the old site can sometimes result in some of the comparison data being unrepresentative and of little value. The period of comparison required for meaningful statistics at any particular location depends on the distribution of climatological differences between the old and new sites.

An example of high quality comparison data for maximum temperatures recorded at old and new observation sites at Perth is shown in Fig. 1(a). Over ten years of complete comparison data were available, allowing reliable differences between the two sites to be calculated. Figure 1(a) is also a good example of how the climatological difference between an old and new site can differ throughout the year, with the mean differences during June to September being of reversed sign to the remainder of the year. Despite the good quality comparison data the spread of daily differences during each month is still quite high, reflecting real climatological differences between the two sites.

Fig. 1 For each calendar month, the mean (dots), inter-quartile range (shaded) and highest/lowest values (line bars) of difference between new and old observation sites for (a) Perth maximum temperature, (b) Strathalbyn minimum temperature, (c) Stawell maximum temperature and (d) Strahan minimum temperature. Based on (a) 12 years of good quality, (b) five years of good quality, (c) 22 months of good quality and (d) 18 months of poor quality overlap observations. Note change in vertical scale in (d).



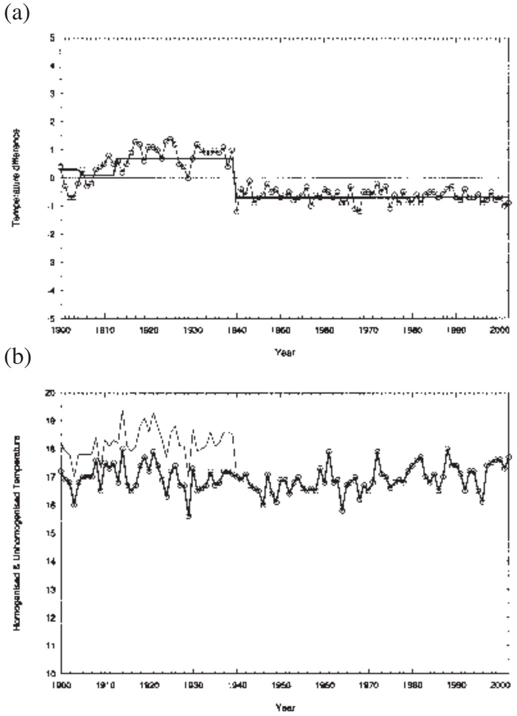
The adjustments calculated using overlap data were compared to the shifts determined by the objective method where possible. Generally the objective method did not detect a discontinuity due to a site move unless it was of relatively large magnitude, about 0.5°C or larger. Where the discontinuity was detected there was usually good agreement in magnitude of the shift. This highlights the importance of good quality metadata and comparison data, and suggests that it is likely that small discontinuities in the records have not been detected by the objective technique. For site moves where no parallel data were available the shift value calculated by the objective technique was used. If this was not available an estimated value was used based on subjective techniques.

Applying homogeneity adjustments

The objective technique requires a test window of data ten years in length (five years before and after a test point) so is unable to detect a discontinuity occurring within the last five years of record. Consequently, for this latest assessment any post-1988 homogeneity adjustment detected by the objective method, and supported by metadata, was applied to the updated record. Some adjustments not supported by metadata were also made. Where a discontinuity was evident from the reference series or DTR test, but not detected by the objective method, the magnitude of adjustment was estimated. An example of differences between the candidate and homogeneous reference series for Launceston maximum temperature series is shown in Fig. 2(a), along with the series homogeneity adjustments applied. The corresponding unadjusted and adjusted temperature series are shown in Fig. 2(b). For the post-1993 assessment a total of 108 new adjustments (48 to maximum temperature series and 60 to minimum temperature series) were made across the network. Overall, the updated dataset includes an average of approximately six adjustments per record since 1910.

The decision of whether or not to correct for a potential inhomogeneity is often a subjective one. Supporting metadata evidence of a discontinuity makes that decision easier but is not always available. Consequently, it was often impossible to reproduce the exact homogeneity adjustments of Torok and Nicholls (1996). Slightly different techniques, reference stations and source data can apparently produce different results. Rather than correct the entire dataset with new adjustments, the pre-1994 homogeneity adjustments determined by Torok (1996) were generally re-applied for the updated series, except for about 30 adjustments in which an obvious error had been made, such as the wrong sign of adjustment or an incorrect year of adjustment. A few records were completely reassessed due to significant differences between the objective test results based on the two types of reference series used.

Fig. 2 **(a) Differences between candidate and reference series (line with open circles), and the corresponding series of applied homogeneity adjustments (bold line), and (b) unadjusted (thin line) and adjusted (bold line with open circles) annual mean maximum temperature series for Launceston. No adjustments were required since 1940.**



Assessment of potential urban warming

Torok and Nicholls (1996) used a simple definition of an urban station, classifying any location with a population greater than 10 000 as urban. However, some 'urban' stations located near, but not within, a large population centre, such as at an airport or lighthouse, may not necessarily have experienced a significant urban warming signal, particularly if the station was relocated to the airport many decades ago. Torok et al. (2001) show that even relatively small towns with populations less than 10 000 can have a detectable urban heat island signal during the right synoptic conditions. However, such conditions are only prevalent for a small fraction of the year (Torok et al. 2001). The influence of urbanisation is also complicated by poor exposure at many observation sites, such as being located in or near a bituminised carpark. Consequently, it remains unclear how much urban warming is really evident at the annual time-scale for small towns.

This study attempted a number of approaches to determine whether a station was likely to have been affected by urban warming. One of the key tools used was a statistical test of the difference between temperature trends calculated using the actual station data and an interpolated trend value from a spatial analysis based on the non-urban stations defined by Torok and Nicholls (1996). A significantly greater rise in the minimum temperature at the station compared to the 'non-urban' background (using a two-sided Kendall Tau test (Press et al. 1996)) was assumed to be indicative of an urban warming signal. However, this approach assumes that the stations originally classified as non-urban have not been affected by some urbanisation.

Other information used in the assessment of urbanisation included the date at which the town exceeded a population of 10 000 (ABS 1996, 1986, 1981, 1976, 1971), land-use codes defined in SitesDB and the station location relative to the town centre. Ultimately a subjective decision was made for each station as to whether it was likely to have been influenced by urbanisation. Some stations originally classified as urban were re-classified as non-urban and vice-versa, with 39 stations being classified as urban in this latest assessment (Table 1).

The updated Australian high-quality annual temperature network

A total of 31 ceased temperature records of Torok and Nicholls (1996) have been re-opened by matching them onto a continuing record. Also, parts of inland Australia and Tasmania are now better sampled due to the inclusion of five homogenised records located at Giles, Woomera, Tennant Creek, Cape Bruny and Tarcoola (Fig. 3). These records are of shorter duration than other series and hence were not included in the original dataset.

Unfortunately sixteen of the original records were found to have closed with no potential candidate station available to continue the record. Also, due to a decline or reassessment of data quality, 81 stations were removed from the dataset. Of the remaining 133 records ranked fair (rank 3) to very good (rank 1), 34 were classified as being urban, leaving 99 non-urban records on which to calculate Australia's annual mean maximum and minimum temperature. This is equal to the number available prior to the update. However, the removal of poor quality records, the application of post-1993 homogeneity adjustments and the inclusion of additional stations in data-sparse regions means that overall the dataset has been greatly improved. The stations included in the updated Australian annual mean temperature dataset are listed in Table 1, and the distribution of the high-quality subset (records with rank 1, 2 or 3) shown in Fig. 3.

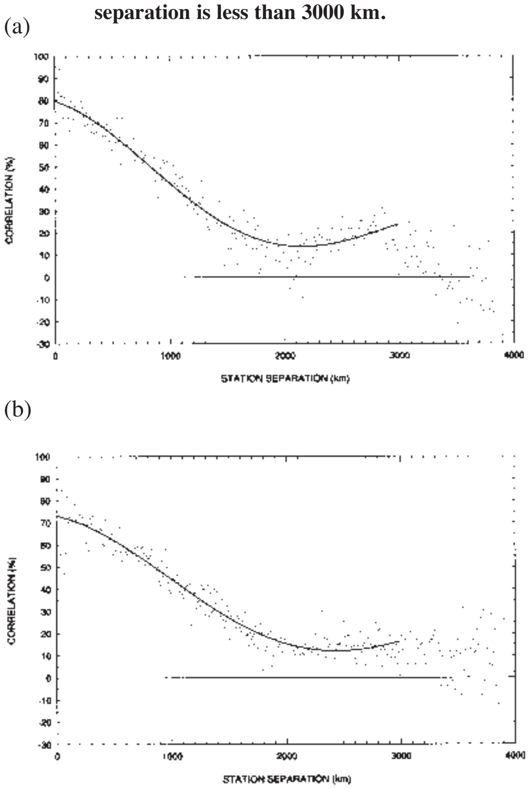
Adequacy of the updated annual temperature network

In order to determine the adequacy of the updated network to resolve interannual climate variability the decorrelation (maximum significant inter-station correlation) length scale of annual de-trended temperature anomalies was calculated using the method of Jones and Trewin (2000), based on the theory of Daley (1993). Figure 4 shows all inter-station correlations over the 1961-1990 period for annual maximum and minimum temperature anomalies as a function of station separation. The correlations show a general decline with inter-station separation up to around 2000 km, then tend to remain around 0.2 or increase slightly before becoming scattered at separations between 3000 and 4000 km.

Plots of the angle (0-π) between stations against inter-station distance indicate that the higher correlations tend to have a north-south orientation at separations between zero and 2000 km, then tend to become more northwest–southeast at separations between 2000 and 3000 km (not shown). The orientation of the inter-station correlations largely reflects the station

Fig. 3 **Updated high-quality temperature network showing non-urban (solid circles) and urban (solid triangles) station locations.**

Fig. 4 **1961-1990 inter-station correlations (expressed as %) for annual mean (a) maximum and (b) minimum temperatures as a function of separation distance. Exponentially damped cosine functions are fitted where the station separation is less than 3000 km.**



distribution (Fig. 3). However, some of the increase in maximum temperature correlations (Fig. 4(a)) and the tapering decline in minimum temperature correlations at separations between 2000 and 3000 km (Fig. 4(b)), could be due to the anisotropic northwest-southeast nature of Australian temperature anomalies (Jones and Trewin 2000; Seaman 1982).

Exponentially damped cosine functions were only fitted to the inter-station correlation data for separations up to 3000 km since the number of inter-station correlations diminishes rapidly beyond this distance. Using a critical correlation value of 0.54 (the one-sided 99.9 per cent significant correlation for a non-autocorrelated series of length 30 (Jones and Trewin 2000)) the fitted curves suggest that for all points within 745 (735) km of a high-quality station, an interpolated annual maximum (minimum) temperature anomaly value will be a reasonable estimate. The inter-station correlations shown in Fig. 4 are based on de-trended data and therefore not inflated by time series having similar trends.

These de-correlation length scales were compared to the average inter-station separation. Koch et al. (1983) calculate a theoretical average station separation given by Eqn 2 that is weighted such that if the station distribution is highly clustered its value will be greater than the simple average of the minimum distance from one station to another.

$$\Delta n_r = A^{1/2}[(1 + M^{1/2})/(M - 1)] \qquad ...2$$

In Eqn 2, A represents the land surface area of Australia (km^2) and M is the number of stations in the network. Using this equation the average station separation of the updated high-quality annual temperature dataset is 316 km, compared to the simple average inter-station separation of 163 km, indicating some clustering of stations. Both these values are well below the de-correlation length scales calculated above using this dataset, as well as the monthly de-correlation length scales determined by Jones and Trewin (2000), suggesting that the high-quality annual temperature network is adequate to resolve the interannual variability of annual mean temperatures. So despite some apparently large gaps remaining in the network the updated distribution is able to provide adequate coverage of the country for the purposes of calculating annual mean temperature anomalies.

Observational errors in the updated network

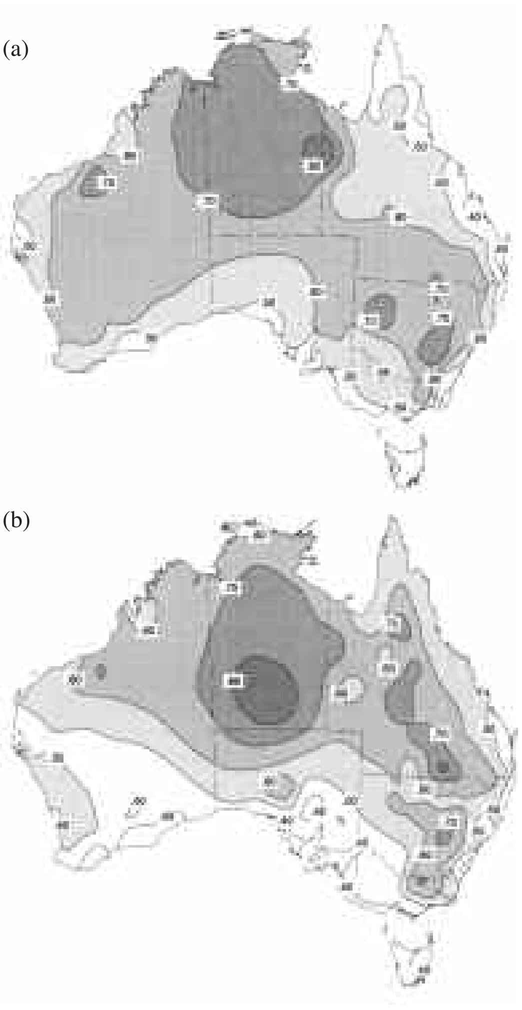
Another approach to assessing the adequacy of the network to resolve interannual variability is to look at the theoretical observation errors and compare them to the total variability of annual maximum and minimum temperature using the techniques of Daley (1993). The study has shown that the y intercept (station separation equals zero) of the exponentially damped cosine function plotted in Fig. 4 can be used to approximate the observation error variance. Theoretically, instruments that have infinitesimally small separation should be sampling the same air temperature and hence their correlation should be 1.0. However, due to instrument, observer and exposure errors this never occurs in practice. Assuming that the homogenised station records in the dataset are free from systematic biases, the inter-station correlations are isotropic in space, and the remaining observation errors are random and not correlated in time or space, then the observation error variance is given by Eqn 3. R_Z is the theoretical correlation at an inter-station distance of zero and $\mathrm{E}_T^2 = R_Z\,\sigma^2$, where σ^2 is the total variance of the station observations which, in this case, is based on the 1961-1990 climatology (see Fig. 5 for plots of σ).

$$E_O^2 = \frac{E_T^2\,(1 - R_z)}{R_z} \qquad ...3$$

Using Eqn 3, the average standard deviation of observational error $\sqrt{E_o^2}$ for annual maximum and minimum temperature is estimated to be 0.25ºC and 0.28ºC respectively. These represent approximately 40 per cent (16 per cent) and 46 per cent (22 per cent) of the all-Australian average maximum and minimum annual temperature standard deviations (variance) shown in Fig. 5. As expected, the average standard deviations of observational errors $\sqrt{E_o^2}$ for annual mean maximum and minimum temperatures (0.25ºC; 0.28ºC) are lower than those for mean monthly (0.45ºC; 0.42ºC) and daily (3.16ºC; 2.88ºC) anomalies obtained using the same method by Jones and Trewin (2000; 2002). The results above can be interpreted as the Australian average observation error in each annual, monthly and daily station observation, respectively. This network is considered to be adequate since the observational errors are less than half the observed standard deviation of the annual maximum and minimum temperature.

Della-Marta (2002) shows similar independent estimates of annual observation error based on the summation of uncertainties associated with bias corrections and measurement error with root mean squared error (rmse) of 0.26ºC and 0.24ºC for annual maximum and minimum temperature, respectively. These figures are not exactly comparable since the method of Daley (1993) assumes that any bias in a station record has been removed and that the error

Fig. 5 **1961-1990 standard deviation of annual mean (a) maximum and (b) minimum temperature anomalies.**



associated with this bias removal is not explicitly taken into account, whereas Della-Marta (2002) removes the bias in station records but retains an estimate of the error of the homogenisation process and explicitly takes these into account. In reality, the process of homogenisation will not remove a bias completely and so the observations used to calculate the theoretical observation error (given by Eqn 3) will contain some component that is due to the remaining bias that will manifest itself as a change in the correlation structure.

Accuracy of the objective analysis

The non-urban annual temperature records ranked 1 to 3 (99 stations detailed in Table 1) represent the high-quality subset used to create all-Australian annual mean temperature anomalies. A thorough investigation of the theoretical analysis errors at each grid point can be achieved through the optimal interpolation techniques of Daley (1993). However, a cross-validation approach was adopted here because the most influential factor on the accuracy of an analysis is the station density (Jones and Trewin 2000) and hence this approach will yield an upper limit to the analysis error. Daley shows that the analysis error is also dependent on the accuracy of the observation and climatological (background) error covariance structures to their true structures. Seaman (1983) shows a good example of how the theoretical analysis error increases as a function of the misspecification of the decorrelation length and observation error.

Initially station anomalies (with respect to 1961-90 normals) were gridded using Statistical Interpolation (SI) given a fitted damped cosine function (Fig. 4) to explain the background covariance structure (Jones and Trewin 2001; Daley 1993). Individual stations were then removed from the analysis and the difference between the observed station temperature anomaly and interpolated station temperature anomaly was used to calculate a rmse due to interpolation at the station location. The spatial plots (not shown) of these errors (hereafter known as interpolation errors) indicate that generally maximum temperature anomalies have lower interpolation errors than minimum temperature. The average rms interpolation errors are 0.31ºC and 0.34ºC for maximum and minimum temperature, respectively.

Figure 6 shows the ratio of rms interpolation errors to the observed standard deviation expressed as a percentage. This ratio shows, averaged over Australia, that interpolation errors are approximately 50 per cent and 58 per cent of the standard deviation for maximum and minimum temperature respectively, a similar result to the 45 per cent determined by Jones and Trewin (2000) using monthly temperatures. The interpolation errors expressed here are a combination of observation error and analysis error and by comparing these figures to the theoretical magnitude of the observation errors detailed in the last section it is clear that the majority of interpolation errors are due to observation error rather than analysis error.

The shading in Fig. 6 denotes regions in which the ratio of rms interpolation errors to observed standard deviation is less than 0.5. For maximum temperatures (Fig. 6(a)) it can be seen that there are consistently higher interpolation errors along the east coast of

Fig. 6 **Ratio (%) of rms error of cross-validated observations and observed standard deviation for annual (a) maximum and (b) minimum temperature using data from 1961-1990.**

Australia and the west coast of Tasmania, probably due to the large gradients of maximum temperature anomaly associated with broadscale topographic features such as the Great Dividing Range. Another region of high noise-to-signal ratio occurs in the northwest of Western Australia due to low station density and high maximum temperature gradients associated with sea-breeze effects (Jones and Trewin 2000). Minimum temperature interpolation errors (Fig. 6(b)) generally show large noise-to-signal ratios where the station density is low. Topographic influences on error calculations for minimum temperature are not as well defined as for maximum temperatures, except in the southeast. Land-sea proximity effects are likely to be of greater influence.

The noise-to-signal ratios shown in Fig. 6 are only an indication of the 'true' analysis errors since interpolation error is only assessed at station locations. These are likely to be an overestimate of the analysis error since, in reality, the dataset contains a station where the cross-validation is taking place. In order to get a reliable estimate of the total analysis error the interpolation error at each grid-point needs to be examined. This was considered beyond the scope of this paper.

Annual mean temperature trends

Torok and Nicholls (1996) calculated Australian annual mean temperature anomalies using a Theissen polygon weighted-average technique (Lavery et al. 1992). For the updated dataset this was done using the successive correction scheme used operationally to produce gridded analyses of temperature and rainfall (Jones and Weymouth 1997). Simple averages of the interpolated grid values were used to produce Australian annual mean maximum and minimum temperature time series (Fig. 7). These show the general rise in temperature during the 20th century which has been previously documented (eg. Torok and Nicholls 1996; Collins and Della-Marta 1999). The similarity to previous results suggests that all-Australian mean temperature anomalies based on a network of about 100 stations are robust to changes in some of the individual stations within that network, as well as changes in the averaging technique. Linear regression over 1910 to 2002 reveals trends of 0.6±0.3°C/(100 years) in the maximum temperature series and 1.1±0.3°C/(100 years) in the minimum temperature series, again consistent with previous studies. The uncertainties shown here are based on the standard error of the linear residuals (Collins and Della-Marta 1999).

The general rise in Australian temperatures during the 20th century is consistent with temperature trends measured by surface instrumentation throughout the globe. IPCC (2001) states that the global average surface temperature increased by 0.6±0.2°C over the 20th century. This global warming trend is supported by independent data records such as those from weather balloons, satellites, sea-ice and glacial extent, and is at least partly attributable to human activities (IPCC 2001).

Throughout most of the country, updated post-1910 trend values (Fig. 8) have changed little compared with those of the original dataset. All of the country displays a warming trend in minimum temperatures, the largest trends being in the northeast. Increases in maximum temperatures have been

Fig. 7 **Time series of Australian annual mean maximum (solid line) and minimum (dashed line) temperatures based on the updated non-urban network (linear trends (1910-2002) also shown). Anomalies are with respect to the 1961-1990 normal.**

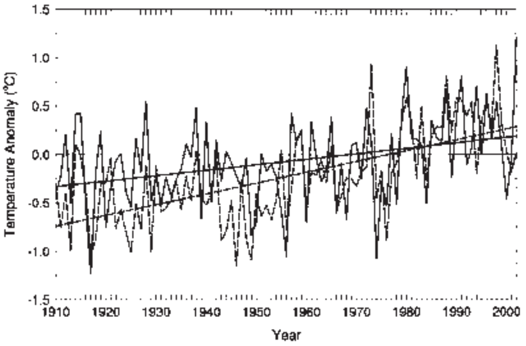
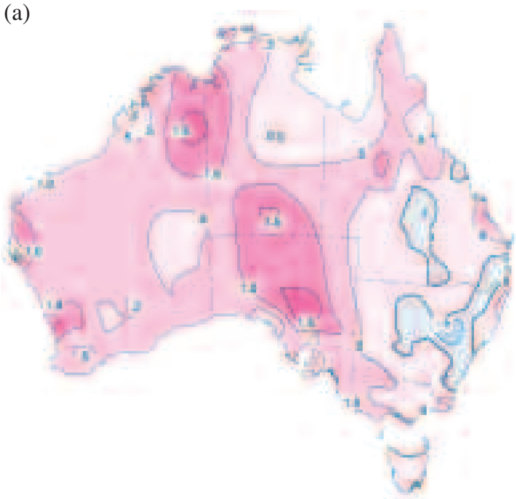
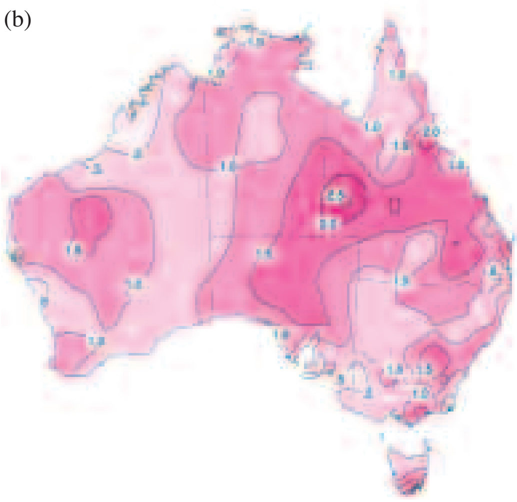


Fig. 8 **Maps of 1910-2002 linear trend in Australian annual mean (a) maximum and (b) minimum temperatures (ºC/100 years) based on the updated, non-urban network.**

(a)



(b)



strongest in the western two-thirds of the country, while some parts of New South Wales and Queensland have actually recorded weak cooling trends since 1910. However, the magnitude and extent of the cooling in central and northern New South Wales is reduced in the updated dataset compared to previous published analyses based on the original dataset (e.g. Collins and Della-Marta 2002). This is due to the removal of many poor quality records in the region.

Most New South Wales stations, and some Queensland stations, were missing digitised data between 1957 and 1964 due to data processing problems (Trewin 2001). In inland New South Wales the annual mean, unhomogenised maximum temperatures prior to 1957 were generally greater than post-1964 and consequently there is a decrease in the median reference series for most stations over this period. Further analyses using only those records with complete 1957-1964 data continued to suggest some maximum temperature cooling in the region since 1910. The distance weighted reference series of Trewin (2001) also displayed cooling through New South Wales and southern Queensland (not shown). The cooling is also consistent with an increase in annual total rainfall in the region (Collins and Della-Marta 2002). Consequently, some maximum temperature cooling is likely to have occurred through parts of New South Wales and Queensland since 1910. A confounding variable in the verification of this cooling trend is the fact that some of the high-quality (New South Wales) stations compiled by Torok and Nicholls (1996) were not exposed in a standard Stevenson screen until the 1940s (Torok 1996; Parker 1994). Where sufficient evidence (metadata) existed, these stations were removed from this dataset. Given this documented non-standard exposure at some candidate stations it is possible that other stations, including those used in the calculation of reference series, might contain a negative bias in their overall trend, which in turn could influence the outcome of the homogenisation process. A more detailed investigation of the available metadata at reference stations and the digitisation of missing 1957-1964 data are required to provide a clearer understanding of the extent and magnitude of this cooling.

Discussion and conclusions

The updated dataset of 99 good quality, non-urban temperature records has been found to provide a reliable measure of Australian annual mean temperatures. This supports the decision to include approximately 100 mainland stations in Australia's RCS network. However, such a network does not meet all the observation needs for climate applications and services. It is too coarse to fully describe local-scale climate or resolve trends in variables with high spatial variability, such as rainfall. Nevertheless, a network of about 100 high-quality stations does provide a backbone on which to base denser networks suitable for other climate purposes.

Despite the suitability of the dataset for national and regional-scale analyses, any individual station record within the dataset should still be treated with caution. The subjectivity inherently involved in the homogeneity process means that two different adjustment schemes will not necessarily result in the same homogeneity adjustments being calculated for individual records. However, if the overall biases of the different approaches are neutral then spatial averages should be highly consistent across a large number of homogenised records. Also, even though a subjective assessment of the likelihood of urban contamination within these records has been made, further work needs to be undertaken to better understand the relative contribution of urban warming to the temperature records of small Australian towns.

Torok (1996) and Trewin (2001) both note bias problems associated with using median reference series based on interannual temperature differences. This technique was compared to the distance weighted reference series of Trewin (2001) by analysing the difference in trend in the reference series for each candidate station. The trends were found to be consistent in sign but not in magnitude, with trends in the median reference series often greater than those of the distance weighted reference series. The total influence of this bias in the pre-1993 homogeneity adjustments is small since the majority of adjustments are verified by documentation. Both reference series in data sparse areas were sometimes unreliable and led to misleading results from the objective test. In these areas (predominately the northern and western parts of the continent) the median reference series sometimes displayed discontinuities associated with changes in the number of reference stations available. Extreme caution was used when making undocumented changes based on the median reference series in these areas. However, it is possible that records in remote location have not been fully homogenised.

The difficulties associated with developing truly homogeneous reference series highlight the importance of good quality station metadata to aid in the identification of discontinuities. This requires climate data users to maintain strong links with observation network managers and to continue to promote the need for accurate, complete and accessible station metadata. Technological developments must be used to ensure that the vast amounts of electronic metadata that are now available are captured for the needs of future generations.

The imperfection of statistical homogeneity techniques also underscores the value of parallel observations for major changes in the observation regime. When artificial discontinuities are small relative to natural variability, objective statistical techniques are often unable to detect the shift. In such cases the only real way to determine the shift is with the aid of complete and accurate comparison data. The period of comparison required is dependent on the broadscale climate and the microclimatic differences of the locations. In this study a good understanding of the climatological impact of a site move was generally only achieved with at least two years of good quality overlap, supporting the Bureau of Meteorology's policy of at least one, preferably two, years of overlapping records for location moves at important climate observation sites.

It is also vital that comparisons between instrument types continue to be undertaken in order to ensure the homogeneity of the climate record. The updated dataset remains limited to the post-1910 period because before this time the use of Stevenson screens was not consistent across Australia (Nicholls et al. 1996). Reference series prior to 1910 are unreliable because of the systematic change in exposure across the network at approximately the same time. Nevertheless, already published exposure comparisons and possible modern comparisons between historical exposure types may eventually allow the extension of some homogeneous annual temperature records backwards in time (Nicholls et al. 1996). However, at most, about 60 unevenly-distributed stations have sufficient data back to 1900, and only about 37 stations back to 1890, suggesting that reliable all-Australian averages will remain restricted to the post-1910 period.

Greater scope exists to extend monthly and seasonal trend analyses further back in time than currently possible. The records in the dataset updated here have been adjusted at the annual time-scale and consequently are inappropriate for analyses at intra-annual time-scales. However, the dates of adjustment and metadata summaries generated could easily form the basis for a high-quality dataset corrected at the monthly time-scale. A high-quality temperature dataset corrected at the daily time-scale has been developed by Trewin (2001) but generally only extends from around 1957 due to a lack of digitised daily data prior to this time. It is hoped that additional digitised data provided by the CLIMARC project (Clarkson et al. 2001) will ultimately enable the extension of the high-quality daily records backwards in time and provide a good coverage of corrected, century-scale daily temperature records across Australia. This would enable longer monthly and seasonal trend analyses than currently available. However, this will require significant resources and consequently the annual temperature dataset updated here will continue to provide the basis for century-scale climate change analyses for some time.

Acknowledgments

The update of station metadata summaries would not have been possible without funding received through the Bureau of Meteorology's special project bids process. The authors wish to thank Dr David Jones for assistance in calculating inter-station correlations and useful discussions on observation and analysis errors, and Dr Blair Trewin for discussions on historical thermometer exposures and homogenisation techniques. The authors are also grateful to Mr Neil Plummer and Dr Blair Trewin for internally reviewing a draft manuscript, and to three external reviewers whose comments were useful in shaping the final form of this paper.

References

ABS 1996. *Selected characteristics for urban centres 2016.0*. Australian Bureau of Statistics, Canberra.

ABS 1986, 1981, 1976, 1971. *Census of population and housing*. Australian Bureau of Statistics, Canberra.

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, September 18-21, 2001, p.49.

Collins, D.A. and Della-Marta, P.M. 1999. Annual climate summary 1998: Australia's warmest year on record. *Aust. Met. Mag., 48,* 273-83.

Collins, D.A. and Della-Marta, P.M. 2002. Atmospheric indicators for the State of the Environment Report 2001. *Technical Report No. 74*, Bur. Met., Australia.

Daan, H. 2002. Manual on the GCOS surface and upper-air networks: GSN and GUAN. *GCOS-73, WMO/TD No. 1106*, WMO, Geneva.

Daley, R. 1993. *Atmospheric data analysis*. Cambridge University Press, NY, U.S.A., 457pp.

Della-Marta, P.M. 2002. Measuring uncertainty in the Australian annual temperature record. *AMOS 2002: 9th National Australian Meteorological and Oceanographic Society Conference*, 18-20 February 2002, University of Melbourne, p.25.

Easterling, D.R. and Peterson, T.C. 1995. A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol., 15*, 369-77.

IPCC 2001. *Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge University Press, UK, 1000pp.

Jones, D.A. and Trewin, B.C. 2000. The spatial structure of monthly temperature anomalies over Australia. *Aust. Met. Mag., 49*, 261-76.

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*, 237-50.

Jones, D.A. and Weymouth, G.T. 1997. An Australian monthly rainfall dataset. *Technical Report No. 70*. Bur. Met., Australia.

Karl, T.R., Derr, V.E., Easterling, D.R., Folland, C.K., Hofmann, D.J., Levitus, S., Nicholls, N., Parker, D.E. and Withee, G.W. 1995. Critical issues for long-term climate monitoring. *Climatic Change, 31*, 185-221.

Koch, S.E., DesJardins, M. and Kocin, P.J. 1983. An interactive Barnes objective map analysis scheme for use with satellite and conventional data. *Jnl Clim. appl. Met., 22*, 1487-503.

Lane, J. 1995. *Sites Database: Functional Requirements*. Bureau of Meteorology internal report. Bur. Met., Australia.

Lavery, B., Kariko, A. and Nicholls, N. 1992. A historical rainfall data set for Australia. *Aust. Met. Mag., 40*, 33-9.

Lund, R. and Reeves, J. 2002. Detection of undocumented changepoints: a revision of the two-phase regression model. *Jnl climate, 15*, 2547-54.

Nicholls, N., Tapp, R., Burrows, K. and Richards, D. 1996. Historical thermometer exposures in Australia. *Int. J. Climatol., 16,* 705-10.

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

Peterson, T.C. and Easterling, D.R. 1994. Creation of homogeneous composite climatological reference series. *Int. J. Climatol., 14*, 671-9.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S.J., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, M.J., Forland, E., Hanssen-Bauer, I., Alexandersson, H., Jones, P.D. and Parker, D.E. 1998. Homogeneity adjustments of In Situ atmospheric climate data: a review. *Int. J. Climatol., 18,* 1493-517.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. 1996. In: *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd Ed. Cambridge University Press, New York, 630-639pp.

Seaman, R.S. 1982. A systematic description of the spatial variability of geopotential and temperature in the Australian region. *Aust. Met. Mag., 30*, 133-41.

Seaman, R.S. 1983. Objective analysis accuracies of statistical interpolation and successive correction schemes. *Aust. Met. Mag., 31,* 225-40.

Torok, S.J. 1996. The development of a high quality historical temperature data base for Australia. PhD Thesis, School of Earth Sciences, University of Melbourne, Australia.

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

Torok, S.J., Morris, C.J.G., Skinner, C.J., and Plummer, N. 2001. Urban heat island features of southeast Australian towns. *Aust. Met. Mag., 50*, 1-13.

Trewin, B.C. 2001. Extreme temperature events in Australia. PhD Thesis, School of Earth Sciences, University of Melbourne, Australia.