

1 Jan 1910

2 Jan 1910

3 Jan 1910

4 Jan 1910

5 J

1 Feb 1910

2 Feb 1910

3 Feb 1910

4 Feb 1910

5 F

1 Mar 1910

2 Mar 1910

3 Mar 1910

4 Mar 1910

5 M

New Australian daily historical climate surfaces using CLIMARC

Technical Report
July 2004

1 May 1910

2 May 1910

3 May 1910

4 May 1910

5 M

David Rayner, Keith Moodie, Alan Beswick,
Nick Clarkson, Rod Hutchinson

1 Jun 1910

2 Jun 1910

3 Jun 1910

4 Jun 1910

5 J



Queensland Government
Natural Resources, Mines and Energy
Primary Industries and Fisheries



Climate Variability
in Agriculture R&D Program



Australian Government
Bureau of Meteorology
Land & Water Australia

New Australian daily historical climate surfaces using CLIMARC*.

July 2004

**David Rayner¹, Keith Moodie¹, Alan Beswick¹,
Nick Clarkson², Rod Hutchinson³**

1. Queensland Department of Natural Resources, Mines and Energy,
Natural Resource Sciences
QCCA Building, Gate 4, 80 Meiers Rd,
Indooroopilly, Qld. 4064, Australia.
2. Queensland Department of Primary Industries and Fisheries,
Predictive, Precision Systems
PO Box 102,
Toowoomba, Qld. 4350, Australia.
3. Australian Government Bureau of Meteorology,
National Climate Centre
GPO Box 1289K,
Melbourne. Vic. 3000, Australia.

*“CLIMARC - Computerising the Australian climate archives” (QPI43) Nick Clarkson, Queensland Centre for Climate Applications (QCCA)/DPI&F, was a joint project between the Climate Variability in Agriculture Program (CVAP), the Queensland Department of Primary Industries and Fisheries (DPI&F), the Queensland Department of Natural Resources, Mines and Energy (NRM&E), and the Australian Government Bureau of Meteorology.

The Climate Variability in Agriculture R&D Program (CVAP) was funded by the Australian Government, and part of the Agriculture Advancing Australia package. The program was managed for the Department of Agriculture Fisheries and Forestry - Australia by Land & Water Australia (LWA). In addition to LWA, the program was supported by Grains, Rural Industries, Sugar and Dairy R&D Corporations. The Managing Climate Variability R&D Program replaced CVAP in 2003.

Department of Natural Resources, Mines and Energy, Queensland
July 2004

QNRME04247
ISBN 1 920920 86 2 (print)
ISBN 1 920920 87 0 (online)

General Citation

Rayner, D.P., Moodie, K.B., Beswick, A.R., Clarkson, N.M., Hutchinson, R.L. (2004), New Australian daily historical climate surfaces using CLIMARC. Queensland Department of Natural Resources, Mines and Energy.

General Disclaimer

Information contained in this publication is provided as general advice only. For application to specific circumstances, professional advice should be sought.

The Queensland Department of Natural Resources, Mines and Energy has taken all reasonable steps to ensure that the information contained in this publication is accurate at the time of production. Readers should ensure that they make appropriate enquiries to determine whether new material is available on the particular subject matter.

© The State of Queensland, Department of Natural Resources, Mines and Energy 2004

Copyright protects this publication. Except for purposes permitted by the *Copyright Act 1968*, reproduction by whatever means is prohibited without prior written permission of the Department of Natural Resources, Mines and Energy, Queensland.

Enquiries should be addressed to:

David Rayner
David.Rayner@nrme.qld.gov.au
Department of Natural Resources, Mines and Energy
80 Meiers Rd Indooroopilly
Brisbane Qld 4068

Table of Contents

1	ACKNOWLEDGEMENTS.....	6
2	SUMMARY.....	7
3	OVERVIEW.....	9
3.1	CLIMATE CHANGE TRENDS.....	10
4	EXAMPLE CLIMATE DATA: PLOTS AND SURFACES.....	11
5	REVIEW.....	13
5.1	OBSERVED CLIMATE RECORDS.....	13
5.2	SPATIAL CLIMATE DATASETS.....	13
5.3	EXTENSION OF THE NRM&E CLIMATE ARCHIVE USING CLIMARC.....	14
6	CLIMATE OBSERVATION DATA.....	15
6.1	SUMMARY.....	15
6.2	METEOROLOGICAL STATIONS.....	15
6.3	PRE-INTERPOLATION PROCESSING OF CLIMATE DATA.....	18
6.3.1	<i>Data checking.....</i>	18
6.3.2	<i>Systematic errors and discontinuities.....</i>	19
7	SPATIAL INTERPOLATION OF CLIMARC NETWORK DATA.....	21
7.1	SUMMARY.....	21
7.2	OBJECTIVES.....	21
7.3	DESCRIPTION OF THE ANOMALY INTERPOLATION ALGORITHM.....	22
8	OVERALL PROPERTIES OF THE ANOMALY-INTERPOLATED SURFACES.....	25
8.1	GENERAL COMMENTS ON THE RELIABILITY OF EARLY CLIMATE RECORDS.....	25
9	IMPROVEMENTS IN THE CLIMATE SURFACES.....	27
9.1	IMPROVED ESTIMATE OF OBSERVED DAILY CLIMATE: CROSS-VALIDATION ANALYSIS.....	27
9.2	IMPROVED ESTIMATE OF OBSERVED DAILY CLIMATE: A REPRESENTATIVE NETWORK ANALYSIS.....	28
9.2.1	<i>Representative station network.....</i>	28
9.2.2	<i>Improved overall accuracy.....</i>	28

9.2.3	<i>Improved reliability.....</i>	29
9.2.4	<i>Spatial variation of interpolation errors.....</i>	29
9.3	REPRESENTATION OF INTER-ANNUAL CLIMATE VARIABILITY.....	35
9.3.1	<i>Maximum and minimum temperature annual variability.....</i>	35
9.3.2	<i>Dew point calculated from vapour pressure.....</i>	37
9.3.3	<i>Solar radiation and cloud amount.</i>	39
9.4	INCLUSION OF INTER-VARIABLE CORRELATIONS.	40
9.4.1	<i>The value of inter-variable correlations.</i>	40
9.4.2	<i>Inter-variable correlations in Warrego River catchment time-series.</i>	41
10	LIMITATIONS OF THE ANOMALY-INTERPOLATED CLIMATE SURFACES.....	43
10.1	SYSTEMATIC ERRORS IN OBSERVATION RECORDS.	43
10.1.1	<i>Inaccurate trends in the maximum and minimum temperature climate surfaces.....</i>	43
10.1.2	<i>The effect of systematic errors in station temperature records.....</i>	44
10.1.3	<i>Bias in the pre-1949 solar radiation surfaces.....</i>	45
10.1.4	<i>The discontinuity in the NRM&E solar radiation surfaces due to the conversion of cloud observations from tenths to oktas.....</i>	47
10.2	THE EFFECT OF USING ZERO-ANOMALIES IN THE INTERPOLATION.	47
10.2.1	<i>Trends in the early climate surfaces.</i>	47
10.2.2	<i>Statistical analysis of the day-to-day variability in the NRM&E climate surfaces.</i>	48
10.2.3	<i>The effect of zero-anomalies on the differences between the NRM&E surfaces and the Torok and Nicholls time-series.</i>	49
10.3	OTHER LIMITATIONS.....	49
11	FURTHER WORK.....	51
11.1	EXTENDING THE CLASS A PAN EVAPORATION SURFACE COLLECTION.	51
11.2	CORRECTING SYSTEMATIC ERRORS IN CLIMATE RECORDS.	51
11.3	EXTENDING THE RECORD FOR DATA-SPARSE PERIODS.	52
11.4	ERROR DETECTION.....	53
12	CONCLUSIONS.	55
13	REFERENCES.....	57

14	APPENDIX A: THE CLIMARC STATIONS.....	59
15	APPENDIX B: NON-CLIMARC STATIONS WITH LONG TEMPERATURE RECORDS.....	61
16	APPENDIX C: LONG TERM AVERAGE SURFACES.....	62
17	APPENDIX D: REPRESENTATIVE NETWORK AND COMPARISON DETAILS.....	63
17.1	GENERAL COMMENTS.....	63
17.2	NETWORK SELECTION.....	63
18	APPENDIX E: DAY-TO-DAY VARIATIONS IN THE NRM&E CLIMATE SURFACES.....	64
18.1	MAXIMUM TEMPERATURE.....	64
18.2	MINIMUM TEMPERATURE.....	66
18.3	SOLAR RADIATION.....	68
18.4	VAPOUR PRESSURE.....	70
19	APPENDIX F: THE DISCONTINUITY IN THE NRM&E SOLAR RADIATION SURFACES DUE TO THE CONVERSION OF CLOUD OBSERVATIONS FROM TENTHS TO OKTAS.....	72
19.1	SUMMARY.....	72
19.2	INTRODUCTION.....	72
19.3	SIMULATION METHODOLOGY.....	72
19.3.1	<i>Estimating distribution of 9am/3pm cloud tenths.</i>	73
19.4	RESULTS.....	74
19.5	FURTHER WORK.....	76
19.6	CONCLUSION.....	76

1 Acknowledgements.

We would firstly like to thank Greg McKeon, NRM&E, for his most generous assistance exploring the issues raised in this report, and commenting extensively on the draft. We also thank Pete DeVil, Queensland Department of Primary Industries and Fisheries, for making the Windows version of TAMET available, and for checking preliminary gap-filled data files for errors and inconsistencies. We thank Neville Nichols, Bureau of Meteorology Research Centre, for making the high quality annual temperature dataset of Torok and Nicholls 1996 available, and Blair Trewin, Bureau of Meteorology National Climate Centre, for providing the TABS monthly temperature dataset. Thanks also to Christina De Vries, NRM&E, for assistance with the solar radiation analysis, and Ken Brook, NRM&E, for editorial assistance.

Our often wide-ranging discussions with Steve Crimp, Neil Flood and John Carter, NRM&E, were also extremely helpful.

“CLIMARC - Computerising the Australian climate archives” (QPI43) Nick Clarkson, Queensland Centre for Climate Applications (QCCA)/DPI&F, was a joint project between the Climate Variability in Agriculture Program (CVAP), the Queensland Department of Primary Industries and Fisheries (DPI&F), the Queensland Department of Natural Resources, Mines and Energy (NRM&E), and the Australian Government Bureau of Meteorology.

The Climate Variability in Agriculture R&D Program (CVAP) was funded by the Australian Government, and part of the Agriculture Advancing Australia package. The program was managed for the Department of Agriculture Fisheries and Forestry - Australia by Land & Water Australia (LWA). In addition to LWA, the program was supported by Grains, Rural Industries, Sugar and Dairy R&D Corporations. The Managing Climate Variability R&D Program replaced CVAP in 2003.

2 Summary.

The Queensland Department of Natural Resources, Mines and Energy (NRM&E) climate surface collection¹ has been extended to cover the period 1889-1956, incorporating data from the CLIMARC project².

All climate surfaces were derived from spatial interpolation of ground-based observation data onto a 0.05° latitude/longitude grid. Daily surfaces are now available for rainfall, maximum and minimum temperature, solar radiation and vapour pressure, for the period 1889 – present. The reliability of the surfaces degenerates prior to 1910. Daily Class A pan evaporation surfaces are available from 1970 to present.

The maximum and minimum temperature, solar radiation and vapour pressure surfaces for the period 1889-1956 were computed using an anomaly-interpolation spline algorithm. Previously, long-term average climate surfaces were used as surrogate surfaces for this period. The new anomaly-interpolated surfaces represent historical daily and annual climate more accurately and realistically than the long-term average surfaces.

The generation of rainfall surfaces for the period 1890 – present has been described elsewhere.

Caveat: It is important that users of the NRM&E climate surfaces familiarise themselves with the errors, limitations, and assumptions described in this report. In particular, the anomaly-interpolated surfaces gradually revert to long-term average surfaces as the number of observations decreases in the early period of the climate record. Also, there are known to be systematic errors in the daily climate records used to compute the surfaces. **Consequently, the climate surfaces are not suitable for climate change studies.**

¹ Access the climate surfaces through the NRM&E SILO web-service: <http://www.nrme.qld.gov.au/silo/>

² CLIMARC “Computerising the Australian Climate Archives” was a joint project between the Climate Variability in Agriculture R&D Program (CVAP), Queensland Department of Primary Industries and Fisheries (DPI&F), Queensland Department of Natural Resources, Mines and Energy (NRM&E), and the Australian Government Bureau of Meteorology.

3 Overview.

In Australia, the Bureau of Meteorology is the custodian of a vast quantity of data, collected from a variety of instruments and locations. For many commonly required climate variables³, most of the archive data prior to 1957 are stored only as paper records. The considerable cost of extracting, computerising and quality-controlling these data means they are unavailable for climate research.

The "Computerising the Australian Climate Archives" (CLIMARC) project has computerised the daily climate records for 51 key climate locations across Australia. CLIMARC was a joint project between the Climate Variability in Agriculture Program (CVAP), the Queensland Department of Primary Industries and Fisheries (DPI&F), the Queensland Department of Natural Resources, Mines and Energy (NRM&E), and the Bureau of Meteorology.

The increased record length of computerised Australian pre-1957 daily climate data provided by the CLIMARC project will improve climate risk assessment in Australia.

The CLIMARC data, together with other pre-1957 daily climate data, have been spatially interpolated by NRM&E. The production of gridded surfaces makes the pre-1957 daily climate data available for a wide variety of applications.

Daily maximum and minimum temperature, solar radiation and vapour pressure surfaces for the period 1889-1956 were computed using an anomaly-interpolation spline algorithm. Consistent daily Class A pan evaporation data were not available before 1970, and hence no interpolations were performed. Australian Sunken Tank data have not been analysed. The derivation of daily station solar radiation and vapour pressure followed the Jeffrey *et al.* (2001) methodologies. The construction of the post-1957 component of the NRM&E daily climate archive, and the generation of daily and monthly rainfall surfaces, was described by Jeffrey *et al.* (2001).

The anomaly-interpolation spline method involved interpolating deviations from long-term means, rather than interpolating climate data directly. The use of the anomaly-interpolation spline method is appropriate because the resulting climate surfaces:

- represent actual weather,
- reflect inter-annual climate variability,
- do not contain unrealistic values, even when there were few station observations,
- preserve the natural inter-variable correlations in the climate datasets.

An important aspect of the method was to insert zero-anomalies into data-sparse regions prior to interpolation. This prevented minor gradients in the available data causing the surfaces to diverge wildly when extrapolated over large distances into data-sparse regions. A further consequence is that the anomaly-interpolated surfaces gradually revert to long-term average surfaces as the number of observations decreases in the early period of the climate record. This reversion does not happen uniformly across Australia. Instead, the surfaces represent actual daily climate where there were nearby climate observing stations, and long-term averages in areas where there were insufficient data to interpolate.

There are systematic errors in the observed data used to compute the NRM&E climate surfaces, caused by such factors as station relocations or instrumental changes. Because of these errors, and because the anomaly-interpolated surfaces gradually revert to long-term average surfaces in the early period of the climate record, the climate surfaces are not suitable for climate change studies.

³ In this report, rainfall is specifically excluded from the list of "climate variables".

3.1 Climate change trends.

This report documents the generation of Australia-wide, daily-climate surfaces from spatially-sparse observation records. Such daily climate surfaces have not previously been available to the community.

We recognise, however, that the climate data for Australia are likely to include the effects of:

1. decadal and interdecadal variability in global sea surface temperatures and atmospheric mean sea level pressures (Power *et al.*, 1999),
2. long-term trends due to long-term natural climate variability (>60-80 years), and human-induced climate forcings (e.g. greenhouse gas concentration, stratospheric ozone depletion, land use change), and
3. changes in observation procedure and instrumentation.

In this report, we do not attempt to estimate the magnitude of these effects. Hence, at this stage the NRM&E daily climate surfaces should not be used for climate change studies, or other trend analyses.

4 Example climate data: plots and surfaces.

Example data and anomaly-interpolated surfaces for 15 January, 1925, are presented to give readers some sense of what the CLIMARC network and anomaly-interpolated surfaces look like.

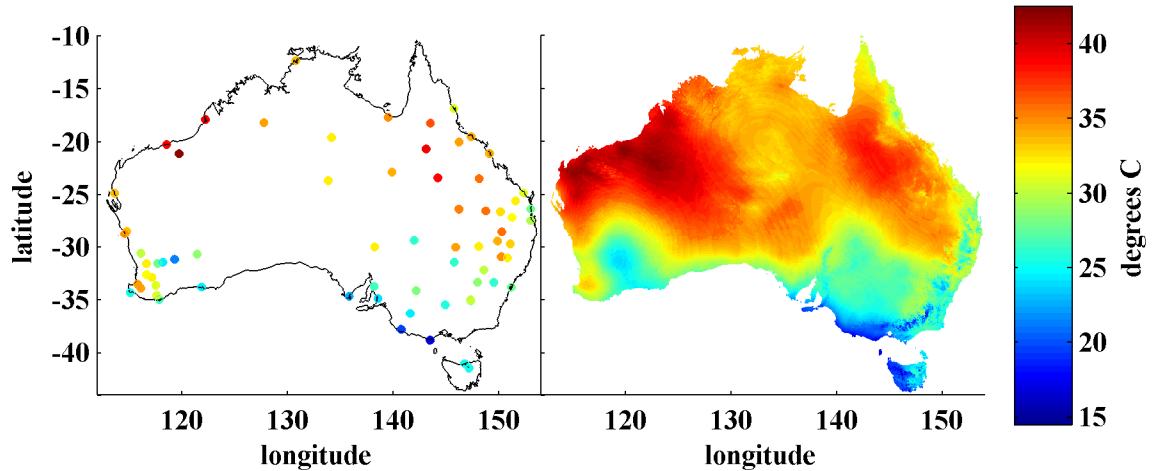


Figure 4.1: Maximum temperature observations (left) for the example date 15 January, 1925, and corresponding anomaly-interpolated surface (right).

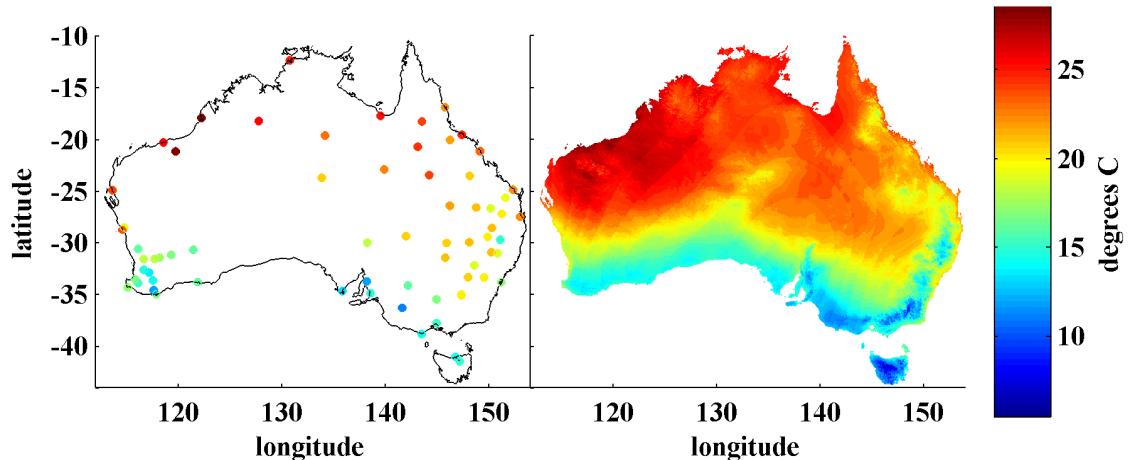


Figure 4.2 Minimum temperature observations (left) for 15 January, 1925, and corresponding anomaly-interpolated surface (right). Note the surfaces represent the low temperatures associated with the highland areas of the Snowy Mountains, even though there are no nearby CLIMARC network observing stations.

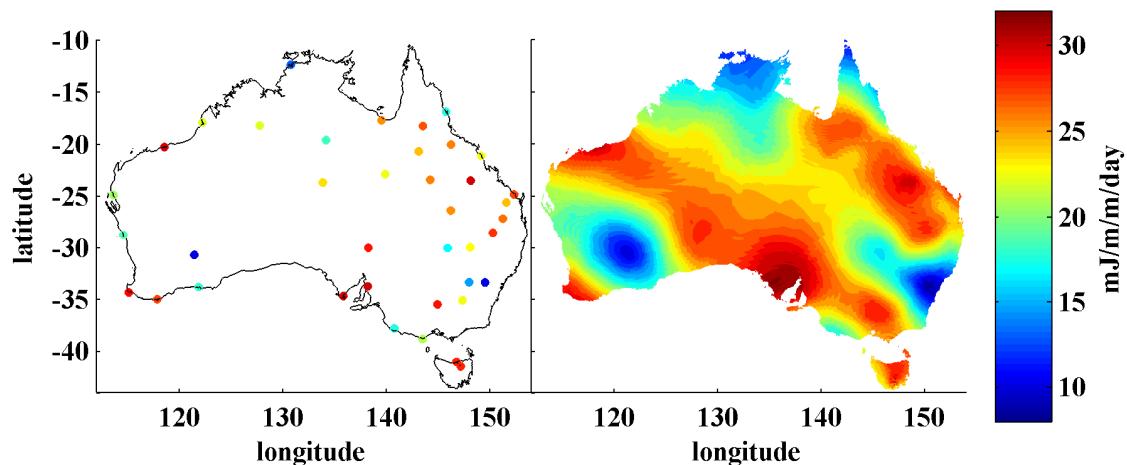


Figure 4.3: Daily solar radiation observations (left) for 15 January, 1925, and corresponding anomaly-interpolated surface (right).

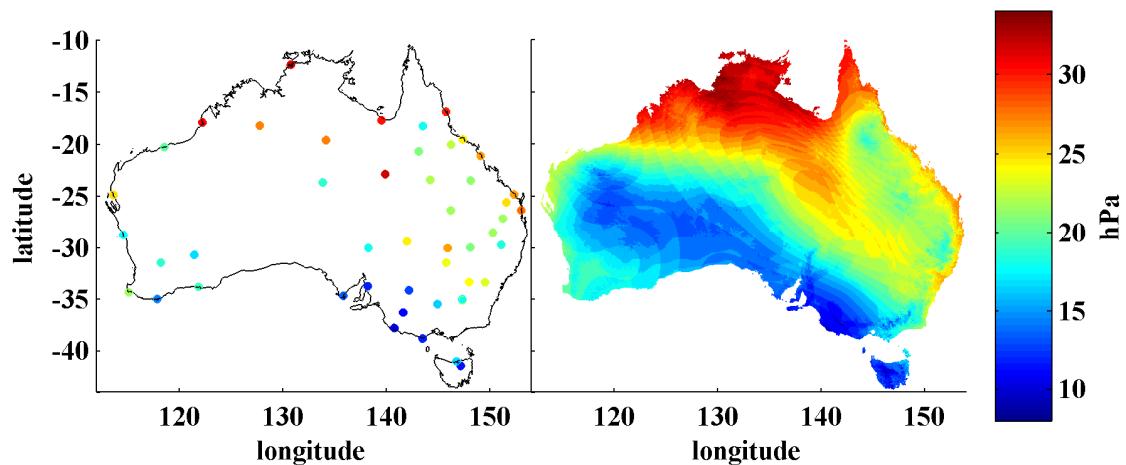


Figure 4.4: Vapour pressure observations (left) for 15 January, 1925, and corresponding anomaly-interpolated surface (right). The anomaly-interpolated vapour pressure surfaces have a numerical-resolution of 1 hPa, which causes the appearance of loops in the displayed image.

5 Review.

Climate risk affects us all. Through its effects on agriculture, water supplies, diseases, pest numbers, tourism, erosion, salinity and countless other factors, our climate has a direct impact on how we manage our economy and our society.

Managing climate risk requires climate knowledge. The recent CLIMARC project has significantly improved the knowledge of Australia's historical daily climate. The spatially-interpolated NRM&E daily climate surfaces, which provide data of significant value to the scientific and agricultural communities, have been extended using the CLIMARC data.

5.1 Observed climate records.

Historical, surface-observed climate data are an essential component of our climate knowledge, and are the principal record of our climate for the era before remote sensing technology.

In Australia, the Bureau of Meteorology is the custodian of a vast quantity of surface-observed climate data, collected from a variety of instruments and locations. Note that, in this report, rainfall is specifically excluded from the list of "climate variables", and the processing of rainfall data is not discussed. Nearly all Bureau of Meteorology rainfall data are available in digital format, and rainfall is processed and interpolated by NRM&E using completely different techniques.

The usefulness of the daily climate archive prior to 1957 is limited, because most of the data are stored only as paper records (Bureau of Meteorology, 1999). The considerable cost of extracting, computerising and quality-controlling these data means they are unavailable for climate research.

The lack of readily available pre-1957 daily climate data reduces our ability to manage climatic risks: it reduces our knowledge of climate extremes; long-term computer models of agriculture and hydrology are less reliable; rare events are not well understood; and our ability to characterise natural climate variability, and to differentiate natural climate variability from induced climate change, is limited.

The CLIMARC project has computerised the daily climate records for 51 key climate locations across Australia. The CLIMARC project represents the most systematic effort yet undertaken to computerise the Australian pre-1957 daily climate archive, and the increase in the record length of the computerised data will provide important improvements in Australian climate risk assessments.

The pre-1957 CLIMARC data, plus the smaller number of pre-1957 daily climate data which were already computerised, are referred to in this report as the "CLIMARC network" (Jones and Trewin, 2002).

5.2 Spatial climate datasets.

Spatial modelling applications require gridded climate datasets (which are estimates of climate variables over large areas) rather than observed climate records. Spatial modelling is required where other spatially-varying factors, such as topography, vegetation, or land-use, must also be included in a computer model. Gridded climate datasets are often constructed by *spatial interpolation* of climate observations. Interpolation methods range in complexity, from simple nearest-neighbor algorithms, up to massive global climate simulation models.

Another important use of spatial interpolation is in the derivation of synthetic climate time-series for modelling applications. Such time-series can be used at locations where there are no observed climate records, or to fill the gaps in incomplete records. All real observation records are incomplete; equipment breaks, observers may be unwell, stations are closed down, and no single-station record goes back to the beginning of recording. In fact, in some cases synthetic time-

series are preferred over observed climate records, because the spatial interpolation algorithm “smooths” out local, site-specific environmental effects.

NRM&E maintains a collection of daily time-step gridded climate surfaces, and provides synthetic and gap-filled climate time-series through the NRM&E SILO web-service⁴. This service has significant value to the scientific and agricultural communities. Prior to the availability of the service, projects which required continuous time-series data had to allocate considerable resources to developing datasets. Further, these datasets were often lost to the community at the conclusion of projects.

Initially, the NRM&E climate surfaces collection contained interpolated daily rainfall surfaces calculated for the period 1889-present, interpolated daily maximum and minimum temperature, solar radiation and vapour pressure surfaces for the period 1957-present, and interpolated daily Class A pan evaporation for the period 1970-present. In the construction of the synthetic and gap-filled time-series, if data were required for the period prior to the start of the interpolations, daily long-term average surfaces were used as a surrogate for interpolated surfaces. Although daily long-term averages do capture overall seasonal and spatial effects, they do not capture natural daily variability, long-term climate variability or climate change, or the naturally correlated variations between climate variables and rainfall.

Jones and Trewin (2002) investigated the likely error in interpolated daily maximum and minimum temperature surfaces for a range of historical station networks. These included the CLIMARC network, and the use of surrogate long-term averages. They used a network-analogue technique to generate hypothetical datasets that replicated the properties of pre-1957 station networks. From studying these hypothetical datasets, Jones and Trewin concluded that interpolated CLIMARC network daily maximum and minimum temperature surfaces would be a substantially better estimate of the daily temperature than surrogate long-term averages.

5.3 Extension of the NRM&E climate archive using CLIMARC.

This report describes the CLIMARC data, the anomaly-interpolation algorithm, and the important properties of the new climate surfaces.

The new spatially-interpolated surfaces have been incorporated into the NRM&E SILO web-service. This significantly improved the quality of the long-term records, by reducing the errors associated with using surrogate long-term averages.

The extension of the NRM&E daily climate surface collection significantly increases the value of the CLIMARC data to the community, and distribution of the improved data through the SILO web-service ensures it is readily accessible.

The new climate data provided by the CLIMARC project, the extension of the NRM&E climate surfaces, and the distribution of climate data through the SILO web-service, will improve the climate knowledge of the researchers, educationalists and decision makers in our community. This combination has the potential to significantly improve climate risk management in Australia.

⁴ The NRM&E SILO web-service: <http://www.nrme.qld.gov.au/silo/>

6 Climate observation data.

6.1 Summary.

The CLIMARC network refers to the stations for which computerised daily climate observations are available for the pre-1957 period. It contains typically 70 stations reporting daily maximum and minimum temperature, and 50 stations from which daily solar radiation and vapour pressure could be derived. The number of stations reporting, and the data quality, gradually decreases, particularly before 1910.

Preliminary checking of the CLIMARC network data has removed the most obvious errors. However, none of the data used to calculate the NRM&E climate surfaces have been corrected for systematic errors or inconsistencies.

6.2 Meteorological stations.

The NRM&E climate surfaces are computed using ground-based meteorological-station observations from the Bureau of Meteorology. The post-1957 climate records, and the daily rainfall records, were available from the Bureau of Meteorology prior to this project. Most of the pre-1957 climate records used were computerised by the CLIMARC project.

The number of stations reporting is shown for each climate variable in Figure 6.1. Vapour pressure is derived from wet/dry bulb temperature observations, and radiation from 9am/3pm cloud okta observations. The comparatively large quantity of computerised daily climate data in the post-1957 period is obvious, as is the gradual reduction in the number of stations in the network from 1910 back to start-of-record.

The CLIMARC dataset consist of hourly and daily records for 64 meteorological stations. The CLIMARC stations are listed in Appendix A. Because meteorological instruments are sometimes moved, the CLIMARC dataset effectively represents the climate history of 51 locations. For example, many weather stations in regional centres have been moved from post-offices to airports. As a result, stations have closed and new stations have opened⁵, and so a description of recorded climate at a location will usually require records from more than one station⁶.

The daily climate records from the Bureau of Meteorology also contain pre-1957 records from non-CLIMARC stations. There are 22 stations that have substantial daily maximum and minimum temperature records for the period 1910-1930, which are listed in Appendix B. The number of stations with temperature records increases after 1939.

The “CLIMARC network” refers to the stations for which computerised daily climate records are available for the pre-1957 period. The locations of the CLIMARC network stations are shown in Figure 6.2. Interpolation algorithms will produce more accurate values for locations close to recording stations, and the proximity to recording stations may be important in deciding whether to use observed, gap-filled, or synthetic time-series for a location.

⁵ Note that the CLIMARC documentation warns that “historically a nearby site (within about 1 mile in earlier days) may have used the same station number”.

⁶ Be aware that the CLIMARC data alone do not contain all the climate data for the 51 locations. For most locations, one or more of the stations which recorded climate data will only have recorded data in the post-1957 period. Because all the data for such stations were already computerised prior to the CLIMARC project, these stations were not included in the CLIMARC project.

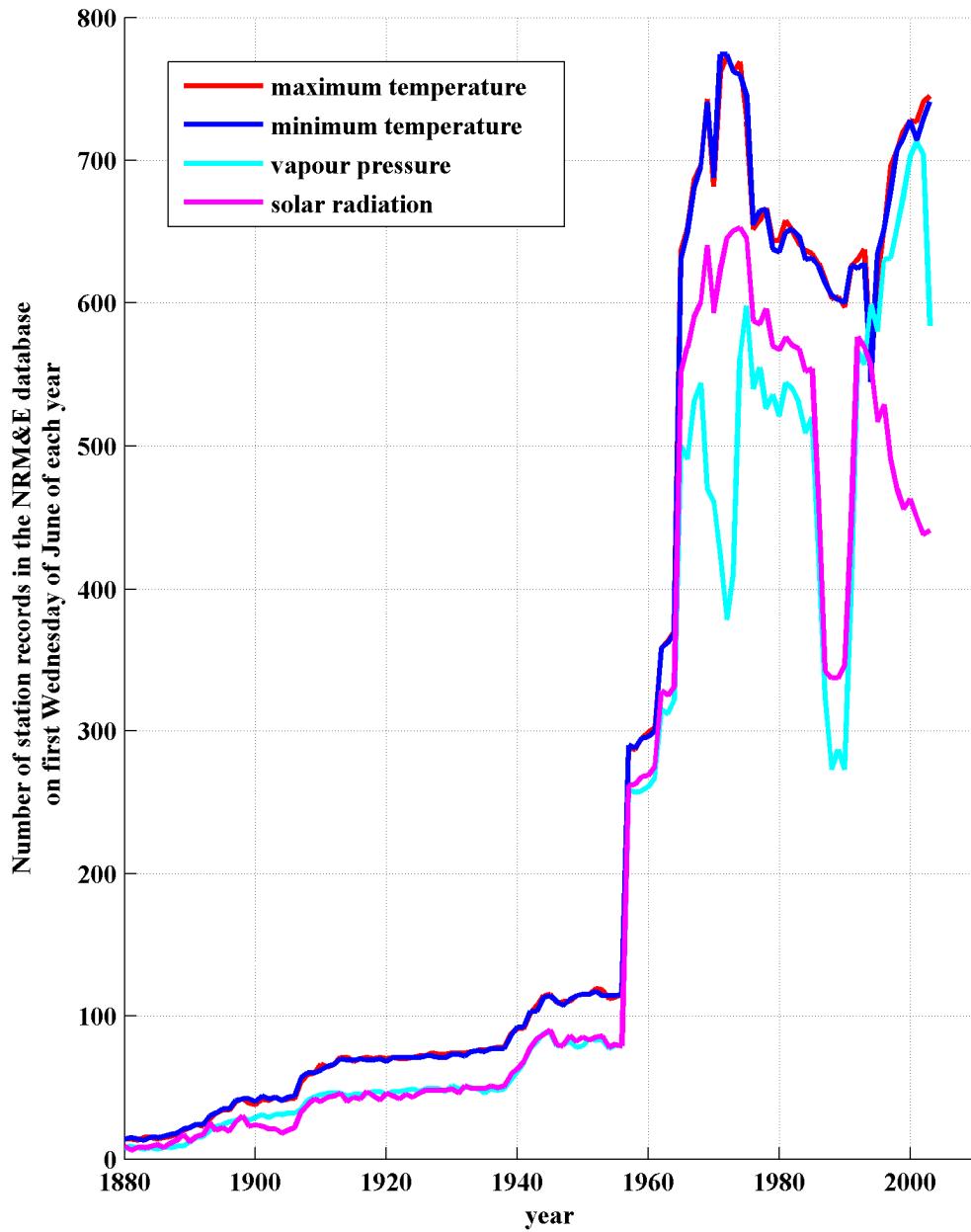


Figure 6.1: The historical number of stations reporting each climate variable. The vapour pressure is derived from wet/dry bulb temperature observations, and radiation from 9am/3pm cloud okta observations (Jeffrey *et al.*, 2001). The comparatively large quantity of computerised daily climate records for the post-1957 period is obvious, as is the gradual reduction in the number of stations in the network from 1910 back to start-of-record.

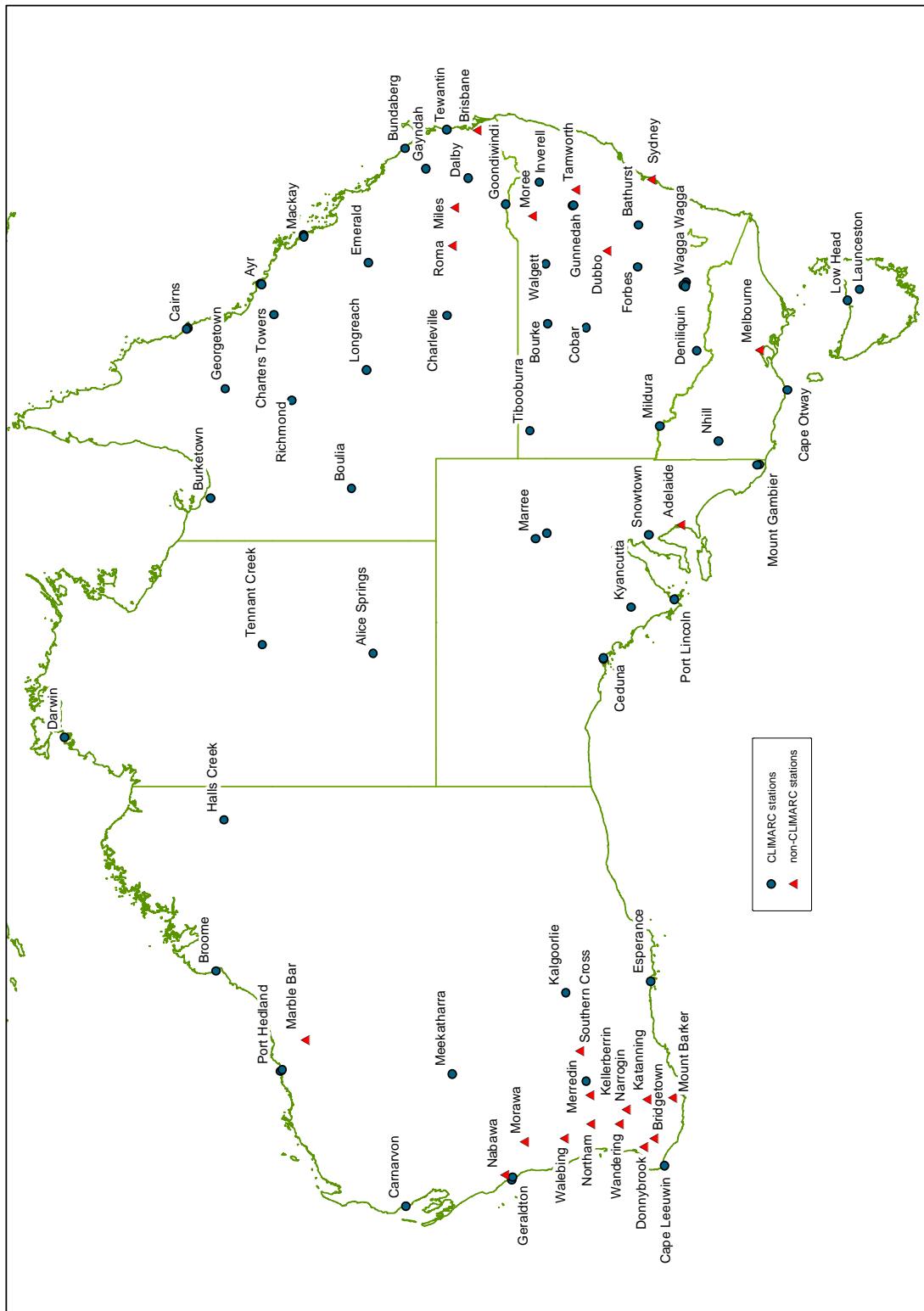


Figure 6.2: The locations of the observing stations in the CLIMARC network. For some locations (e.g. Wagga Wagga, NSW), there is more than one station in the CLIMARC record. The “non-CLIMARC stations” already had substantial computerised daily maximum and minimum temperature records for the pre-1957 period, prior to the CLIMARC project.

6.3 Pre-interpolation processing of climate data.

All recorded data for the selected stations were computerised for the CLIMARC project, and were made available by the Bureau of Meteorology in ASCII table format. So, in addition to commonly used climate variables such as temperature and precipitation, for some stations there may be observations for wind, visibility, “phenomena” (e.g. dust, hail, fog, thunder), etc. However, not all of the climate variables required for interpolation are directly observable. The required climate variables were generated from the observations as follows:

- Minimum and maximum daily temperatures are recorded directly by observers, and so could be used directly.
- Vapour pressure records were derived from wet-bulb/dry-bulb air temperature observations, and station-level air pressure,
- Daily solar radiation records were derived from total cloud amounts. Cloud amounts are observed using standard meteorological procedures, and were supplied in oktas⁷.
- Evaporation records were not used.

Details and references for the vapour pressure and solar radiation calculations are provided in Jeffrey *et al.* (2001).

6.3.1 Data checking.

Preliminary data checking was performed by inspection, using the TAMET error detection algorithms (Wall, 1977), and ad-hoc statistical analyses. These tests were sufficient to identify the most obvious errors.

Simple automatic error detection schemes were investigated, consisting of threshold tests on normalised and un-normalised daily data, and threshold tests on block-maxima/minima of monthly and annual data. These schemes were not sufficiently reliable to be run automatically, and the results were inspected visually before suspect data were excluded from the interpolations.

An exploratory comparison of the CLIMARC and “Big 12” temperature dataset (Stone *et al.*, 1996) did not reveal any major problems with the CLIMARC dataset. The Big 12 dataset contains independently computerised daily maximum and minimum temperature records for 12 locations in Queensland and NSW. Because both the CLIMARC and Big 12 datasets were sourced from the same paper records, any differences between the two datasets should be due to data-entry errors. However, the available Big 12 dataset was not well documented and did not contain any indication of the quality of individual records. The Big 12 dataset contained records which were flagged as accumulated⁸ or as poor quality in the CLIMARC dataset, and clearly contained patched data.

The comparison with the Big 12 dataset demonstrated that there are data-entry errors in the CLIMARC dataset, but that a formal comparison Big 12 dataset would not be a very efficient way of finding them. In a trial comparison of the pre-1957 data for Dalby (Station 041023), which has a long, high-quality observation record, there were 34 days where the CLIMARC and Big 12 maximum or minimum temperature observations differed by at least 5°C. Visual inspection of the data records for all stations for the dates in question suggested five of these

⁷ An okta is the number of eights of the sky obscured by cloud.

⁸ “Accumulated data climate data” refers to observations made after instruments were recording for more than one day. So, accumulated maximum temperature is the highest maximum temperature of two or more days, and accumulated minimum temperature is the lowest minimum temperature of two or more days. Accumulated rainfall, by way of contrast, would be the sum of the rainfall over the days.

differences were due to erroneous CLIMARC records, 17 were due to erroneous Big 12 records, while for the remainder the source of the error was not obvious.

Improved error detection is an area for future work. Even so, it will not be possible to devise statistical error detection schemes which can comprehensively identify data errors, because of the sparseness of the current CLIMARC network.

6.3.2 Systematic errors and discontinuities.

As discussed by Torok and Nicholls (1993), the Australian daily climate records have not been corrected for systematic errors or inconsistencies caused by such factors as station relocations, instrumental changes, or site changes (vegetation changes, new buildings, etc.). Researchers using the NRM&E climate surfaces should be aware of this issue, especially because a full record of the climate at a CLIMARC location will usually comprise data from more than one station.

Torok and Nicholls (1996) derived adjustments for a set of Australian annual-average maximum and minimum temperature station records. The adjustments for the annual records were typically 0.5°C-0.8°C, for the pre-1957 period. These adjustments indicate the typical level of systematic errors in the daily temperature records too, because the annual temperature records were derived by averaging monthly temperature records, which were in turn created by averaging daily temperature records. Correcting the climate data for systematic errors would be a valuable future project.

Cloud amounts are observed at 9am and 3pm local time, and are recorded as oktas, where an okta is the number of eighths of the sky obscured by cloud. However, prior to 1949, observations were actually made in *tenths* of the sky obscured by cloud (Jones and Henderson-Sellers, 1992). In the data supplied by the Bureau of Meteorology, the cloud amounts have been converted from tenths to eighths by multiplying by 0.8, and then rounding. The effect of this conversion is given in Table 1. Because of this conversion, cloud okta values of 2 and 6 are over-represented in the pre-1949 observations.

Table 1: Table used for converting cloud amount observations made in tenths (pre-1949) to cloud amounts in eighths (oktas). Because of this conversion, cloud okta values of 2 and 6 are over-represented in the pre-1949 observations.

Cloud amount in tenths	0	1	2	3	4	5	6	7	8	9	10
Cloud amount in eighths	0	1	2	2	3	4	5	6	6	7	8

Another potential source of discontinuities is caused by daylight-savings time. During daylight savings time, observations may have been made at local time, or standard time⁹. Daily maximum and minimum temperature observations will not usually be affected by the observation time, unless the minimum temperature occurred close to 9am. However, the daily vapour pressure (derived from 9am wet/dry bulb thermometer observations), and the daily solar radiation (derived from 9am/3pm cloud observations), will be affected.

The NRM&E daily pan evaporation surfaces were created from interpolating standard Class A pan evaporimeter records. Daily evaporation measurements are included in the CLIMARC dataset, but Class A pans were not in use in Australia at this time. The variety of measurement devices in use before the introduction of Class A pans, and their associated problems, were described by Hounam (1961). Early evaporimeters were generally sunken tanks, and initially these were completely non-standard. Even after the introduction of a standard Australian Sunken Tank following the formation of the Bureau of Meteorology in 1908, maintenance problems and

⁹ The observation practices under daylight savings time are described on the Bureau of Meteorology website at http://www.bom.gov.au/climate/averages/tables/dst_obs.shtml

inconsistent exposure lead to inconsistent and unreliable evaporation observations. Following the caution of Jeffrey *et al.* (2001), the CLIMARC pan evaporation data were not interpolated. Possible future approaches for estimating pan evaporation are discussed in Section 11.1.

7 Spatial interpolation of CLIMARC network data.

7.1 Summary.

NRM&E provides a collection of gridded climate surfaces, primarily intended for natural resource management and biophysical modelling applications. These climate surfaces are used in the generation of synthetic and gap-filled time-series for the SILO web-service.

Originally, for many applications including the construction of the synthetic and gap-filled time-series, daily long-term average surfaces were used as a surrogate for interpolated surfaces for the period prior to the start of the interpolations. The principal reason for interpolating the CLIMARC data was to overcome the problems associated with this practice. Analyses clearly demonstrated that it is better to use the anomaly-interpolated surfaces than long-term average surfaces.

7.2 Objectives.

Originally, where climate data were required and there were not enough data to spatially interpolate using the thin-plate spline algorithm, seasonally-dependent, long-term averages were used instead. Long-term averages are a poor representation of the real climate, in that they:

- do not capture natural daily variability,
- do not capture long-term climate variability or climate change,
- do not represent the natural inter-variable correlations in climate data.

A further problem occurs when long-term averages are used to fill gaps in incomplete observation records. If an observation is missing during a run of above- or below-average weather conditions, unrealistically large discontinuities can occur between the real observations and the gap-filling data.

Spatially-interpolated observation data can overcome these deficiencies. However, there is no guarantee that spatial interpolation will produce surfaces which represent daily climate better than long-term average surfaces! To be an improvement, the interpolated surfaces must:

- estimate the actual daily climate more accurately than long-term averages i.e. have a lower overall error,
- be reliable, in the sense that unrealistic values should not be produced from sensible inputs.

The anomaly-interpolated CLIMARC surfaces were assessed using these criteria, and were found to be an improvement over the surrogate long-term average surfaces. In addition, it was explicitly demonstrated that the anomaly-interpolated surfaces represented annual climate variability and natural inter-variable correlations. The analyses are discussed in Section 9.

It is important to note that spatial consistency and temporal consistency (e.g. in the relative weighting of stations, autocorrelation length scales, annual variability, etc.) were *not* considered primary criteria when assessing the interpolated CLIMARC surfaces. This was simply because the number of daily observations changes so dramatically over the period of the record. Enforcing complete spatial and temporal consistency could have been attempted using complex weather-generation style algorithms to “create” station time-series for the pre-1957 period. However, this would have been incompatible with the NRM&E philosophy of the providing climate data which are transparently related to real observations where possible. The alternative would have been to substantially dilute the information content of the surfaces, by enforcing a “cut-off” date for interpolation in the early climate record, and drastically reducing the number of stations used for interpolation in the post-1957 period. Clearly, neither of these options makes efficient use of the observation records.

7.3 Description of the anomaly interpolation algorithm.

Jeffrey *et al.* (2001) proposed “anomaly interpolation” as a scheme for the interpolation of sparse climate datasets. With this technique, the differences between observed values and long-term averages (the “anomalies”) were interpolated, rather than actual observed values.

The anomaly interpolation technique was as follows:

1. extract the values of the long-term average surface¹⁰, for the appropriate day-of-year, for the locations of the observed data,
2. form the anomalies, which are the differences between the observed data and the values of the long-term average surface,
3. insert local mean values (“zero-anomalies”) into data-sparse regions (see below),
4. interpolate the anomaly data using a thin-plate smoothing spline (see below), with latitude and longitude as independent variables,
5. add the long-term average surface to the interpolated surface from step 4, yielding the anomaly interpolated surface for the daily climate observations.

The anomaly-interpolation technique provides a simple method for incorporating the spatial climate patterns derived from the post-1957 station network into surfaces derived using the much sparser CLIMARC network. For example, the post-1957 climate surfaces were calculated using a trivariate thin-plate spline algorithm to incorporate the effect of terrain elevation (i.e. the lapse-rate). The CLIMARC network, however, is not dense enough to derive a daily, spatially-varying lapse-rate. Using anomaly-interpolating generates surfaces which represent the seasonally and spatially-varying, long-term average lapse-rate.

Local mean values of the long-term average surfaces (zero-anomalies) were inserted into datasets containing very data-sparse regions. For example, the CLIMARC network datasets prior to 1907 contain little or no observations in the western half of the continent. To prevent minor gradients in the available data from causing the surfaces to diverge wildly when extrapolated over large distances into data-sparse regions, a regular grid covering the continent was used to identify data-sparse regions, and a zero-anomaly was supplied for all grid cells that did not contain an observation. These values must not interfere with the surface in regions where observational data are available, so the regular grid used in the procedure must be low resolution. A five-degree grid was used.

The need to constrain the interpolation of sparse datasets using zero-anomalies is illustrated in Figure 7.1. The surface on the *left* was produced using a thin-plate spline algorithm; the surface on the *right* was produced using the anomaly interpolation algorithm. The surfaces were calculated using the daily vapour pressure records for 15 January, 1900. The locations of the recording stations are shown as open circles; there were no records for the southwest of Australia at this time. The presence of an anomalously-low record in central Australia has caused a steep gradient in the spline surface, which has resulted in unrealistically high vapour pressure values along the west coast. In contrast, in the data-sparse regions the anomaly-interpolated surface is essentially the same as the long-term average surface.

¹⁰ A description of the derivation of the NRM&E long-term average surfaces is given in Appendix C.

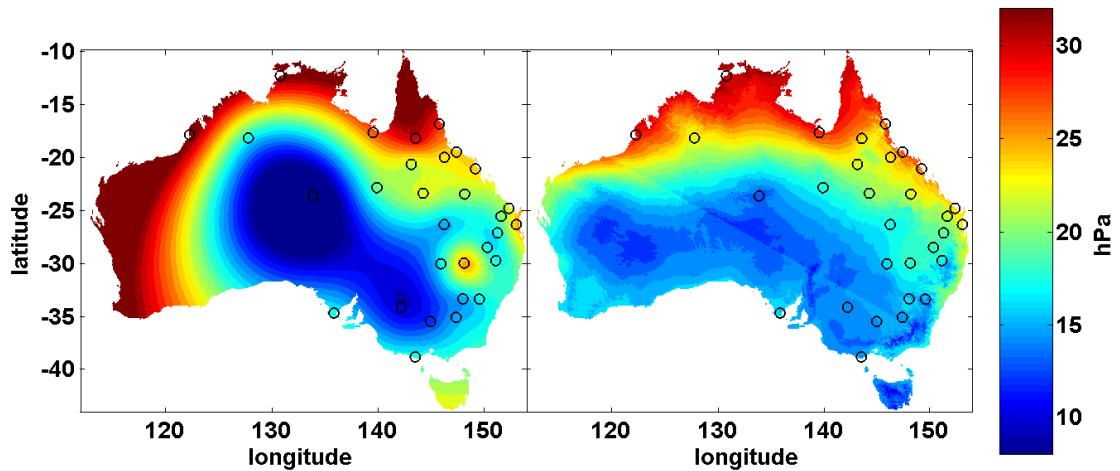


Figure 7.1: Vapour pressure surfaces for 15 January 1900, produced using the thin-plate spline algorithm (*left*) and the anomaly-interpolation algorithm (*right*). The locations of the recording stations are shown as open circles; there were no records for the southwest of Australia at this time. The presence of an anomalously-low record in central Australia has caused a steep gradient in the spline surface, which has resulted in unrealistically high vapour pressure values along the west coast. In contrast, in the data-sparse regions the anomaly-interpolated surface is essentially the same as the long-term average surface.

Clearly, there will be days where the departures from the long-term averages will be smooth and slowly-varying, and on such days inserting zero-anomalies into a dataset may reduce the accuracy of the interpolated surfaces. Also, the use of zero-anomalies makes the surfaces unsuitable for climate change studies, as discussed in Section 10.2. Tests clearly demonstrated, however, that overall the use of zero-anomalies provided a substantial improvement in the accuracy of the anomaly-interpolated climate surfaces.

The anomaly data were interpolated using a bivariate thin-plate smoothing spline (Wahba, 1979). Surfaces were fitting using the ANUSPLIN 4.0 package (Hutchinson, 1999). The degree of stiffness in the fitted surfaces was derived by minimizing the generalised cross validation (GCV) error (Wahba, 1990). This minimisation may fail when surfaces are generated from very few observations. In such cases, the ANUSPLIN 4.0 package generates a very simple, highly smoothed surface, usually with a single maximum or minimum. Inspection and statistical analysis demonstrated the anomaly-interpolation algorithm even provided an improvement over the use of long-term average surfaces in these cases. When data-availability is higher, anomaly-interpolation produces more complex surfaces, eventually becoming similar to surfaces derived using direct interpolation.

8 Overall properties of the anomaly-interpolated surfaces.

The principal reason for interpolating the CLIMARC observations was to overcome the problems associated with the use of long-term average surfaces as surrogate daily climate surfaces, as discussed in Section 7.2. The anomaly-interpolation spline algorithm does overcome most of these problems:

- The surfaces represent actual weather, as shown by the substantial reduction in the typical error of the surfaces, compared to errors from using long-term averages. The reduction in error has been demonstrated by cross-validation example (Section 9.1) and by representative network analyses (Section 9.2.2).
- The surfaces reflect inter-annual climate variability, as demonstrated by the comparison with independent time-series (Section 9.3).
- The surfaces preserve the natural inter-correlations within the climate datasets, as demonstrated by the effect of rainfall on climate variables in the Warrego River catchment (Section 9.4).
- The use of zero-anomalies prevents unrealistic values being generated for data-sparse regions. The reduction in the number of “large” errors was demonstrated as part of the representative network analysis (Section 9.2.3).

Users should be aware of the main limitations of the anomaly-interpolated surfaces:

- The input observation records have not been corrected for systematic errors or inconsistencies. Evidence for such systematic errors is discussed in Section 10.1
- The use of zero-anomalies in the interpolation algorithm (Section 7.3) means the anomaly-interpolated surfaces gradually revert to long-term average surfaces as the number of observations decreases in the early period of the climate record. The effect of using zero-anomalies in the interpolation is discussed in Section 10.2.

Because of these limitations, the NRM&E climate surfaces should not be used for climate change studies.

8.1 General comments on the reliability of early climate records.

The NRM&E climate archive data are used in a wide variety of applications. Whether the climate data are accurate enough to be considered “useable” will be entirely application-specific. For example, long-term average climate data are considered adequate for some rainfall-runoff models (when used with observed rainfall), whereas long-term averages are clearly useless for climate change studies. Further, the accuracy of the anomaly-interpolated surfaces varies both spatially and temporally. These factors preclude giving a definitive start date for “useable” data.

Even so, some general comments should assist potential users of the CLIMARC records and anomaly-interpolated surfaces.

- Jones and Trewin (2002) note the raw maximum and minimum temperature records prior to 1900 are subject to substantial uncertainty. Also, the number of station records decreases steadily prior to 1900.
- There are very few daily maximum or minimum temperature observations in Western Australia prior to 1907. See Appendix E.
- The quality of the early CLIMARC wet/dry bulb thermometer readings are presumably subject to the same uncertainty as the minimum and maximum temperature readings prior to 1900, and again there are no observations for southwestern Australia prior to 1907.
- The quality of the interpolated solar radiation, as determined from the distribution of observing sites, was the subject of an independent study (not presented). The quality of the anomaly-interpolated surfaces was found to revert rapidly to the quality of the long-term averages prior to 1910.

9 Improvements in the climate surfaces.

Comprehensive statistical analyses clearly demonstrated the anomaly-interpolated surfaces were superior to surrogate long-term average surfaces. The surfaces had a lower overall error, and so were a more accurate estimate of the actual daily climate. There was a reduction in the number of “large” errors, confirming that the algorithm was not generating unrealistic values from sensible inputs. Also, the interpolated surfaces were shown to represent the inter-annual climate variability and natural inter-variables correlation.

9.1 Improved estimate of observed daily climate: cross-validation analysis.

The example maximum temperature time-series shown in Figure 9.1 demonstrates that the anomaly-interpolated surfaces represent actual daily climate variability, whereas long-term averages do not. The data are for Longreach Post Office (station 36030) for 1925. The blue dots show the observed values, the red line shows cross-validation values; and the green line shows the long-term average values for Longreach. The cross-validation values were calculated by anomaly-interpolation, but *without* using the Longreach observations in the interpolation. Without the Longreach observations, the nearest station was 300km away. The cross-validated, anomaly-interpolated time-series is clearly a much better representation of the observed time-series than the long-term averages. It can also be seen from Figure 9.1 that the cross-validated time-series underestimates the most extreme departures from the long-term averages. This is a well-known issue with thin-plate smoothing spline interpolation, especially in regions of low data-density.

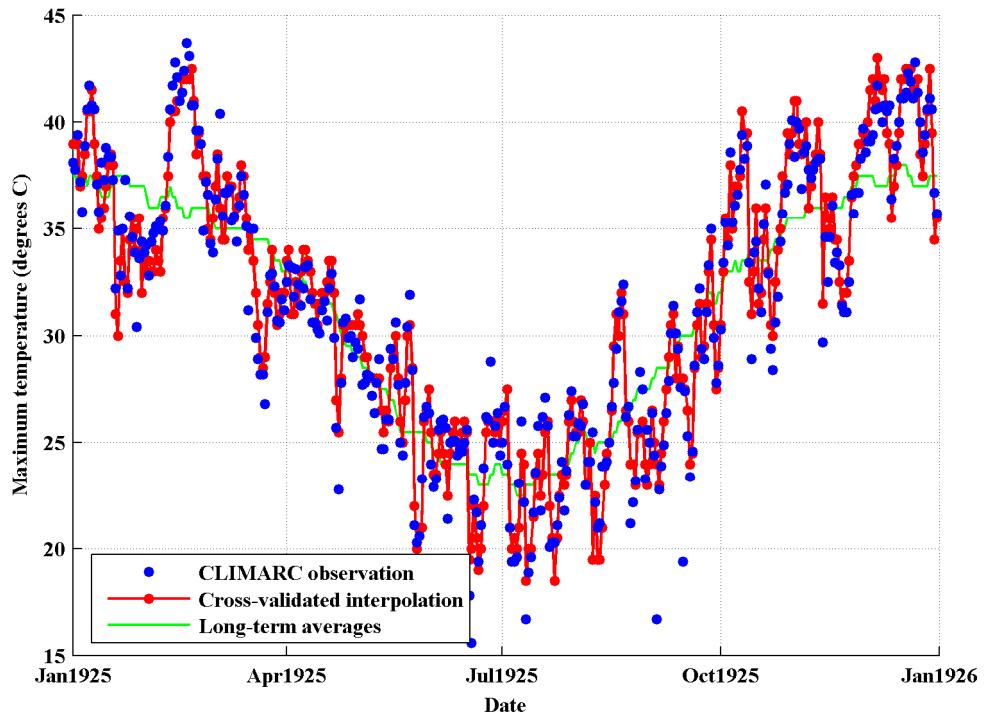


Figure 9.1: Longreach Post Office (36030) maximum temperature for 1925. Blue dots show CLIMARC observations; the red line shows cross-validation values (calculated by anomaly-interpolation *without* using the Longreach observations as input); and the green line shows the long-term average values for Longreach. The anomaly-interpolated surfaces represent daily weather, unlike long-term averages.

The mean-absolute-error for the surrogate long-term average time-series is 2.4°C, and the mean-absolute-error for the cross-validated time-series is 1.0°C. Although a decrease in the mean-

absolute error of 1.4°C may not sound significant, Figure 9.1 shows that it actually implies an important improvement, in that the interpolated surfaces represent natural daily climate variability.

9.2 Improved estimate of observed daily climate: a representative network analysis.

A representative network analysis verified that the anomaly-interpolated surfaces represent observed daily climate better than long-term average surfaces. The representative network analysis used post-1957 observation records to estimate the likely errors in pre-1957, anomaly-interpolated surfaces. This result applied to interpolated maximum temperature, minimum temperature, daily solar radiation and daily vapour pressure. The analysis also demonstrated that anomaly-interpolations gave fewer “large” errors, and illustrated the spatial variability of interpolation errors.

9.2.1 Representative station network.

Section 9.1 showed an example of how anomaly-interpolated surfaces represent daily climate variability, using cross-validated data. Cross-validation is a very useful technique for estimating the typical accuracy of interpolated surfaces, because cross-validation error estimates can be obtained from the input data alone, and because they incorporate the effects of data errors. However, cross-validation has two significant disadvantages: it is very computationally intensive, and it only gives an estimate of the interpolation error for locations where there is a historical observation.

An alternative technique that can be used for incomplete datasets is *representative network analysis*. Jones and Trewin (2002) used this technique to investigate the likely error in interpolated daily maximum and minimum temperature surfaces for a range of historical station networks, which included the CLIMARC network. The analysis described here involved selecting a set of stations and observation records from the existing, post-1957 climate database that were *representative* of the CLIMARC network. Anomaly-interpolated surfaces were then created using only this sub-set of stations. Comparing the records for the unused (*reserved*) stations with the interpolated surfaces gave an estimate of the error in the surfaces. For comparison, the same reserved stations were used to obtain an estimate of the error resulting from using surrogate long-term average surfaces.

Representative network analysis was an appropriate technique, because there were many more stations reporting daily climate post-1957 than in the CLIMARC network. The difference has been illustrated in Figure 6.1. For example, for the period 1970 to 1990, there were typically 700 stations reporting daily maximum and minimum temperature, 500 stations from which daily solar radiation was derived, and 500 stations from which daily vapour pressure was derived. In comparison, there were only approximately 70 stations in the CLIMARC network with maximum and minimum temperature records, and only 50 stations from which daily solar radiation and vapour pressure were derived.

Details of the representative network analysis are given in Appendix D.

9.2.2 Improved overall accuracy.

To get an overall picture of the accuracy of the anomaly-interpolation algorithm, the results were averaged across all reserved stations, and over the entire analysis time-period. The overall mean-absolute-errors are shown in Table 9.1, and the errors were clearly lower for the anomaly-interpolated surfaces than for long-term average surfaces, for all four climate variables.

Table 9.1: Overall mean-absolute-errors in anomaly-interpolated surfaces and surrogate long-term average surfaces, derived from representative network analysis. The anomaly-interpolated surfaces were significantly more accurate.

	Long term averages mean-absolute-error	Anomaly-interpolation mean-absolute-error
Maximum temperature	2.7 °C	1.4 °C
Minimum temperature	2.6 °C	1.6 °C
Solar radiation¹¹	11.8 %ET	7.7 %ET
Vapour pressure	2.8 hPa	1.9 hPa

Note that, because the long-term average surfaces were determined using post-1957 data, the use of long-term average surfaces to directly estimate pre-1957 climate data is likely to give significantly *larger* errors in practice than indicated by the analysis presented here.

9.2.3 Improved reliability.

The anomaly-interpolated surfaces were also more reliable than the surrogate long-term average surfaces, in that the number of estimates of reserved station observations with “large” errors was drastically reduced. This is demonstrated by Table 9.2, which shows the percentage of estimated reserved station data that had errors of 5 °C or more for maximum and minimum temperature, of 20% ET or more for solar radiation, and of 5hPa or more for vapour pressure.

Table 9.2 The percentage of estimates with “large” errors, for anomaly-interpolated surfaces and surrogate long-term averages, derived from representative network analysis. There were significantly less “large” errors when using anomaly-interpolated surfaces, demonstrating the improved reliability of the anomaly-interpolated surfaces.

	Long term averages	Anomaly interpolation.
Maximum Temperature with error ≥ 5 °C.	15%	2%
Minimum Temperature with error ≥ 5 °C.	13%	2%
Solar radiation with error ≥ 20 %ET.	15%	8%
Vapour pressure with error ≥ 5 hPa.	17%	6%

9.2.4 Spatial variation of interpolation errors.

The errors in the interpolated surfaces were not uniform. For example, errors would generally be higher where there was no observing station nearby, in areas of complex terrain, or where local weather phenomena such as sea breezes were significant.

The mean-absolute-errors, from the representative network analysis, are shown for each reserved station in Figure 9.2 to Figure 9.9. The mean-absolute-errors for surrogate long-term average

¹¹ The solar radiation was analyzed in units of percentage extra-terrestrial radiation (%ET), rather than $MJ\ m^{-2}\ day^{-1}$. This meant that the analyses were not biased by the seasonal and latitudinal components of the solar radiation surfaces.

surfaces are also shown for comparison. For all of maximum temperature, minimum temperature, solar radiation and vapour pressure, it is easy to see that the errors in the anomaly-interpolated surfaces were significantly lower over most of Australia.

The major improvements occurred in inland regions. There was less improvement along the coast, presumably because the ocean significantly moderates extremes in climate, and because local weather phenomena limit the precision of interpolation algorithms in coastal areas. Similarly, anomaly-interpolation provides little statistical improvement in far-north Australia, where there is little daily variation in weather for much of the year.

The errors in the maximum temperature, minimum temperature and solar radiation anomaly-interpolated surfaces are highest in southeast Australia. This is probably due to a combination of two factors. Firstly, the CLIMARC network is very sparse in coastal southeast Australia, with only three coastal stations in all of Victoria and New South Wales (see Figure 6.2). Further, the representative station network used for this analysis was based solely on the CLIMARC stations, and did not include the pre-1957 stations already contained in the NRM&E climate archive. So in fact, the representative station network contained only one coastal station in Victoria, and none in NSW! Secondly, southeast Australia is more mountainous than inland Australia. Complex local weather patterns, which occur in mountainous regions, mean the interpolation errors for these regions will always be higher than for the relatively flat inland regions.

Maximum Temperature:

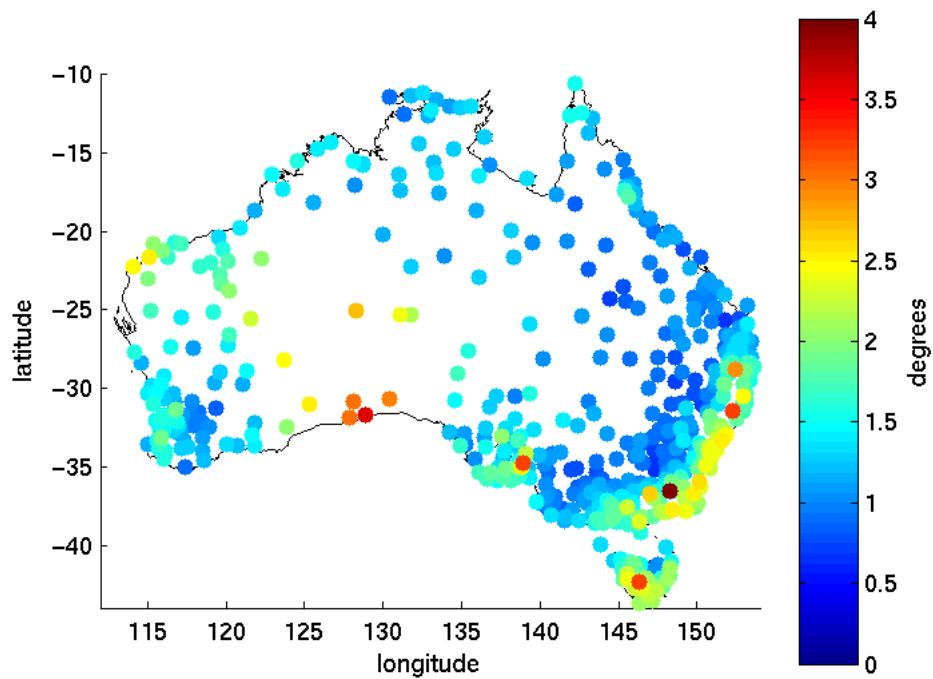


Figure 9.2: Mean-absolute-error using anomaly interpolation of representative stations to estimate maximum temperature.

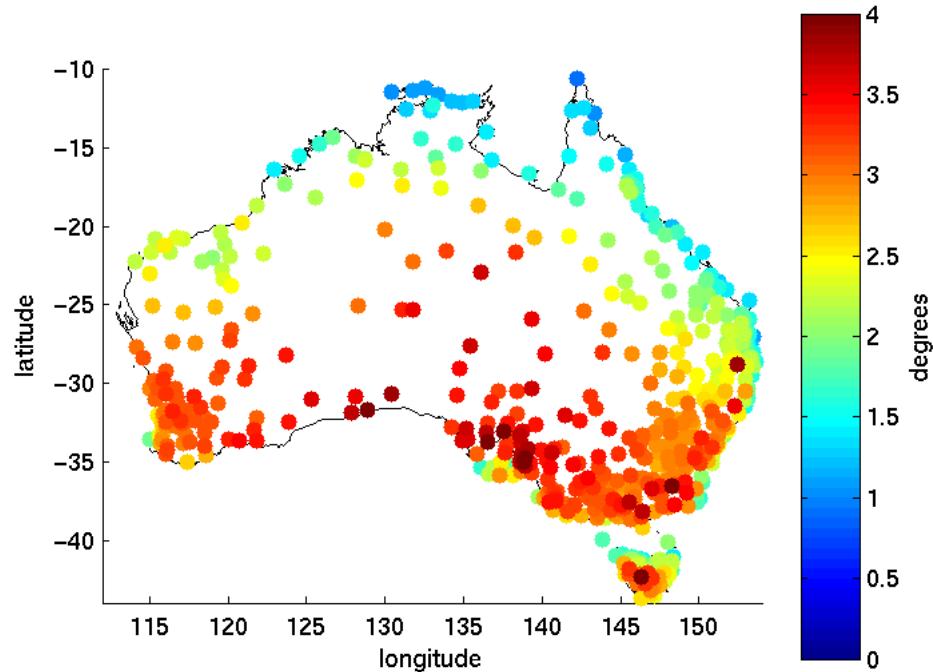


Figure 9.3: Mean-absolute-error using long-term averages to estimate maximum temperature.

Minimum Temperature:

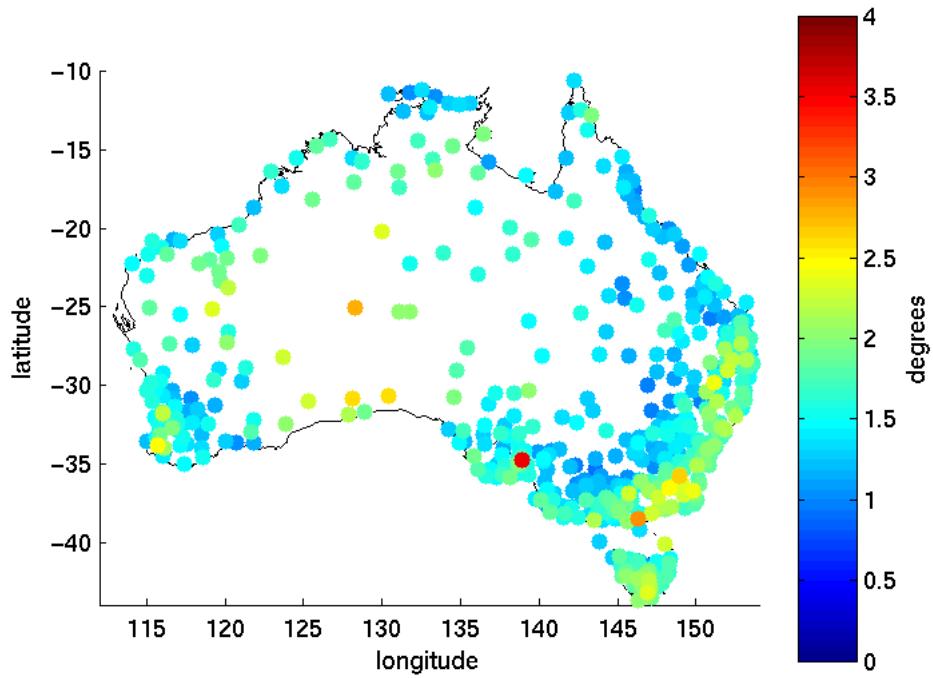


Figure 9.4: Mean-absolute-error using anomaly interpolation of representative stations to estimate minimum temperature.

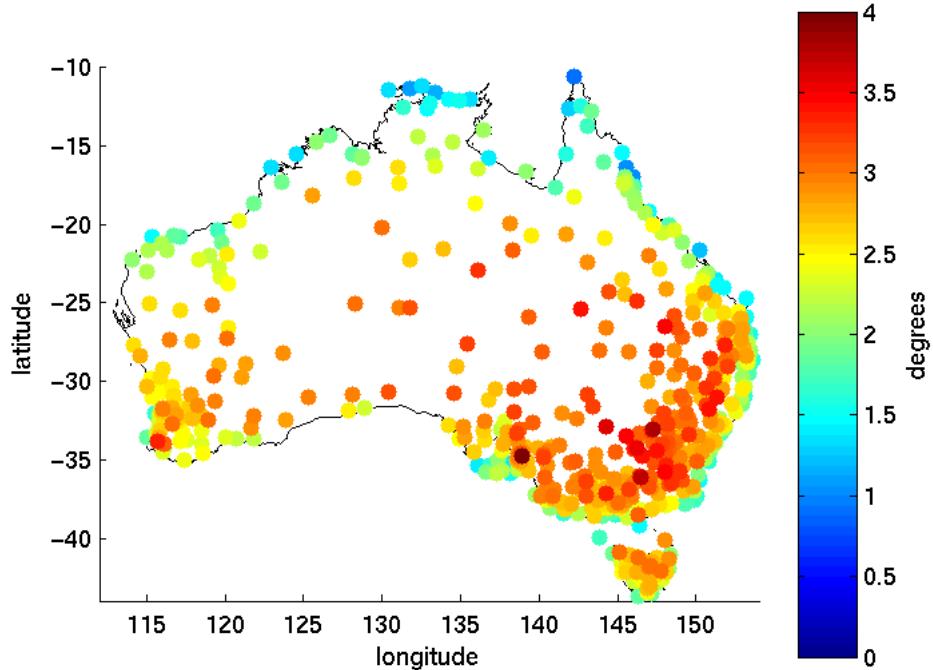


Figure 9.5: Mean-absolute-error using long-term averages to estimate minimum temperature.

Solar Radiation:

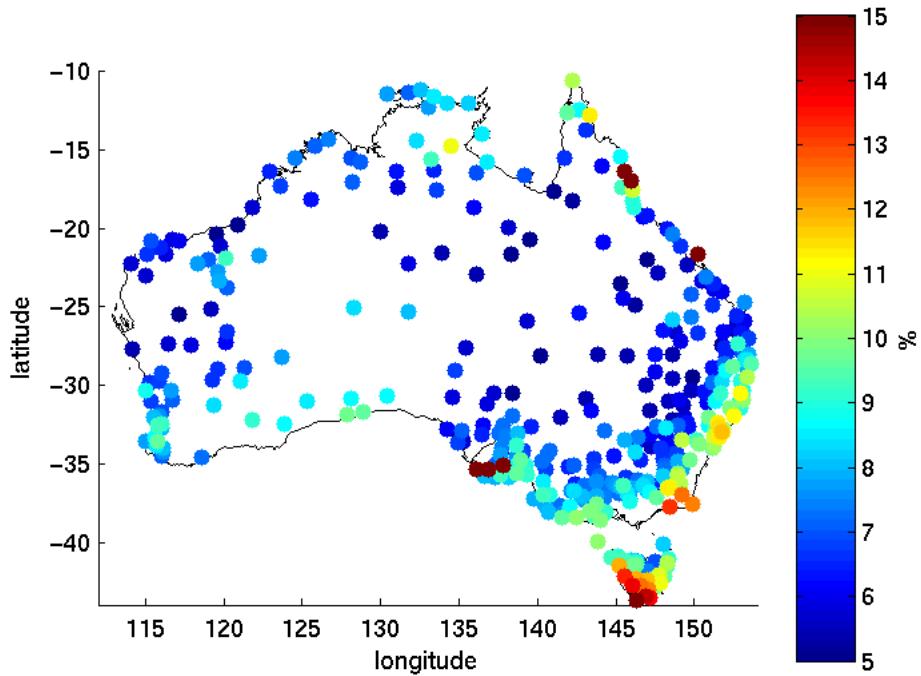


Figure 9.6: Mean-absolute-error using anomaly interpolation of representative stations to estimate %ET solar radiation.

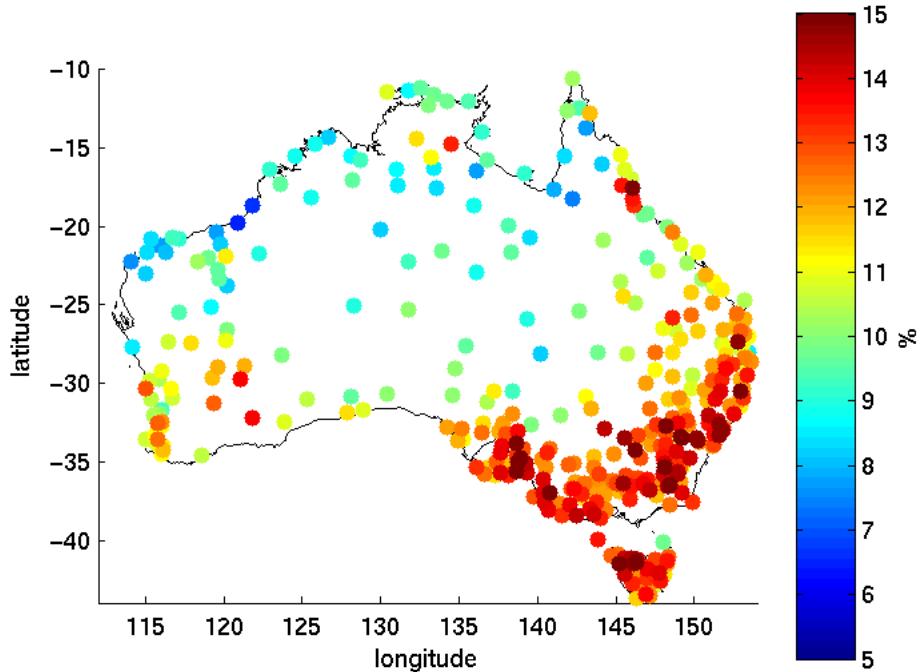


Figure 9.7: Mean-absolute-error using long-term averages to estimate %ET solar radiation.

Vapour Pressure:

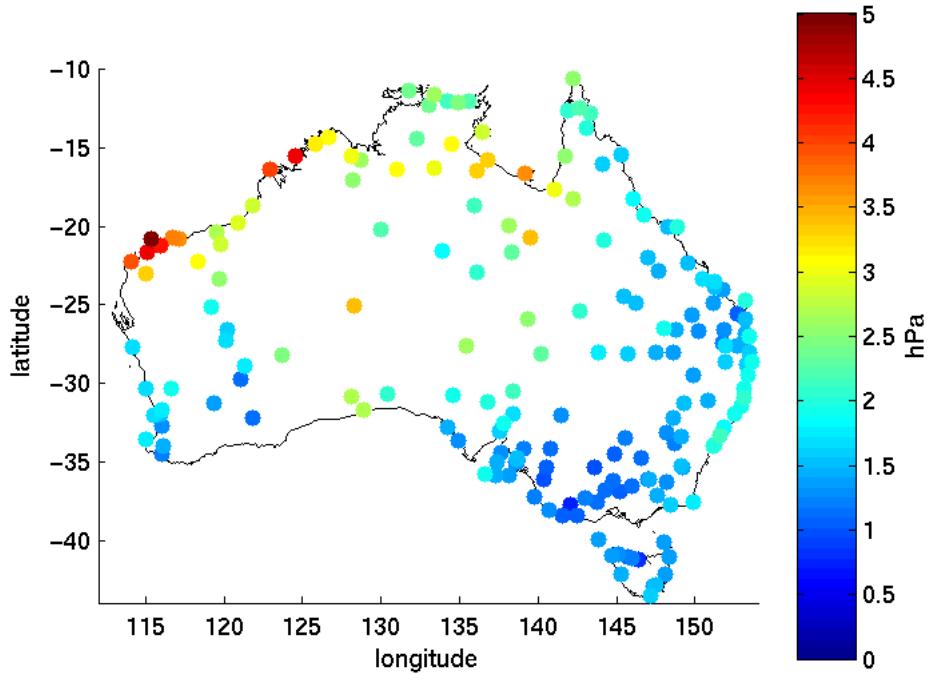


Figure 9.8: Mean-absolute-error using anomaly interpolation of representative stations to estimate vapour pressure.

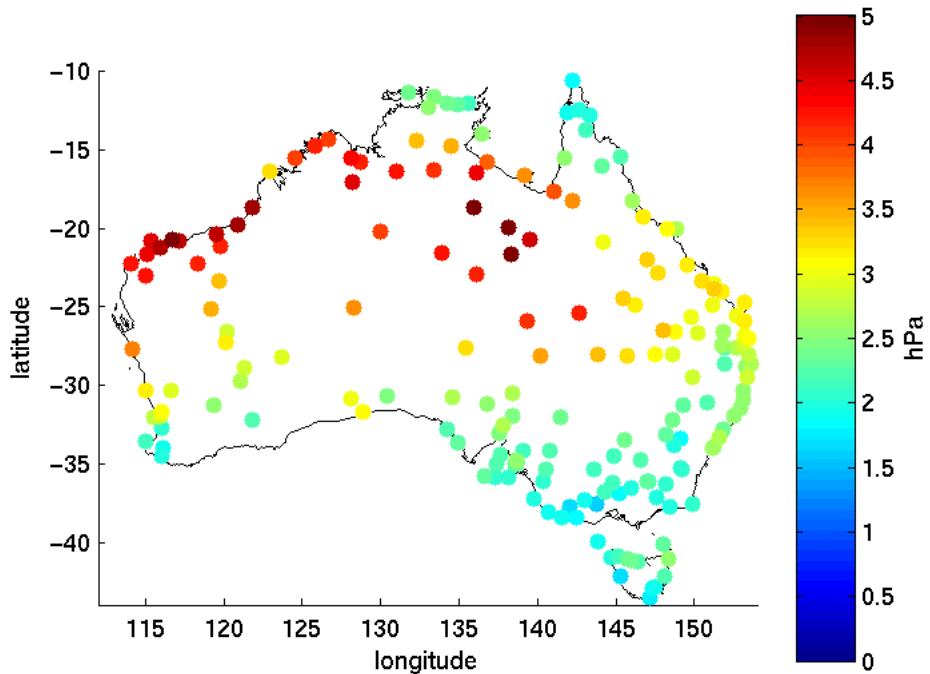


Figure 9.9: Mean-absolute-error using long-term averages to estimate vapour pressure.

9.3 Representation of inter-annual climate variability.

One of the problems with long-term averages is that they do not represent historical inter-annual climate variability. The anomaly-interpolated surfaces do represent inter-annual climate variability, as demonstrated by the correlations between:

- the inter-annual variations in minimum and maximum temperature with those from an independent, high-quality annual dataset,
- the inter-annual variations in dew point, calculated from the vapour pressure surfaces, and the inter-annual variations in minimum temperature, and
- the inter-annual variations in solar radiation and cloud amount.

9.3.1 Maximum and minimum temperature annual variability.

Torok and Nicholls (1996) produced time-series of annual mean maximum and mean minimum temperature for 224 Australian stations with long observation records. They derived the time-series from the Bureau of Meteorology's mean-monthly (TABS) datasets¹², and adjusted for discontinuities using statistical techniques together with station history documentation. The time-series is nearly complete from 1910 to 1993.

The NRM&E maximum and minimum temperature surfaces were compared against the Torok and Nicholls annual temperature time-series. The comparison series was created from the NRM&E climate surfaces by sampling the daily maximum and minimum temperature surfaces at the locations of the 224 Torok and Nicholls stations, and then averaging the data to create annual time-series. Both the NRM&E annual time-series and the Torok and Nicholls time-series were then averaged over all stations.

The maximum temperature comparison is plotted in Figure 9.10 and the minimum temperature comparison in Figure 9.11. The NRM&E time-series is composed of two components: the NRM&E spline surfaces for the post-1957 period, and the anomaly-interpolated surfaces for the pre-1957 period. These plots demonstrate that the anomaly-interpolated surfaces accurately represent inter-annual variations in maximum and minimum temperature at the Australia-wide scale. There are systematic differences between the time-series, however, which are discussed in Section 10.1.

The comparison of the Torok and Nicholls time-series with the NRM&E time-series for *individual stations* gave more diverse results. Generally, there was a good correlation between the inter-annual variations¹³ for both minimum and maximum temperature. However, for many stations the correlation between the actual time-series was poor. This implies that applying the Torok and Nicholls corrections significantly affects climate trend calculations for some stations. Because the observation records used to calculate the NRM&E surfaces have not had these corrections applied, the surfaces should not be used to calculate climate trends.

¹² For a discussion of the Bureau of Meteorology mean monthly temperature dataset, see Jones and Trewin (2002).

¹³ “Correlation between the variations” means the correlation between the time-series of successive differences, rather than the direct correlation of the time-series. For example, for a daily time-series, the time-series of successive differences is the value of each day, minus the value of the day before. The correlation between successive difference time-series measures the short time-scale relationship between variables, not the relationship between long time-scale trends.

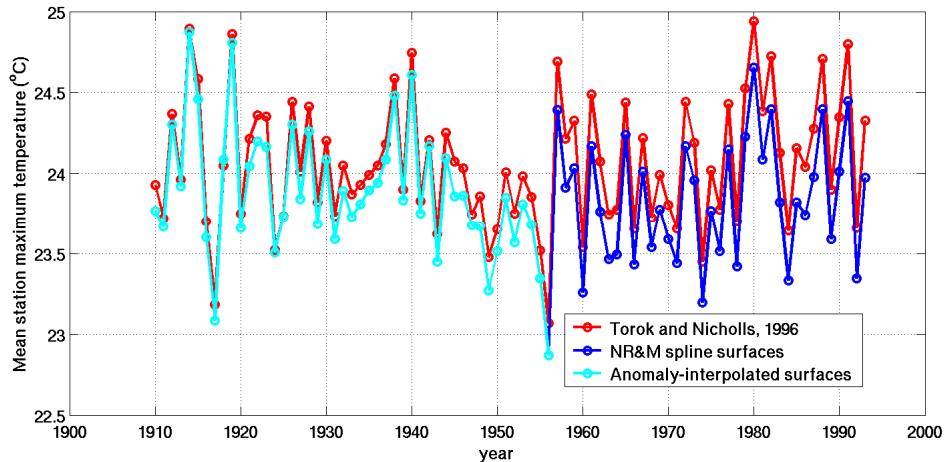


Figure 9.10: The mean maximum-temperature for the high-quality historical annual series from Torok and Nicholls (1996), time-series derived from the NRM&E spline surfaces, and time-series derived from the anomaly-interpolated CLIMARC network surfaces. The anomaly-interpolated surfaces accurately represent inter-annual variations, but the Torok and Nicholls series becomes systematically warmer than the NRM&E times-series.

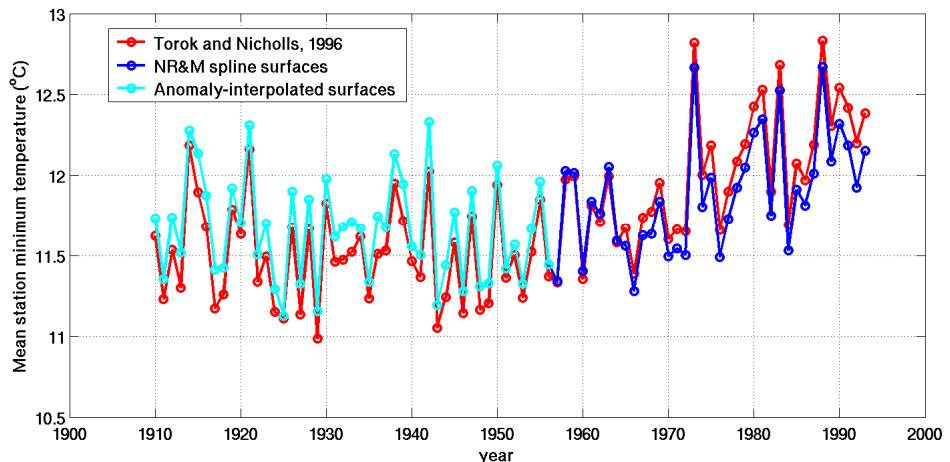


Figure 9.11: The mean minimum-temperature for the high-quality historical annual series from Torok and Nicholls (1996), time-series derived from the NRM&E spline surfaces, and time-series derived from the anomaly-interpolated CLIMARC network surfaces. Again, the anomaly-interpolated surfaces accurately represent inter-annual variations. The Torok and Nicholls time-series is initially cooler than the NRM&E time-series, but becomes systematically warmer.

9.3.2 Dew point calculated from vapour pressure.

On a calm, clear night, a meteorological “rule-of-thumb” is that the minimum temperature should not be significantly less than the dew point (Bohren and Albrecht 1998, pp 190). There are two processes that act to prevent temperature falling below the dew point. Firstly, when the temperature decreases to the point where condensation begins, the net removal of water-vapour from the air lowers the dew point. Secondly, the release of latent-heat of condensation acts to prevent the temperature falling lower.

The dew-point is related to the water vapour pressure, defined as (Emanuel, 1994):

$$\text{dew point} = \frac{243.5 \times \ln\left(\frac{\text{vapour pressure}}{6.2112}\right)}{17.67} - 1$$

The NRM&E vapour pressure surfaces are calculated from wet/dry bulb thermometer measurements (Jeffrey *et al.* 2001), and are independent of minimum temperature measurement. Comparing correlation between dew point and minimum temperature for the pre-1957 and post-1957 climate surfaces is thus a good test of the temporal consistency of the vapour pressure surfaces.

Synthetic time-series of minimum temperature and dew point were constructed by sampling the NRM&E minimum temperature and vapour pressure surfaces at the locations of the CLIMARC-network stations. For much of Australia, the relationship between minimum temperature and dew point in the post-1957 period was found to be quite poor. This is presumably because the air is sufficiently dry that the minimum temperature cannot fall to the dew point overnight by radiative cooling. However, there was a good correlation between variations in southeast Queensland, NSW and Victoria during May, June, and July. The May-June-July dew point and minimum temperature averages, calculated from the synthetic time-series, were used for this analysis. Only those locations for which the correlation in the variations was $r^2 > 0.65$ in the post-1957 section of time-series (21 locations) were used.¹⁴

Time-series of annual variations were calculated for the selected locations by subtracting the mean of the entire time-series. The use of mean-subtracted time-series removed the offset between the minimum temperature and dew point. Finally, the mean-subtracted time-series were averaged to produce a single eastern Australian, May-June-July dew point time-series, and a matching minimum temperature time-series.

The average minimum temperature and dew point time-series are shown in Figure 9.12. The high correlation which was known to exist between the post-1957 time-series is clearly maintained in the pre-1957 time-series. The inter-annual variations in the pre-1957 time-series are well-correlated, which demonstrates that the anomaly-interpolated vapour pressure surfaces represent inter-annual variability, which the long-term average surfaces do not.

The differences between the mean minimum temperature and dew point time-series are shown in Figure 9.13. This plot shows the dew point regression residuals, derived from regressing the dew-point time-series shown in Figure 9.12 onto the minimum temperature time-series. The relationship is generally consistent over the period 1910 to 2000. The changes in the relationship prior to 1910, and especially prior to 1900, may be due to the increasing sparseness of the early CLIMARC network. There is no discontinuity at 1957, which demonstrates the consistency

¹⁴ We would have liked to examine the correlation between Torok and Nicholls (1996) minimum temperature time-series and the dew point calculated using NRM&E surfaces. However, the Torok and Nicholls time-series is an annual time-series, and monthly time-series were required to form the May-Jun-July averages.

between the pre-1957 anomaly-interpolated vapour pressure surfaces and post-1957 vapour pressure spline surfaces.

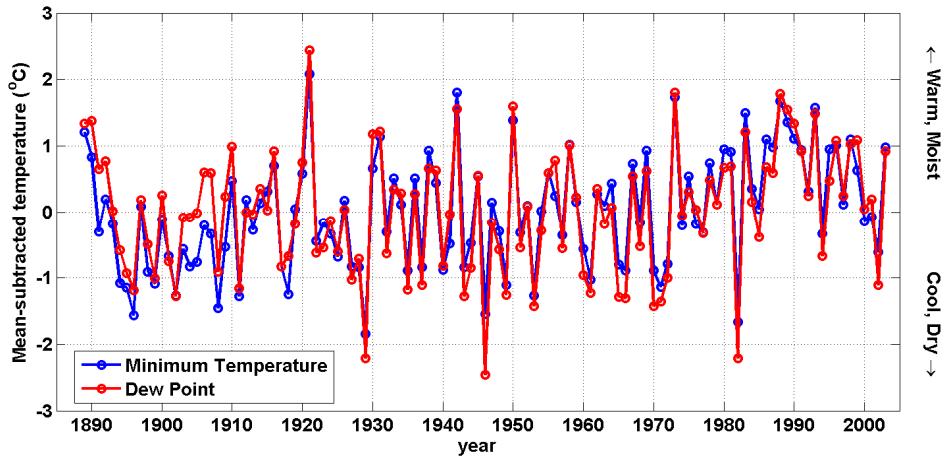


Figure 9.12: Annual variations in May-June-July minimum temperature and dew point time-series averaged over selected eastern Australian locations. The time-series were constructed from sampling the NRM&E surfaces at locations that had a high correlation between minimum temperature and dew point in the post-1957 period. The inter-annual variations in the pre-1957 time-series are well-correlated, which demonstrates that the anomaly-interpolated vapour pressure surfaces represent inter-annual variability.

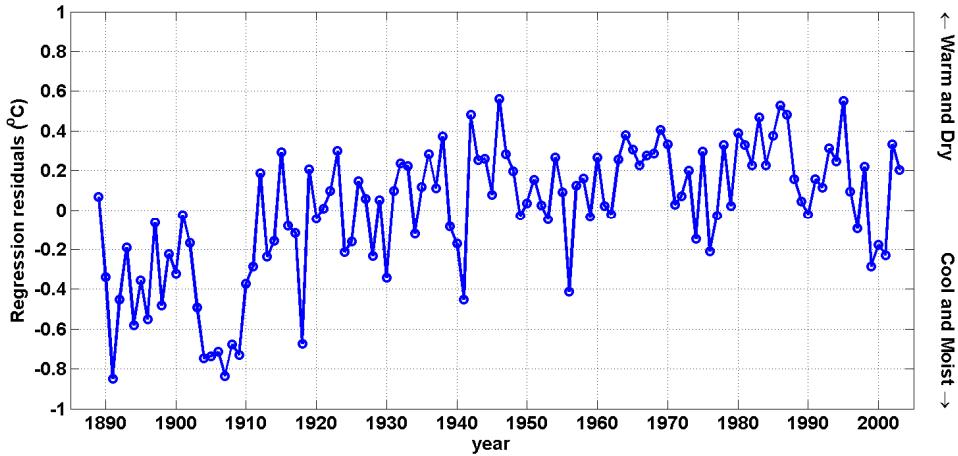


Figure 9.13: Dew point regression residuals derived from regressing the dew-point time-series shown in Figure 9.12 onto the minimum temperature time-series. The differences prior to 1910, and especially prior to 1900, may be due to the increasing sparseness of the early CLIMARC network. There is no discontinuity at 1957, which demonstrates the consistency between the pre-1957 anomaly-interpolated vapour pressure surfaces and post-1957 vapour pressure spline surfaces.

9.3.3 Solar radiation and cloud amount.

Solar radiation is highly correlated with cloudiness¹⁵, and in fact the NRM&E solar radiation surfaces are calculated from cloud okta observations using this relationship (Jeffrey *et al.*, 2001). Historical Australian cloud records were analyzed by Jones and Henderson-Sellers (1992). Their analysis used monthly-average cloud observations for 318 stations, from the Bureau of Meteorology's TABS dataset. The NRM&E solar radiation surfaces were compared to the average cloud data presented by Jones and Henderson-Sellers. This comparison provided independent confirmation that the NRM&E solar radiation surfaces do represent actual historical inter-annual variability.

Monthly average cloud time-series from the TABS dataset, averaged over all stations, are given in Jones and Henderson-Sellers (1992, Figure 1). The time-series is a “simple” average of the mean-subtracted station time-series.

A comparison time-series was created from the NRM&E surfaces. The NRM&E surfaces are calculated using cloud observations from stations which record both 9am and 3pm cloud oktas. Firstly, synthetic time-series were created by sampling the daily solar radiation surfaces at the locations of those CLIMARC-network stations which record both 9am and 3pm cloud oktas. The daily time-series were averaged to annual time-series, the annual mean was subtracted from each time-series, and finally an Australian time-series was calculated by averaging all the station time-series.

Calculating the Australian time-series from a “simple” average of synthetic mean-subtracted time-series was appropriate because the spatial distribution of the CLIMARC-network stations is similar to, although much sparser than, the distribution of stations used by Jones and Henderson-Sellers. Further, Jones and Henderson-Sellers state that time-series created using simple averaging, and time-series created using area-weighted averaging, were almost indistinguishable.

The annual solar radiation and cloud time-series are shown in Figure 9.14. Note the inverted y-scale for the cloud time-series. The line at 1949 indicates where the observing convention changed from tenths to oktas, as discussed in Section 6.3.2. Although it is clear that the solar radiation surfaces are systematically lower in the post-1949 period than in the pre-1949 period, it is not possible to determine from Figure 9.14 alone whether this represents a climate trend, or a discontinuity associated with the change in observation practice. This issue is investigated in more detail in Sections 10.1.3 and 10.1.4. The inter-annual variations in the solar radiation and cloud time-series in Figure 9.14 are well-correlated, however, providing independent confirmation that the NRM&E solar radiation surfaces do represent actual historical inter-annual variability.

¹⁵ Of course, the correlation between solar radiation and cloudiness is negative i.e. solar radiation is high when cloudiness is low.

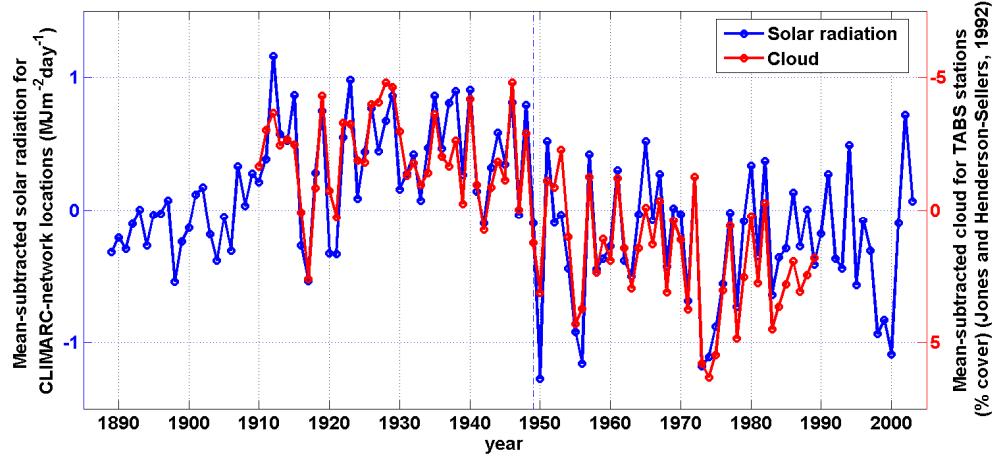


Figure 9.14: Mean-subtracted, annual solar radiation and cloud time-series. The solar radiation time-series was calculated by sampling the NRM&E solar radiation surfaces at the locations of the CLIMARC-network stations. The cloud time-series was calculated from 318 stations from the TABS dataset by Jones and Henderson-Sellers (1992, Figure 1). Note the inverted y-scale for the cloud time-series. The line at 1949 indicates where the observing convention for cloud amounts changed from tenths to eighths. The inter-annual variations in the solar radiation and cloud time-series are well-correlated, providing independent confirmation that the NRM&E solar radiation surfaces do represent actual historical inter-annual variability.

9.4 Inclusion of inter-variable correlations.

Many biophysical models use more than just one climate variable, or just rainfall. These models will benefit from using anomaly-interpolated climate surfaces, because the surfaces preserve the natural inter-variable correlations in climate data. These correlations are not preserved in long-term average surfaces.

9.4.1 The value of inter-variable correlations.

The principal applications of the NRM&E climate surfaces are daily time-step pasture growth modelling and hydrological modelling. Within both applications there are good examples of specific processes which can be modelled more realistically using climate data with preserved inter-variable correlations.

In rainy conditions, pasture growth modelled using daily climate data is generally higher than pasture growth modelled using long-term average data. There are several factors that contribute to this. During rainy conditions, there is generally more cloud cover than the long-term average. The higher cloud cover increases the ratio of diffuse light to direct light, and plants use diffuse light more efficiently than direct light; there is less self-shadowing of leaves, for example. Also, the vapour-pressure is generally higher than the long-term average and the temperature lower. This leads to a lower-than-average vapour-pressure deficit and hence lower potential water-loss from leaves. Plants respond to such conditions, by opening stomata and increasing leaf-nitrogen for maximum photosynthesis and growth.

Similarly, in rainy conditions, runoff and drainage are generally higher when modelled using daily climate data than when modelled using long-term average data. Drainage is typically modelled to occur after initial rainfall has saturated a soil moisture store. In days after initial rainfall, the evaporation is likely to be lower than the long-term average, due to the factors outlined above. When models are run using long-term average data, the excess evaporation must

be replaced by follow-up rainfall before runoff or drainage occur, leading to potentially large systematic errors in the water balance calculations¹⁶.

9.4.2 *Inter-variable correlations in Warrego River catchment time-series.*

A time-series analysis of the Warrego River catchment, located in the upper Murray-Darling catchment, demonstrated that the inter-variable correlations are preserved in the anomaly-interpolated surfaces. Catchment-average daily rainfall, minimum and maximum temperate, solar radiation and vapour pressure were first generated for the period 1910 to 2003. Each day in the period was categorised according to the number of days that had passed since a day of 15mm catchment-average rainfall. For each rainfall category, the means of the pre-1957 and post-1957 climate variables were calculated using the days in that category. Finally, for each climate variable, the mean calculated using all days was subtracted from the category means. The results are presented in Figure 9.15.

Figure 9.15 demonstrates the correlations in the climate and rainfall time-series. For days with catchment-average rainfall of 15mm or more, mean maximum temperature was 4°C below average, mean minimum temperature was 2-3°C above average, mean vapour pressure was 3hPa above average, and solar radiation 6 MJ/m²/day below average. Over the week following rainfall, the climate variable means approached the overall average again.

Note that Figure 9.15 shows departures from the overall catchment means, calculated for the 1910 to 2003 period. The offsets between the pre-1957 and the post-1957 curves, which are most pronounced for minimum temperature, are presumably caused by systematic differences in the climate, or in the observation records, of the region between the two periods.

¹⁶ Note that model calibration may mitigate the effects of such errors.

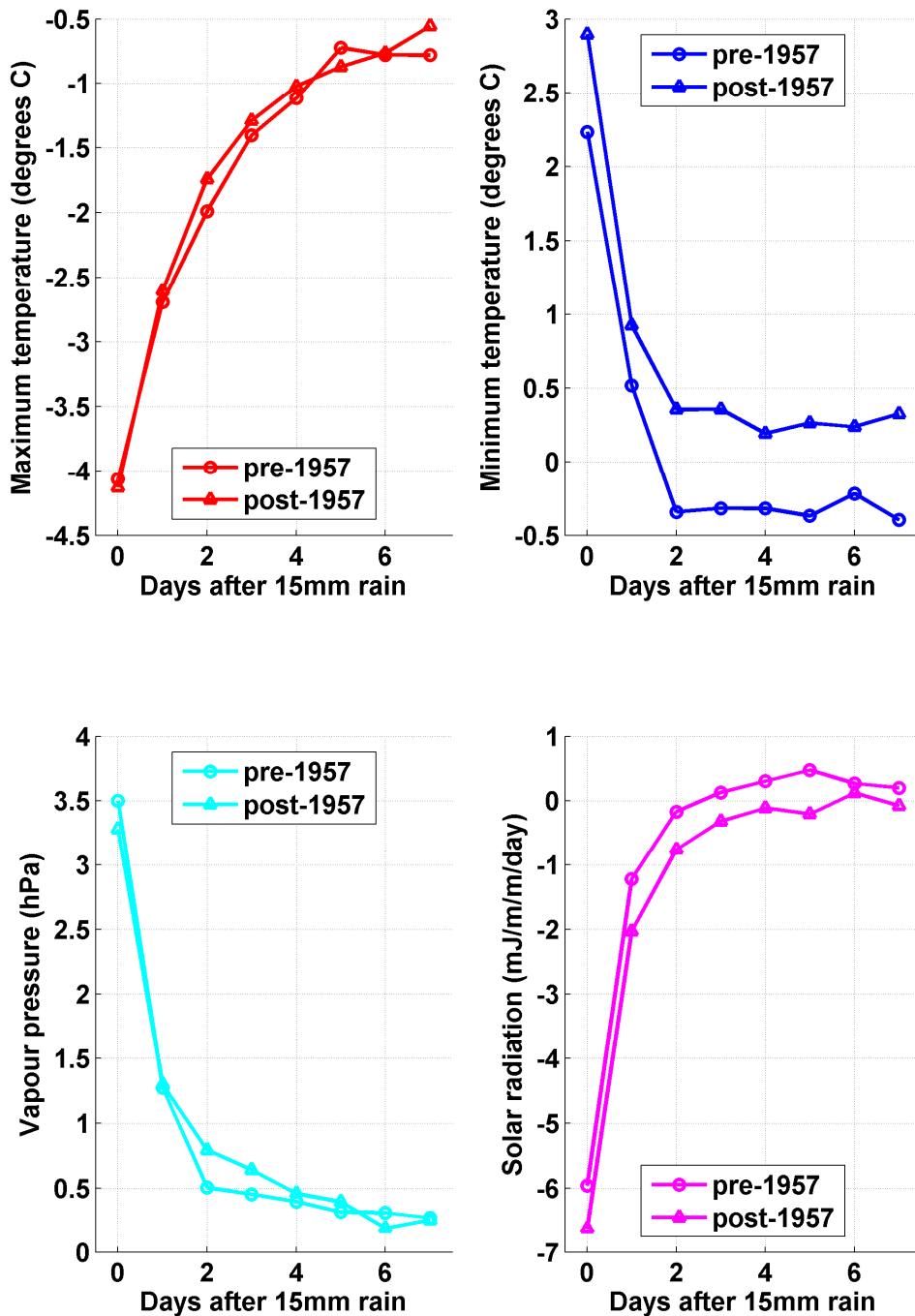


Figure 9.15: Pre-1957 and post-1957 mean-subtracted, category-average climate for the Warrego River catchment. The data were calculated by categorizing the daily, catchment-average climate for the period 1910-2003 according to the number of days that had passed since a day of 15mm catchment-average rainfall. The effect of rainfall on the climate demonstrates that correlations between the climate and rainfall time-series are preserved in the anomaly-interpolated surfaces. The offsets between the pre-1957 and the post-1957 curves are presumably caused by systematic differences in the climate, or in the observation records, of the region between the two periods.

10 Limitations of the anomaly-interpolated climate surfaces.

The NRM&E interpolated surfaces do not accurately represent climate change. This is due to two factors:

- the input observation records have not been corrected for systematic errors, and
- the anomaly-interpolated surfaces gradually revert to long-term average surfaces as the number of observations decreases in the early period of the climate record.

10.1 Systematic errors in observation records.

10.1.1 Inaccurate trends in the maximum and minimum temperature climate surfaces.

The comparison between the Torok and Nicholls (1996) annual maximum and minimum temperature time-series and the NRM&E surfaces was described in Section 9.3.1. The differences between the time-series are plotted in Figure 10.1. The differences are systematic, and there are clear trends. The effect of these differences is to incorrectly *reduce* the climate change trends in the NRM&E time-series compared to the Torok and Nicholls time-series.

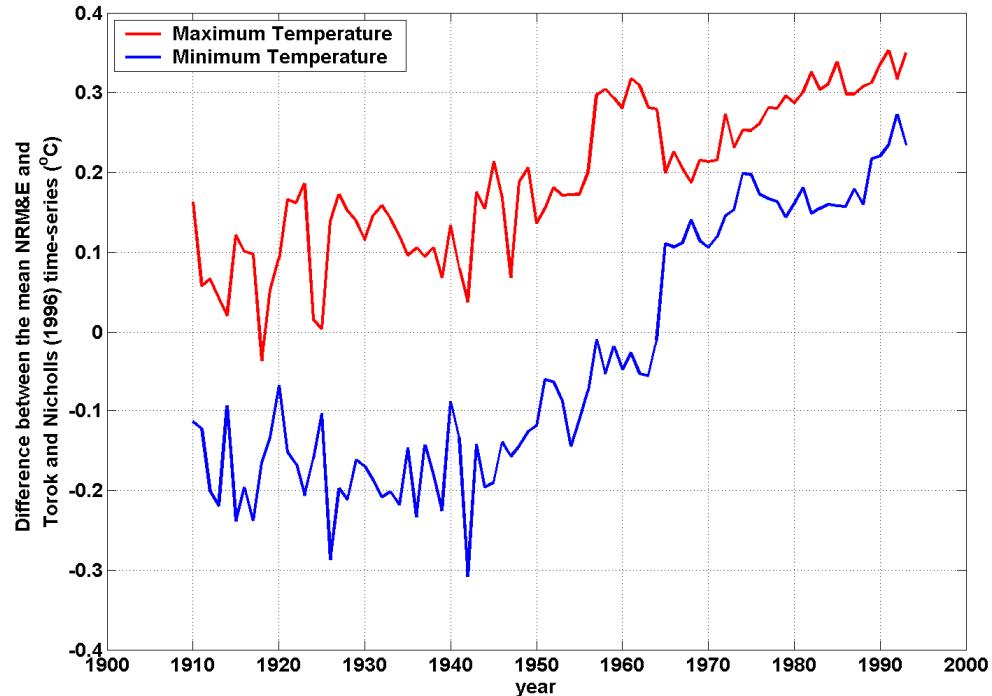


Figure 10.1: The differences between the mean annual series from Torok and Nicholls (1996), and the mean time-series derived from the NRM&E surfaces, as described in Section 9.3.1. Where the differences are positive, the Torok and Nicholls time-series is warmer than the NRM&E time-series. There are systematic trends in the differences, but there are no anomalous discontinuities at 1957.

It is important to note that there do not appear to be any discontinuities in the differences at 1957. This is interpreted to mean that the systematic differences in the time-series occur because the daily station observations used to construct the surfaces have not been corrected for systematic errors, and not because of any deficiency in the anomaly-interpolation algorithm.

10.1.2 The effect of systematic errors in station temperature records.

Because the corrections applied to the Torok and Nicholls time-series are entirely station-specific, it might be expected that they would cancel out when averaged across all of the stations in the CLIMARC network. If this were the case, then systematic errors in station observations could not be the cause of the differences between the NRM&E time-series and the Torok and Nicholls time-series.

The mean cumulative station corrections, for all 224 of the stations corrected by Torok and Nicholls, are plotted in Figure 10.2. The average corrections clearly do not cancel out. Earlier records require increasingly negative corrections, especially for minimum temperature. This is probably due to a systematic relocation of stations away from urbanisation. To illustrate the effects of such relocations, consider a station that is currently at a historically cool location (e.g. a regional airport), but in an earlier period was located at a regional post-office where it was affected by an urban “heat island”. Now, the Torok and Nicholls corrections are defined to be zero for the present station location; this is why the corrections are zero for 1993 in Figure 10.2. So, to correct the observations from the post-office location to make them comparable with the current airport location, a *negative* correction would be required.

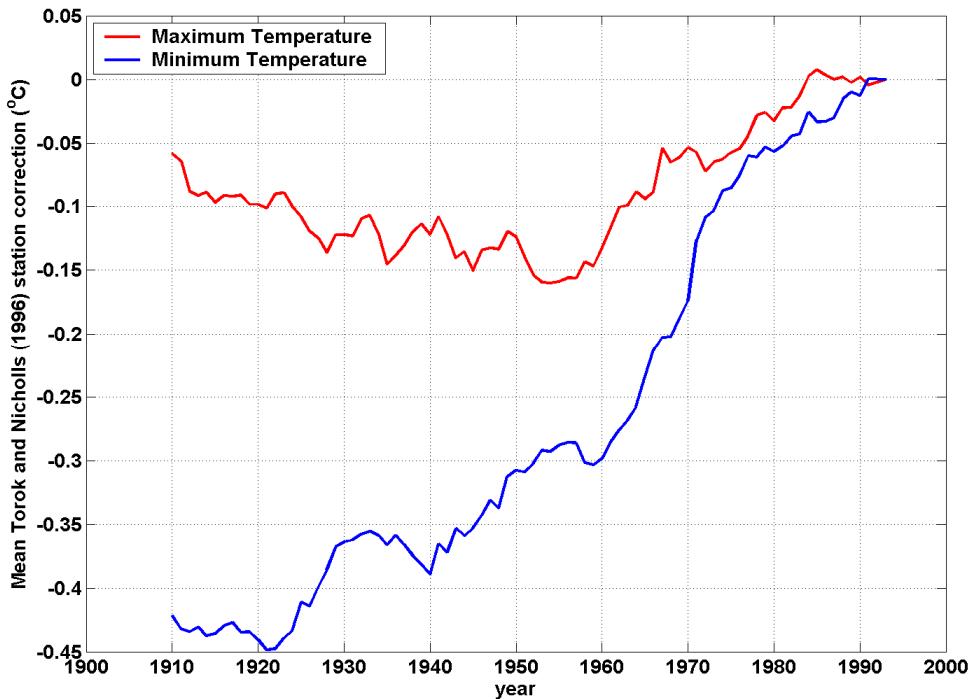


Figure 10.2: Mean station corrections for all 224 stations in the Torok and Nicholls (1996) high-quality annual-temperature dataset. Negative corrections imply the uncorrected historical records are too warm.

The magnitude of the trend in the mean station corrections for minimum temperature shown in Figure 10.2 closely matches the trend in the differences in the mean minimum temperature series shown in Figure 10.1. This is interpreted to mean that the overall effects of systematic station relocations or equipment changes are similar for both the CLIMARC network records and the Torok and Nicholls station records.

10.1.3 Bias in the pre-1949 solar radiation surfaces.

As discussed in Section 6.3.2, cloud amounts are observed at 9am and 3pm local time, and are recorded as oktas, where an okta is the number of eighths of the sky obscured by cloud. However, prior to 1949, observations were actually made in *tenths* of the sky obscured by cloud. The solar-radiation and cloud amount time-series shown in Figure 9.14 suggest a systematic difference between the pre-1949 and post-1949 solar radiation surfaces. A comparison of solar radiation with other climate variables suggested that this systematic difference is at least partly associated with the change in observing procedure.

Over much of inland Australia, there is a natural correlation between solar radiation and diurnal temperature range¹⁷, because the atmospheric temperature increase during the day is related to the available solar energy. In addition, solar radiation is inversely-correlated with rainfall. This is because there is a very good inverse-correlation between solar radiation and cloudiness, and for obvious reasons cloudiness is correlated with rainfall. So, diurnal temperature range and rainfall can be used as independent estimates of solar radiation. These independent estimates will also contain any genuine climate trends or discontinuities in solar radiation, but will not be affected by the change in cloud amount observing procedure.

Synthetic solar radiation, rainfall and diurnal-temperature range time-series were created by sampling the NRM&E climate and rainfall surfaces at the locations of the CLIMARC-network stations¹⁸. The correlation between the solar radiation and diurnal temperature range is only strong for inland Australian locations. Only the 15 locations for which the correlation in the differences was $r^2 > 0.55$ were used in this analysis. The daily time-series were summed to form annual time-series, and mean-subtracted annual time-series were calculated. Finally, an Australian time-series was calculated by taking the mean of all the synthetic time-series. Prior to 1910, very few stations were recording cloud amounts, so data prior to 1910 were omitted from the analysis.

The relationships between solar radiation and diurnal temperature range, and between solar radiation and rainfall, are shown in Figure 10.3. Pre-1949 years are shown as blue, years 1949 and later are shown as red. There is a clear separation between the pre-1949 and post-1949 observations in the relationship between solar radiation and rainfall. The separation is not as well-defined for the relationship between solar radiation and diurnal-temperature range.

Regression residual time-series derived from regression of mean annual solar radiation onto diurnal temperature range and rainfall are shown in Figure 10.4. The regression residuals show the variations in annual solar radiation which are not related to changes in diurnal temperature range, or rainfall. Figure 10.4 suggests that the break-point in the overall relationship between solar radiation and rainfall is 1950/1951. This is very close to the year when the cloud observation procedure changed from tenths to oktas. Figure 10.4 also shows a trend for solar radiation to be decreasing relative to rainfall in the post-1949 period.

Taken together, Figure 10.3 and Figure 10.4 suggest that the relationships between daily solar radiation and diurnal temperature range, and between daily solar-radiation and rainfall, are different for the periods 1910 to 1949, and 1950 to 2003. They are evidence that the solar radiation data in CLIMARC include both long-term trends in cloud cover (e.g. Jones and Henderson-Sellers, 1992; Henderson-Sellers, 1992), as well as the effects of a change in the observation procedure. Users of CLIMARC data should be aware of these factors, especially when solar radiation is used to calculate other variables, such as synthetic Class A Pan evaporation, or plant growth.

¹⁷ Diurnal temperature range is daily maximum temperature minus daily minimum temperature.

¹⁸ CLIMARC-network stations which did not actually record 9am and 3pm cloud amounts in the pre-1957 period were discarded, however.

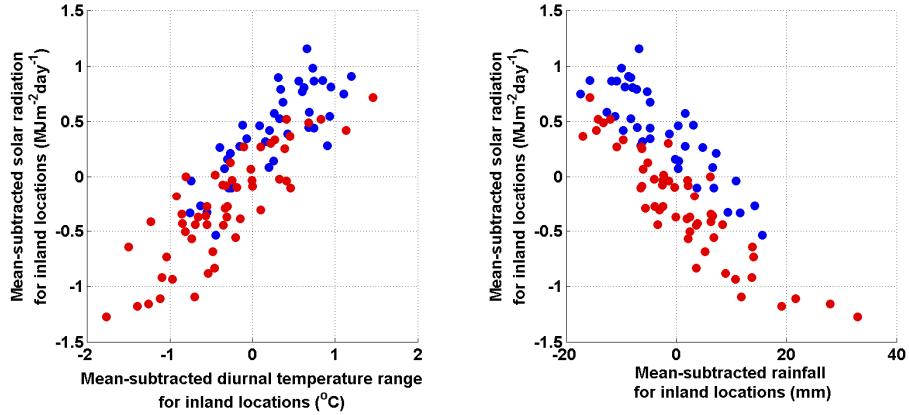


Figure 10.3: Scatter plots showing the correlation between mean-subtracted annual solar radiation and mean-subtracted annual diurnal temperature range (*left*), and between mean-subtracted annual solar radiation and mean-subtracted annual rainfall (*right*). The blue dots represent pre-1949 years, the red dots represent years 1949 and later. For the relationship between solar radiation and rainfall, there is a clear separation between the pre-1949 and post-1949 observations. The separation is not as well-defined for the relationship between solar radiation and diurnal temperature range.

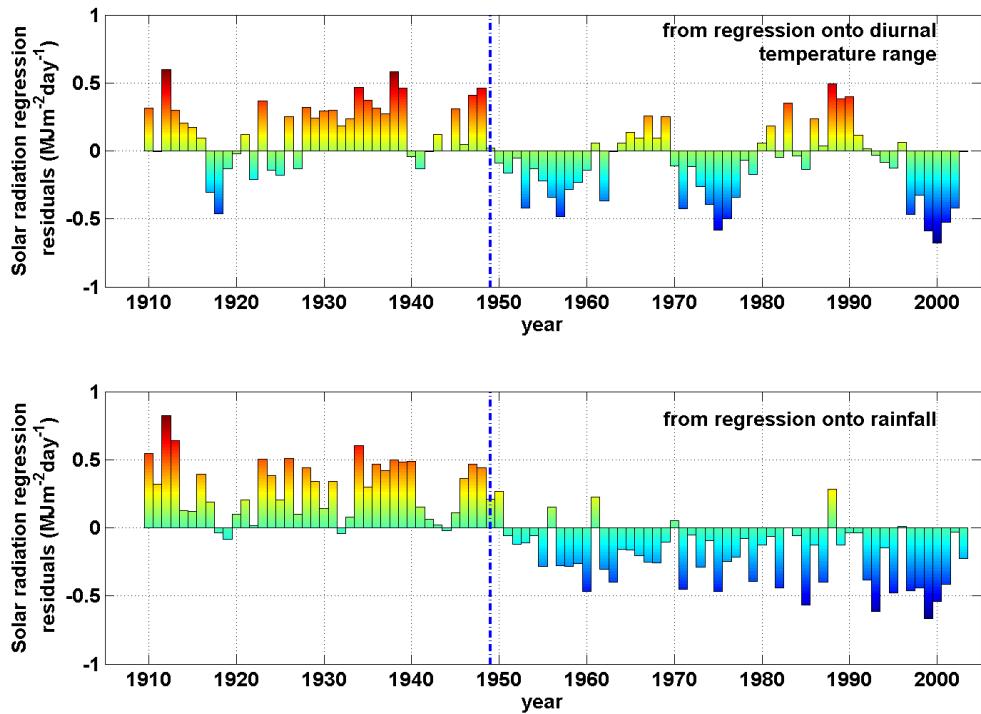


Figure 10.4: Regression residual time-series derived from regression of mean annual solar radiation onto diurnal temperature range (*top*), and of mean annual solar radiation onto rainfall (*bottom*). The regression residuals show the variations in annual solar radiation which are not related to changes in diurnal temperature range/rainfall. The residuals from the regression onto rainfall show that a systematic change in the relationship with solar radiation occurred in 1950-1951.

10.1.4 The discontinuity in the NRM&E solar radiation surfaces due to the conversion of cloud observations from tenths to oktas.

Cloud amounts are currently recorded as oktas, where an okta is the number of eighths of the sky obscured by cloud. Prior to 1949, observations were made in units of tenths of the sky obscured by cloud (Jones and Henderson-Sellers, 1992). In the CLIMARC dataset supplied by the Bureau of Meteorology, however, the pre-1949 cloud amounts have been converted from tenths to eighths by multiplying by 0.8, and then rounding.

The NRM&E solar radiation surfaces are calculated from a lookup-table of 9am/3pm cloud okta observations (Jeffrey *et al.*, 2001). The effect on the NRM&E solar radiation surfaces of the conversion of cloud amounts from tenths to oktas was investigated using simulated cloud amounts. The details of the simulation are given in Appendix F.

For most locations, the conversion of tenths to oktas would cause a systematic difference between the pre-1949 and post-1949 solar radiation surfaces. This “conversion discontinuity” contributes to, but does not fully account for, the apparent discontinuity in the NRM&E solar radiation surfaces which occurred around 1949 (as discussed in Section 10.1.3).

Note that the conversion discontinuity is a mathematical artifact of the way in which the NRM&E solar radiation surfaces are derived from cloud amount observations, and exists independently of whether any changes occurred in the real or recorded average cloud amounts at this time. The conversion discontinuity could be removed by deriving the daily solar radiation from the originally-observed cloud tenths.

10.2 The effect of using zero-anomalies in the interpolation.

The anomaly-interpolation spline method involved interpolating observed deviations from long-term average surfaces, rather than interpolating climate data directly. An important aspect of the method was to insert zero-anomalies into data-sparse regions prior to interpolation. The use of zero-anomalies prevented unrealistic values being generated for data-sparse regions, even when the interpolation was run with very few input stations. Thus, the surfaces represent the actual daily climate where there were nearby climate observing stations, and the long-term average where there were no nearby stations. As a result, interpolated data were used for the entire historical time-series, and there was no need to impose an artificial “cut-off” date for interpolated data. This gradual cut-off in the climate signal is an advantage for many modelling applications, but means the surfaces are not suitable for climate change studies.

It is important to realise that, as a consequence of using zero-anomalies, the *observation signal* in the NRM&E climate surfaces – the extent to which the surfaces represent actual daily climate observations rather than long-term averages – is not spatially or temporally uniform. Although this has always been the case with the NRM&E climate surfaces, in that the observation signal was zero prior to 1957, the use of anomaly-interpolation means the surfaces *gradually* revert to long-term averages, as the number of observations decreases in the early period of the climate record.

The gradual decrease in the observation signal in the NRM&E climate surfaces has two consequences:

- there will be spurious temporal trends or jumps in the early climate surfaces, and
- the day-to-day variations in the climate time-series will be lower in the early climate surfaces.

10.2.1 Trends in the early climate surfaces.

Because only observations from 1957 and later were used to calculate the long-term average surfaces, there will be spurious temporal trends in the early climate surfaces. For example, over much of Australia there has been an increase in minimum temperature since 1957 (Figure 9.11).

As a consequence, in data-sparse regions the early anomaly-interpolated minimum temperature surfaces will be systematically warmer than was really the case.

The decrease in observation signal can be qualitatively demonstrated by plotting derived time-series. For example, plots of annual-average maximum temperature, derived from the anomaly-interpolated surfaces for five locations in Western Australia, are given in Figure 10.5. Prior to 1907, the Carnarvon and Meekatharra time-series show no variation, so they basically contain only long-term average data. There is a systematic jump in the Carnarvon time-series at 1907, which is clearly spurious. The Halls Creek, Broome and Port Hedland series are influenced by observations back to the beginning of the time-series, although there is less inter-annual variation in the early period. Further, the initial cooling trend for these series (from about 1889 to 1915) is presumably the reversion to long-term averages due to the use of zero-anomalies in the interpolation algorithm.

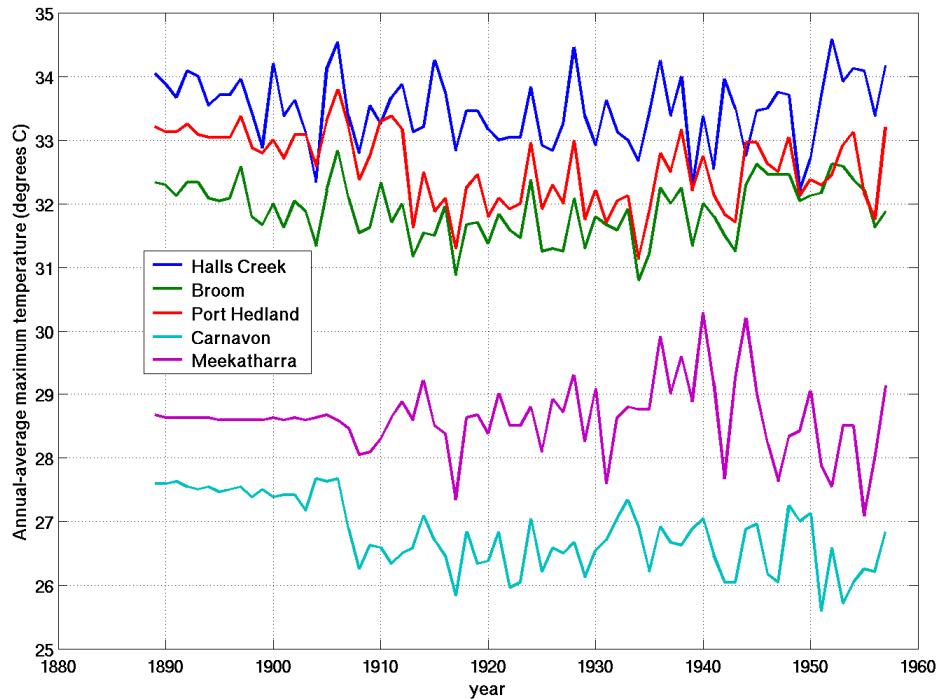


Figure 10.5: Annual-average maximum temperature, derived by sampling the anomaly-interpolated surfaces at five locations in Western Australia. Prior to 1907, the Carnarvon and Meekatharra time-series are basically long-term average data. The Halls Creek, Broome and Port Hedland series are influenced by observations, but the initial cooling trend is presumably a gradual reversion to long-term averages.

10.2.2 Statistical analysis of the day-to-day variability in the NRM&E climate surfaces.

The second consequence of the decrease in the observation signal is that the day-to-day variations in the climate time-series will decrease. This is because the long-term averages represent smoothly-varying, seasonal climate conditions.

The plots in Appendix E give an indication of where and when the surfaces revert to the long-term averages, as indicated by the day-to-day climate variation in the surfaces. The plot for each year shows the ratio of the root-mean-square (*rms*) daily variation, for that year, to the median *rms* daily variation for 1957 to 2003. Both the seasonal and inter-annual climate variations were subtracted prior to the analysis.

The *dark regions* are where there is little daily climate variation, and hence the surfaces represent long-term averages. The *light regions* are where the daily climate variation is similar to the 1957 to 2003 surfaces, and so represent daily climate observations. *Regions which are white* have at least 80% of the median daily variation in the 1957 to 2003 surfaces.

Finally, note that the images only indicate the amount of day-to-day climate variability in the surfaces. They should not be taken as an indication of data-quality.

10.2.3 *The effect of zero-anomalies on the differences between the NRM&E surfaces and the Torok and Nicholls time-series.*

The differences between the Torok and Nicholls (1996) annual maximum and minimum temperature time-series and the NRM&E surfaces were discussed in Section 10.1. The use of zero-anomalies is unlikely be a significant factor in these differences. Firstly, there are relatively few stations in the Torok and Nicholls dataset in the areas which the plots in Appendix E indicate have little daily climate signal. Secondly, there is no indication in Figure 9.10 or Figure 9.11 that the average time-series revert back to the post-1957 averages.

Note that the completeness of the Torok and Nicholls time-series only goes back to 1910. The use of zero-anomalies will have a significant effect on time-series derived from the NRM&E temperature surfaces in the period before 1910.

10.3 Other limitations.

There will be trends in the NRM&E temperature surfaces caused by urbanisation. Although Torok and Nicholls excluded stations with populations greater than 10,000 for the high quality annual dataset (Torok, 1996), the daily climate dataset is much sparser, and so all available daily station records were used to generate the NRM&E temperature surfaces. As a further consequence of the sparseness of the CLIMARC network, the anomaly-interpolated surfaces will presumably show urbanisation trends over areas much larger than are actually affected.

11 Further work.

Research and development to improve the accuracy and relevance of the NRM&E climate surfaces is ongoing.

For historical climate applications, the most significant problems with NRM&E climate surfaces are that:

- there are no daily interpolated Class A pan evaporation surfaces prior to 1970,
- the station records used to generate the climate surfaces were not corrected for systematic errors,
- there are regions and time periods where anomaly-interpolated daily climate surfaces are inadequate, and
- NRM&E has not applied consistent error-detection tests to the data.

11.1 Extending the Class A pan evaporation surface collection.

Pan evaporation records prior to 1970 are not used to generate interpolated surfaces. This is because problems with the various measuring devices that were in use before 1970 lead to inconsistent and unreliable records (Hounam, 1961). The use of substituted long-term averages prior to 1970 places significant limitations on the accuracy any historical water-balance models which use the pan evaporation data.

A possible solution is to develop *synthetic Class A pan evaporation* surfaces, where Class A pan is modelled using other climate variables. There is no universally accepted method for estimating Class A pan evaporation from other climate variables, and pan evaporation measurements are notorious for incorporating site-specific and operational biases. However, a synthetic pan evaporation model has been investigated by Richards *et al.* (2001), as part of the *Aussie GRASS* project (Carter *et al.*, 2000). For selected Queensland rangeland locations, this model accounted for 70-80% of the variability of measured pan evaporation. NRM&E is currently seeking funding to develop gridded synthetic pan evaporation using a similar methodology.

11.2 Correcting systematic errors in climate records.

Correcting systematic errors in the daily climate records, caused by changes in instrumentation, location, observation practices or environment, would give more confidence in identifying trends in historical model outputs.

The annual corrections derived by Torok and Nicholls (1996) for long-record Australian temperature observing stations would provide a good starting point. There are only seven long-record, temperature-recording stations in the NRM&E climate archive (CLIMARC and non-CLIMARC) that are not included in the Torok and Nicholls analysis¹⁹, so it is possible to apply direct corrections to most of the pre-1957 component of the temperature archive. Deriving corrections for the post-1957 component of the archive would require different techniques, because many of the stations do not have the long records required by statistical correction techniques. A hybrid technique could be investigated, where the NRM&E daily climate surfaces are adjusted so that the annual means match surfaces derived from the Torok and Nicholls time-series.

The NRM&E vapour pressure and daily solar radiation records should also be corrected for inconsistencies, although there has been relatively little research published on such corrections. There are presumably biases in the early CLIMARC wet/dry bulb thermometer readings due to

¹⁹ The stations not included are Nabawa, Morawa, Tennant Creek, Ceduna, Kyancutta, Maree and Tewantin.

non-standardised instrumentation. Vapour pressure has less spatial variation than maximum or minimum temperature, however, so discontinuities introduced by location changes will presumably not be as severe.

The NRM&E daily solar radiation records are derived from cloud okta observations. As mentioned in Section 9.3.3 and discussed further in Section 10.1.3, there appears to be a systematic discontinuity in the historical cloud records around 1949, caused by a change in observation practice. Also, as discussed in Appendix F, the conversion of cloud amounts from tenths to eighths introduces a mathematical artifact into the NRM&E solar radiation surfaces, independent of whether any changes occurred in the real or recorded average cloud amounts. This mathematical artifact could be removed by deriving the daily solar radiation from the originally-observed cloud tenths.

Apart from the discontinuity around 1949, cloud okta observations are by their nature an average over a local area, so they should not be overly sensitive to local changes in station location. Also, changes in instrumentation will not be an issue for okta records, although there may be observer biases.

11.3 Extending the record for data-sparse periods.

The anomaly-interpolation algorithm involves interpolating deviations from long-term average surfaces. This provides a simple method for incorporating the spatial climate patterns derived from the post-1957 station network into the surfaces derived from the much sparser CLIMARC network. Even so, this analysis has identified regions and time periods where anomaly-interpolated daily climate data will be inadequate for some applications (Appendix E). In particular, the climate for the period 1890 to 1910, which included early degradation/recovery episodes associated with the development of the pastoral industry (McKeon, Hall and Henry *et al.*, 2004), is not accurately represented in the anomaly-interpolated surfaces. A more complex approach to estimating historical daily climate is required.

A number of sources of historical climate information are available.

- The actual climate records are the most important record of historical daily climate, and must be the most important input when creating pre-1957 daily climate surfaces.
- The post-1957 daily climate records and derived surfaces contain information on climate patterns with a higher spatial-resolution than can be derived from the pre-1957 observations alone (see Figure 4.2).
- The monthly TABS datasets have been used to determine climate variability and trends (e.g. Henderson-Sellers, 1992; Torok and Nicholls, 1996).
- The daily rainfall record may provide an alternative estimate of daily climate in areas where there are no daily climate records (see Figure 9.15).
- The inter-correlations within the climate variables may be useful where not all climate variables are recorded at a station (e.g. only daily maximum and minimum temperature).
- There are other records within the CLIMARC dataset (e.g. phenomena), and other unrelated datasets (e.g. interpolated sea-surface temperature records) which may be useful.

The challenge is to integrate these sources of information.

The anomaly-interpolation algorithm provides a method whereby observed daily climate records are used to perturb a default, or “pattern”, climate surface. Thus, the algorithm does not have the scope to elegantly incorporate multiple sources of information, and must eventually be superseded by a more rigorous mathematical approach. However, it should not be difficult to modify the long-term average surfaces used in the maximum and minimum temperature anomaly-interpolations to incorporate the long-term climate variability and trends which were identified in the Torok and Nicholls (1996) annual time-series. This would remove one of the main limitations of the NRM&E climate surfaces.

11.4 Error detection.

All climate observation records contain errors. These errors are generally small, however, and the observations remain a valuable approximation to the real state of the world. However, there are some observations in the climate datasets (both CLIMARC and non-CLIMARC) for which the errors are so large that the observations are meaningless.

As discussed in Section 6.3, preliminary error-checking has been performed on the CLIMARC network dataset. Despite requiring a large effort in terms of manual inspection, these checks were limited to removing the most obvious random and systematic errors.

From the perspective of an NRM&E climate surface user, an automatic error-detection system would have several benefits:

- The errors would be identified in a uniform manner for all data.
- The process of error identification could be properly documented.
- The process would be repeatable if the climate data were re-processed.
- It may be possible to improve the reliability of error identification.

Clearly, there are also significant benefits for the NRM&E surface providers, in that errors could be identified much more rapidly than by manual checking, especially when applied to new data.

As mentioned in Section 6.3, simple automatic error detection schemes were investigated, such as single-variable threshold tests. Suspect observations were compared graphically with the observations from nearby stations and with observations from the days before and afterwards. In general, it was found that the threshold tests detected a lot of observations which, when inspected, could not be confidently identified as errors. But setting the threshold higher, to reduce the number of spurious error detections, meant that genuine errors were missed. The same comments apply to the TAMET error detection criteria (Wall, 1977), which is why TAMET outputs are always inspected manually.

Some automatic threshold tests are applied to observed station data, to remove data that are physically impossible. For example, observed minimum temperature cannot exceed the maximum temperature on the same day. However, setting up a more sensitive automatic error-detection system will require the development of more sophisticated multi-variable, spatial techniques.

As mentioned in Section 6.3, it will not be possible to devise statistical error detection schemes which can comprehensively identify data errors, because of the sparseness of the current CLIMARC network. Improving the station density, through further computerisation of historical climate records, would improve the reliability of error detection systems, and also improve the quality of the NRM&E interpolated surfaces.

12 Conclusions.

The NRM&E daily climate surface collection has been extended to cover the period 1889-1956, incorporating the newly computerise station records from the CLIMARC project. Interpolated daily climate surfaces are now available for rainfall, maximum and minimum temperatures, solar radiation and vapour pressure, for the period 1889 – present. Daily Class A pan evaporation surfaces are available from 1970 to present.

The maximum and minimum temperature, solar radiation and vapour pressure surfaces for the period 1889-1956 were computed using an anomaly-interpolation spline algorithm. Previously, long-term average climate surfaces were used as surrogate surfaces for this period. Substituted long-term average surfaces were inadequate for many applications.

The new anomaly-interpolated surfaces represent historical climate more accurately and realistically than the long-term average surfaces, because they:

- represent actual daily weather,
- represent inter-annual climate variability,
- do not contain unrealistic data, even when there were few station observations,
- preserve the natural inter-variable correlations in the climate datasets.

Users should be aware that there are systematic errors in the observation records that were used to compute the NRM&E climate surfaces, caused by such factors as station relocations, or changes in instrumentation or observing procedures. Also, the anomaly-interpolated surfaces gradually revert to long-term average surfaces as the number of observations decreases in the early period of the climate record. This reversion does not happen uniformly across Australia; the surfaces represent actual daily climate where there were nearby climate observing stations, and long-term averages in areas where there were not enough data to interpolate.

Because of systematic errors in the observation records, and because the anomaly-interpolated surfaces revert to long-term average surfaces in the early period of the climate record, the NRM&E climate surfaces are not suitable for climate change studies.

Research and development to improve the accuracy and relevance of the NRM&E climate surfaces is ongoing. NRM&E is currently seeking funding to overcome the most significant problems associated with the use of the NRM&E climate surfaces.

13 References.

- Bureau of Meteorology, 1999, *Bureau of Meteorology, Annual Report 1998-1999*, Australian Government Bureau of Meteorology, pp 102.
- Bohren, C.F. and Albrecht, B.A., 1998, *Atmospheric Thermodynamics*, Oxford University Press.
- Carter, J.O., Hall, W.B., Brook, K.D., McKeon, G.M., Day, K.A., Paull, C.J., 2000, "Aussie GRASS: Australian grassland and rangeland assessment by spatial simulation", in Hammer, G.L., Nicholls, N., and Mitchell, C. (eds.) *Applications of Seasonal Climate Forecasting in Agricultural and Natural Ecosystems*, 329-349, Kluwer Academic Publishers.
- Henderson-Sellers, A., 1992, "Continental Cloudiness Changes this Century", *GeoJournal*, **27.3**, 255-262.
- Hounam, C.E., 1961, *Evaporation in Australia: A critical Survey of the Network and Methods of Observation together with a Tabulation of the Results of Observations*, Bulletin **44**, Bureau of Meteorology.
- Hutchinson, M.F., 1999, *ANUSPLIN Version 4.0 Users Guide*, Centre for Resource and Environmental Studies, Australian National University, Canberra, ACT.
- Jeffrey, S.J., Carter, J.O., Moodie, K.B. and Beswick, A.R., 2001, "Using spatial interpolation to construct a comprehensive archive of Australian climate data", *Environmental Modelling and Software*, **16(4)**, 309-330.
- Jones, P.A. and Henderson-Sellers, A., 1992, "Historical Records of Cloudiness and Sunshine in Australia", *Journal of Climate*, **5**, 260-267.
- Jones, A.J. and Trewin, B., 2002, "On the adequacy of digitized historical Australian daily temperature data for climate monitoring", *Aust. Met. Mag.*, **51**, 237-250.
- McKeon, G.M., Hall, W.B., Henry, B.K., Stone, G.S., Watson, I.W., 2004, *Pasture degradation and recovery in Australia's rangelands: Learning from History*, Queensland Department of Natural Resources, Mines and Energy.
- Power, S., Casey, T., Folland, C., Colman, A., Mehta, V., 1999, "Inter-decadal modulation of the impact of ENSO on Australia", *Climate Dynamics*, **15**, 319-324.
- Richards, R. et al., 2001, *Australian Grassland and Rangeland Assessment by Spatial Simulation (Aussie GRASS). Southern Pastures Sub-project.*, Queensland Department of Natural Resources and Mines.
- Stone, R., Nicholls, N., and Hammer, G., 1996, "Frost in northeast Australia: trends and influences of the phases of the Southern Oscillation", *J. Clim*, **9**, 1896-1909.
- Torok, S.J., 1996, "The development of a high quality historical temperature data base for Australia", PhD Thesis, University of Melbourne.
- Torok, S.J., and Nicholls, N., 1996, "A historical annual temperature dataset for Australia", *Aust. Met. Mag.*, **45**, 251-260.
- Wahba, G., 1979, "How to smooth curves and surfaces with splines and cross-validation" *Proc. 24th Conference on the Design of Experiments*. US Army Research Office 79-2. Research Triangle Park, NC, 167-192.
- Wahba, G. 1990, *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia.
- Wall, B.H., 1977, "TAMET: A computer program for processing meteorological data.", *Tropical Agronomy Technial Memorandum* (4).

14 Appendix A: The CLIMARC stations.

Station Number	Latitude	Longitude	Station Name	State
2011	-18.2519	127.7817	OLD HALLS CREEK	WA
3002	-17.95	122.25	BROOME POST OFFICE	WA
4002	-20.3139	118.5742	PORT HEDLAND POST OFFICE	WA
4032	-20.3725	118.6317	PORT HEDLAND AIRPORT	WA
6062	-24.9	113.65	CARNARVON POST OFFICE	WA
7046	-26.5933	118.4883	MEEKATHARRA POST OFFICE	WA
8050	-28.7758	114.6	GERALDTON WATER SUPPLY	WA
8051	-28.7953	114.6975	GERALDTON AIRPORT	WA
9500	-35.0289	117.8808	ALBANY	WA
9518	-34.3728	115.1358	CAPE LEEUWIN	WA
9541	-33.85	121.8833	ESPERANCE POST OFFICE	WA
10093	-31.4994	118.2242	MERREDIN RESEARCH STATION	WA
12039	-30.75	121.4667	KALGOORLIE POST OFFICE	WA
14016	-12.4	130.8	DARWIN POST OFFICE	NT
15087	-19.6475	134.1896	TENNANT CREEK POST OFFICE	NT
15540	-23.71	133.8683	ALICE SPRINGS POST OFFICE	NT
17024	-30.066	138.2739	MARREE (FARINA)	SA
17031	-29.6494	138.0625	MARREE	SA
18011	-32.1267	133.6725	CEDUNA POST OFFICE	SA
18012	-32.1297	133.6976	CEDUNA AMO	SA
18044	-33.1332	135.5553	KYANCUTTA	SA
18070	-34.7225	135.8558	PORT LINCOLN	SA
21046	-33.7844	138.2133	SNOWTOWN	SA
26020	-37.8333	140.7833	MOUNT GAMBIER POST OFFICE	SA
26021	-37.7474	140.7739	MOUNT GAMBIER AERO	SA
29004	-17.7425	139.5475	BURKETOWN POST OFFICE	QLD
30018	-18.2922	143.5483	GEORGETOWN POST OFFICE	QLD
30045	-20.7289	143.1425	RICHMOND POST OFFICE	QLD
31010	-16.9333	145.7833	CAIRNS POST OFFICE	QLD
31011	-16.8736	145.7458	CAIRNS AERO	QLD
33001	-19.5775	147.4075	BURDEKIN SHIRE COUNCIL	QLD
33002	-19.6169	147.3758	AYR DPI RESEARCH STN	QLD
33045	-21.1706	149.1794	MACKAY AERO	QLD
33046	-21.15	149.1833	MACKAY POST OFFICE	QLD
33047	-21.1642	149.1192	TE KOWAI EXP STN	QLD
34002	-20.0781	146.2614	CHARTERS TOWERS POST OFFICE	QLD
35027	-23.5267	148.1617	EMERALD POST OFFICE	QLD
36030	-23.45	144.25	LONGREACH POST OFFICE	QLD
38003	-22.9117	139.9039	BOULIA AIRPORT	QLD
39015	-24.8667	152.3467	BUNDABERG POST OFFICE	QLD
39039	-25.6258	151.6094	GAYNDAH POST OFFICE	QLD
40264	-26.3919	153.0408	TEWANTIN POST OFFICE	QLD

41023	-27.1839	151.2639	DALBY POST OFFICE	QLD
41038	-28.5481	150.3075	GOONDIWINDI POST OFFICE	QLD
44022	-26.4025	146.2381	CHARLEVILLE POST OFFICE	QLD
46037	-29.4358	142.0083	TIBOOBURRA POST OFFICE	NSW
48013	-30.0917	145.9358	BOURKE POST OFFICE	NSW
48030	-31.5	145.8	COBAR POST OFFICE	NSW
52026	-30.0236	148.1218	WALGETT COUNCIL DEPOT	NSW
55023	-30.9841	150.254	GUNNEDAH POOL	NSW
55024	-31.0261	150.2687	GUNNEDAH SCS	NSW
56017	-29.7783	151.1114	INVERELL COMPARISON	NSW
63004	-33.4167	149.55	BATHURST GAOL	NSW
65016	-33.3892	148.0081	FORBES (CAMP STREET)	NSW
72150	-35.1583	147.4573	WAGGA WAGGA AMO	NSW
72151	-35.1333	147.3667	WAGGA WAGGA (KOORINGAL)	NSW
73127	-35.0517	147.3493	WAGGA WAGGA AGRICULTURAL INSTITUTE	NSW
74114	-35.1311	147.3091	WAGGA WAGGA RESEARCH CENTRE	NSW
74128	-35.55	144.952	DENILIQUIN POST OFFICE	NSW
76077	-34.1833	142.2	MILDURA POST OFFICE	VIC
78031	-36.3364	141.6353	NHILL	VIC
90015	-38.8556	143.5128	CAPE OTWAY LIGHTHOUSE	VIC
91049	-41.5	147.2	LAUNCESTON (PUMPING STATION)	TAS
91057	-41.0567	146.7883	LOW HEAD (COMPARISON)	TAS

15 Appendix B: Non-CLIMARC stations with long temperature records.

Station Number	Latitude	Longitude	Station Name	State
10073	-31.6183	117.7217	KELLERBERRIN	WA
10111	-31.6408	116.6703	NORTHAM	WA
10579	-33.6886	117.5553	KATANNING COMPARISON	WA
10614	-32.9342	117.1797	NARROGIN	WA
10648	-32.6814	116.6756	WANDERING COMPARISON	WA
12074	-31.2319	119.3281	SOUTHERN CROSS	WA
23000	-34.9333	138.5833	ADELAIDE WEST TERRACE	SA
40214	-27.4778	153.0306	BRISBANE REGIONAL OFFICE	QLD
42023	-26.6581	150.1844	MILES POST OFFICE	QLD
43030	-26.5719	148.7897	ROMA POST OFFICE	QLD
53027	-29.5	149.9	MOREE POST OFFICE	NSW
55054	-31.0867	150.8467	TAMWORTH AIRPORT	NSW
65012	-32.2388	148.6089	DUBBO (DARLING STREET)	NSW
66062	-33.8607	151.205	SYDNEY (OBSERVATORY HILL)	NSW
86071	-37.8075	144.97	MELBOURNE REGIONAL OFFICE	VIC

16 Appendix C: Long term average surfaces.

This section briefly describes how the NRM&E long-term average climate surfaces were created.

Long-term averages were computed for each day-of-year. For a given station and day-of-year, a "pool" of observations was created. The pool consisted of all available observations taken at that station, on the day-of-year itself, and on the two days immediately before and after. Thus, if there were N years of observations at the station, there would have been $\sim Nx5$ observations in the "pool". For some years there were not observations for all the 5 days, so the actual number would be less than $Nx5$. The averages and variances were then calculated simply as the averages and variances of all observations in the pool.

These station climate averages were then interpolated using a thin-plate spline, as described in Jeffrey *et al.* (2001). The stations were weighted with a function of the variance, which ensured that stations with highly variable climate, or stations with only a short observing history, did not corrupt the surface. For a given station and day-of-year, a minimum of 15 observations were required for the station average to be included in the interpolation.

Only observation records from 1957 or later were used to calculate the long-term averages. Hence, the long-term average surfaces will include the impacts of any recent climate trends or variability.

The long-term average surfaces are re-computed on an irregular basis, to take account of new data and corrections.

17 Appendix D: Representative network and comparison details.

17.1 General comments.

Estimating errors using representative network analysis implicitly assumes that there are no significant differences the statistical properties of typical daily climate fields between the post-1957 and CLIMARC epochs. It also assumes that instrumental and observer errors are similar; anecdotal evidence suggests that daily temperature observations from the 1800's are significantly less reliable than contemporary observations. Representative network analysis also implicitly assumes that all errors are due to interpolation, whereas in practice there is also an observational error term.

One of the primary objectives of the analysis was to verify that the anomaly-interpolated surfaces estimate daily climate more accurately than substituted long-term average surfaces. Because the long-term average surfaces were determined using post-1957 data, the use of long-term average surfaces to directly estimate pre-1957 climate data is likely to give significantly *larger* errors in practice than indicated by the analysis presented here. The anomaly-interpolation technique also makes use of the long-term average surfaces, but the spline fitting compensates for errors in the long-term average surfaces.

In summary, it should be remembered that errors derived from representative network analysis are only estimates of the interpolation errors for actual pre-1957 climate data.

17.2 Network selection.

Fifty-one stations were chosen to represent the CLIMARC station network. Although the actual number of stations changes considerably in the period 1889-1956, it was computationally much easier to use a fixed representative network. The representative network closely matches the actual station set used to determine solar radiation and vapour pressure surfaces. However, there are typically 70 records per day for minimum and maximum temperature from 1907 onwards, and so the representative analysis errors will most likely overestimate the actual errors in those surfaces.

The representative stations were either CLIMARC stations, or "successors" to closed CLIMARC stations. The temperature records for the representative stations were basically complete from 1970 to 1990.

To minimise computational requirements, only the 15th day of each month from 1970 to 1990 was used in the analysis. For each of the days analyzed, the observed climate data for the representative CLIMARC stations were extracted from the climate database, and anomaly-interpolated. The observed data for all other stations (the *reserved* stations) were then extracted from the database, and compared to the values of the interpolated surface at the same locations. The differences are estimates of the errors in the interpolation for those locations.

18 Appendix E: Day-to-day variations in the NRM&E climate surfaces.

These plots give an indication of where and when the anomaly-interpolated surfaces revert to the long-term averages. The plot for each year shows the ratio of the root-mean-square (*rms*) daily variation, for that year, to the median *rms* daily variation for 1957 to 2003. Both the seasonal and inter-annual climate variations were subtracted prior to the analysis.

The *dark regions* are where there is little daily climate variation, and hence the surfaces represent long-term averages. The *light regions* are where the daily climate variation is similar to the 1957 to 2003 surfaces, and so represent daily climate observations. *Regions which are white* have at least 80% of the median daily variation in the 1957 to 2003 surfaces.

Finally, note that the images indicate the amount of daily climate variability in the surfaces. They should not be taken as an indication of data quality.

18.1 Maximum Temperature.

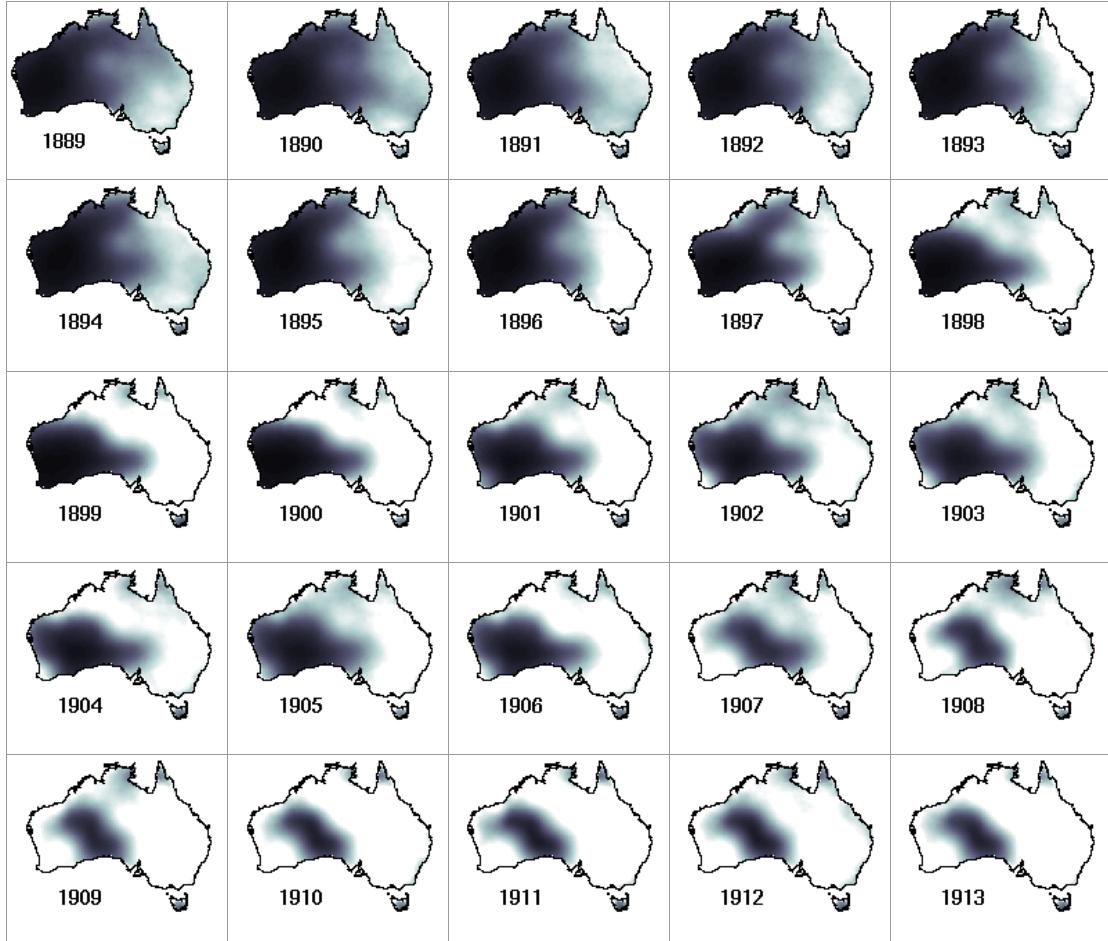


Figure 18.1: Day-to-day variations in the maximum temperature surfaces.

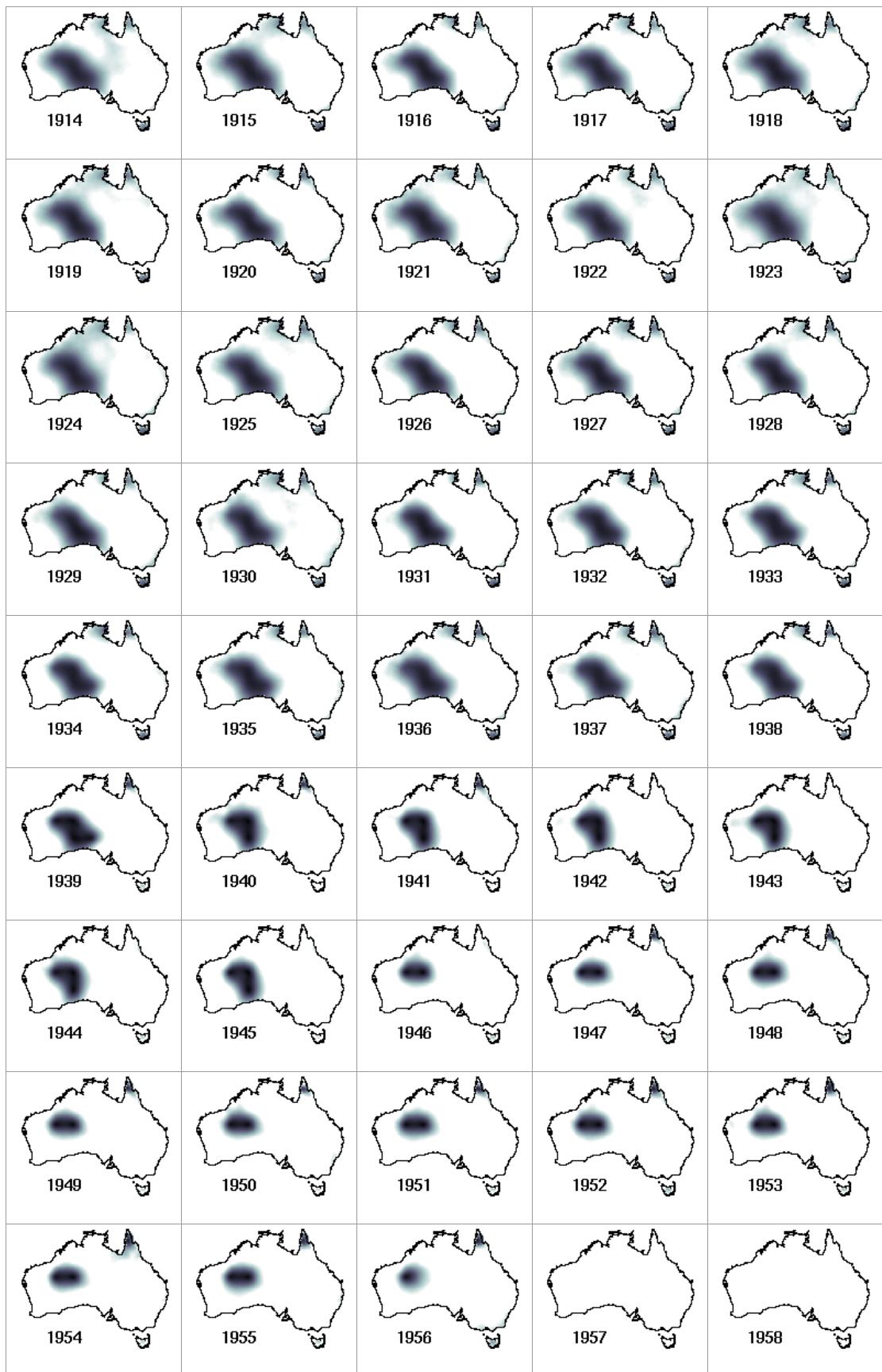


Figure 18.1: Day-to-day variations in the maximum temperature surfaces. (cont.)

18.2 Minimum Temperature.

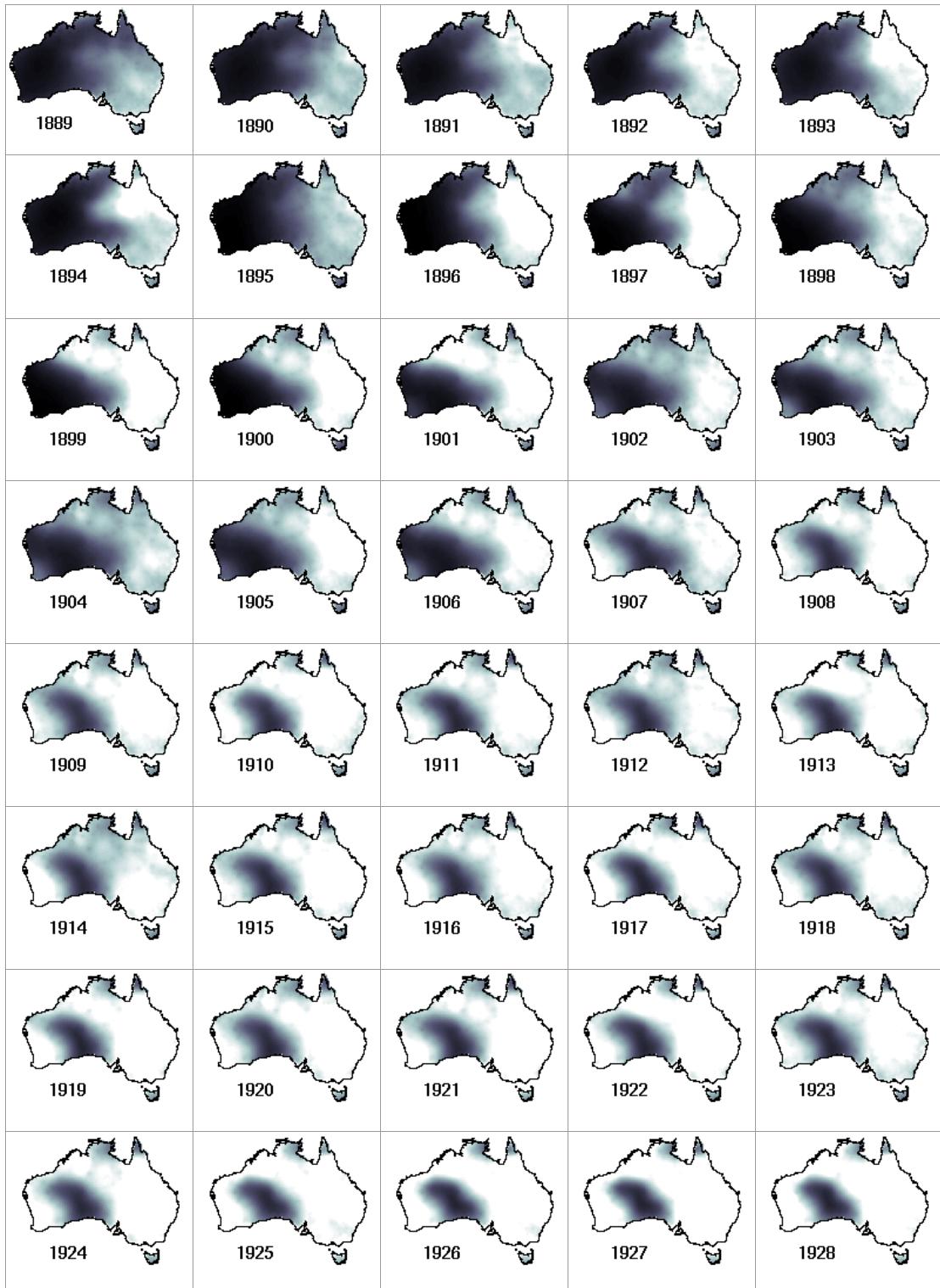


Figure 18.2: Day-to-day variations in the minimum temperature surfaces.

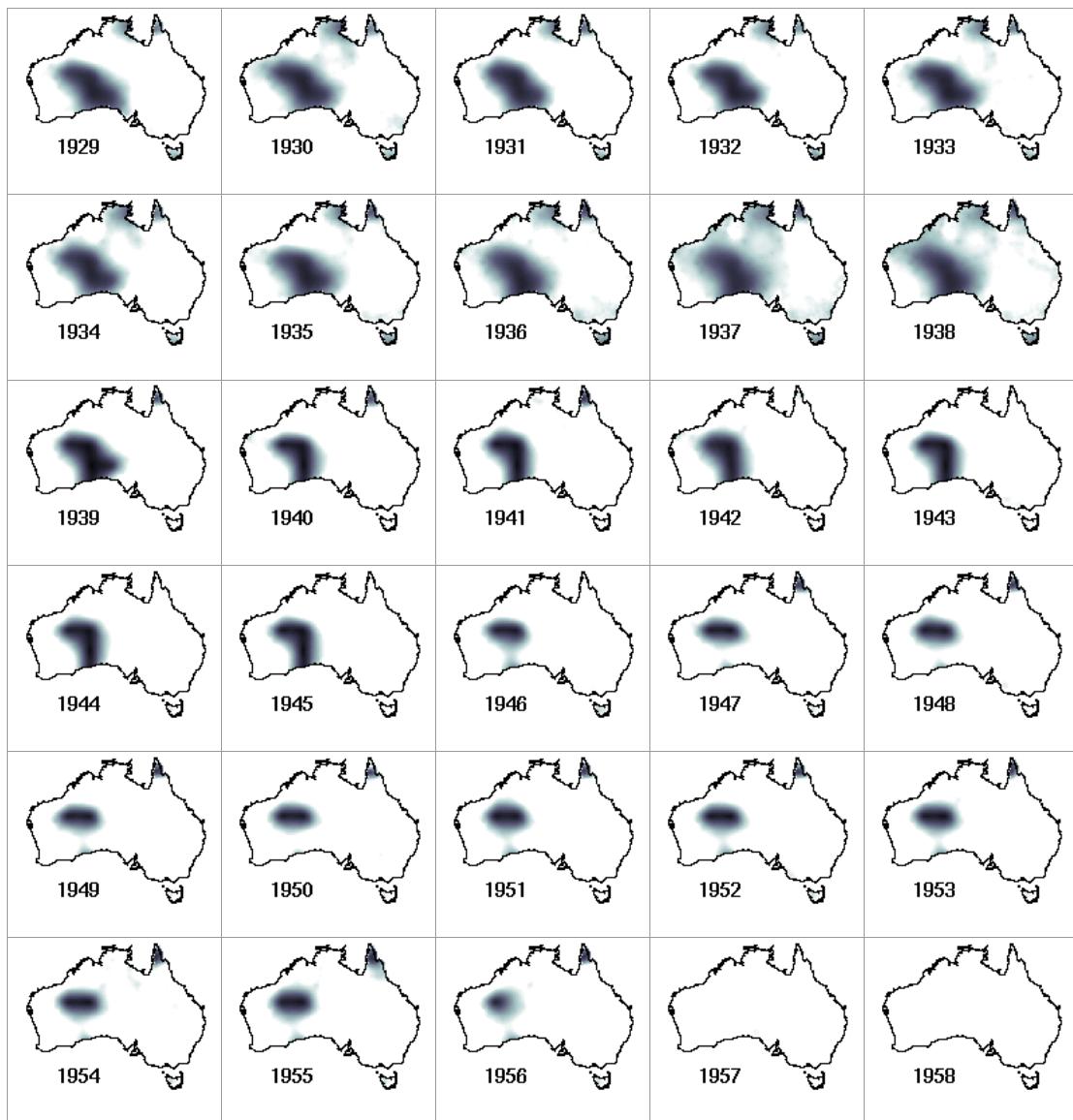


Figure 18.2: Day-to-day variations in the minimum temperature surfaces. (cont.)

18.3 Solar Radiation.

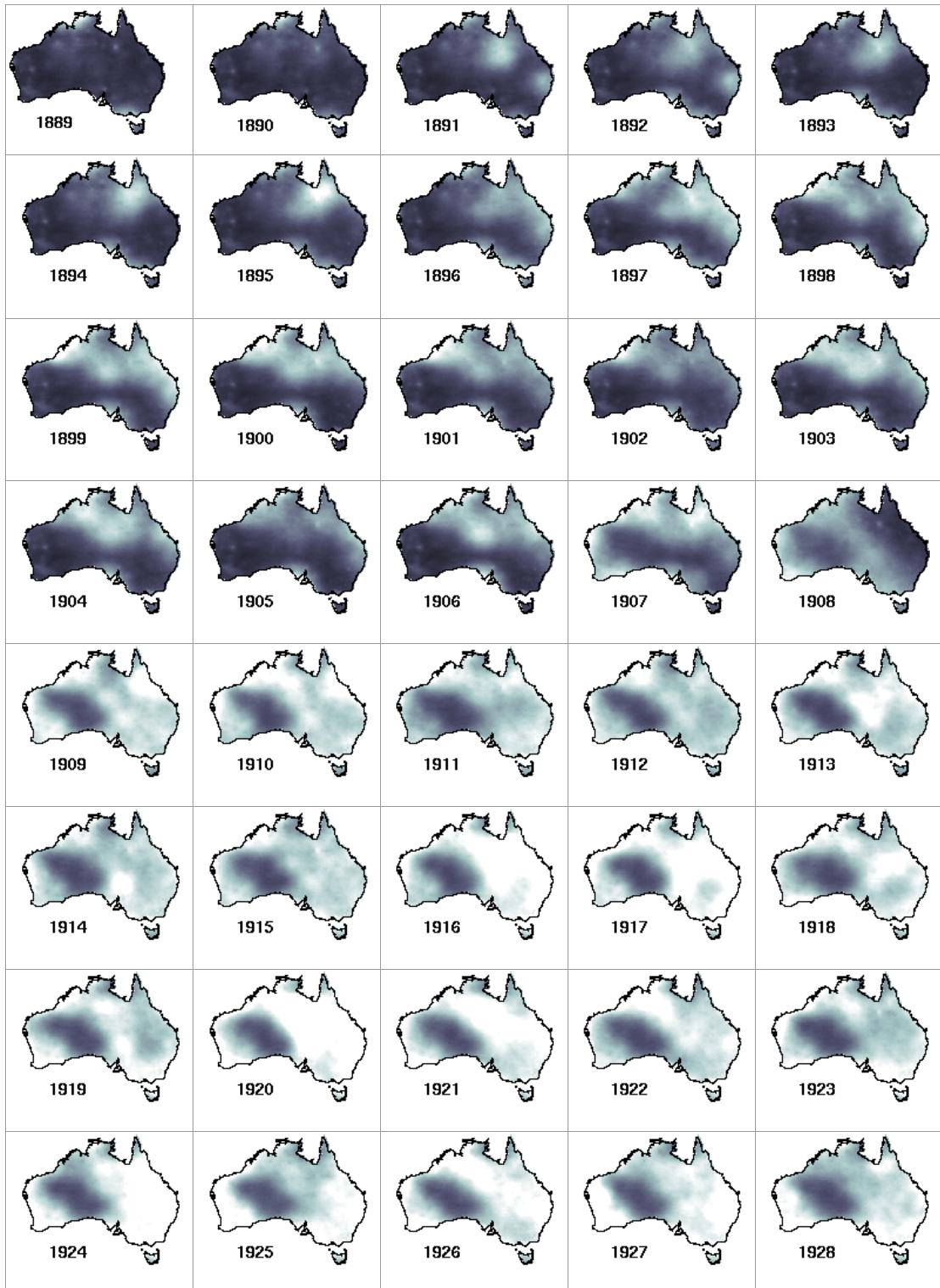


Figure 18.3: Day-to-day variations in the solar radiation surfaces.

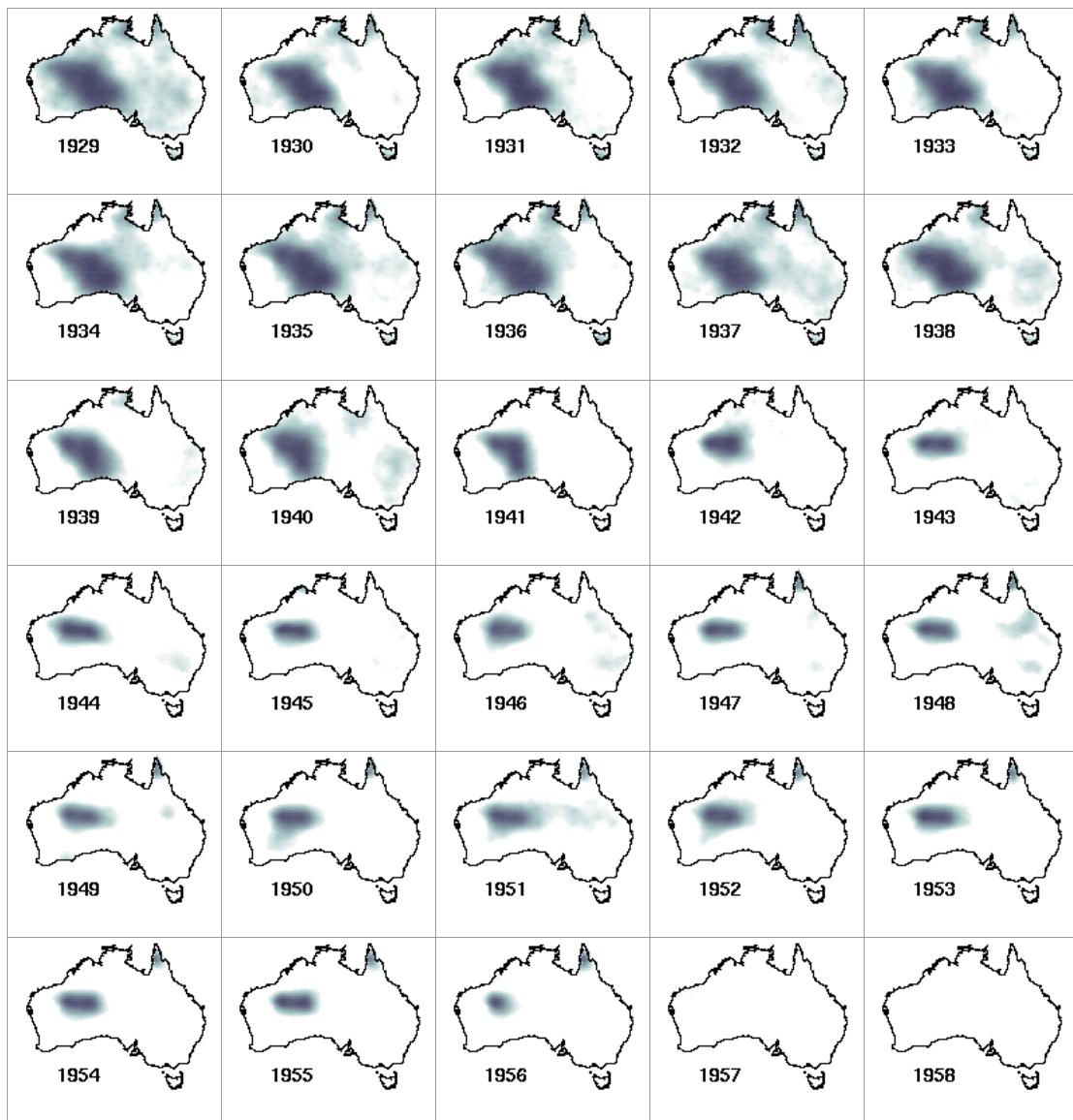


Figure 18.3: Day-to-day variations in the solar radiation surfaces. (cont.)

18.4 Vapour Pressure.

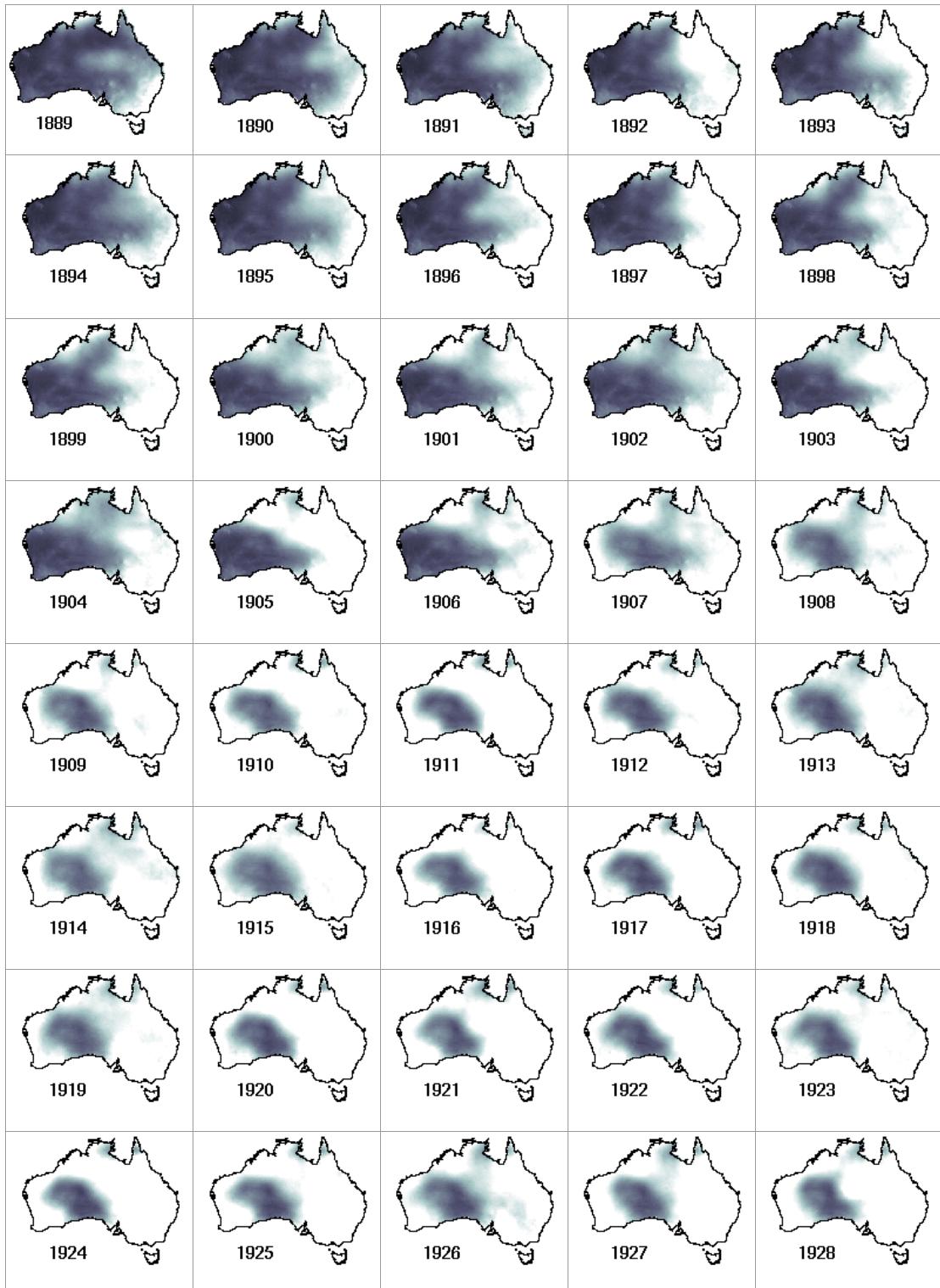


Figure 18.4: Day-to-day variations in the vapour pressure surfaces.

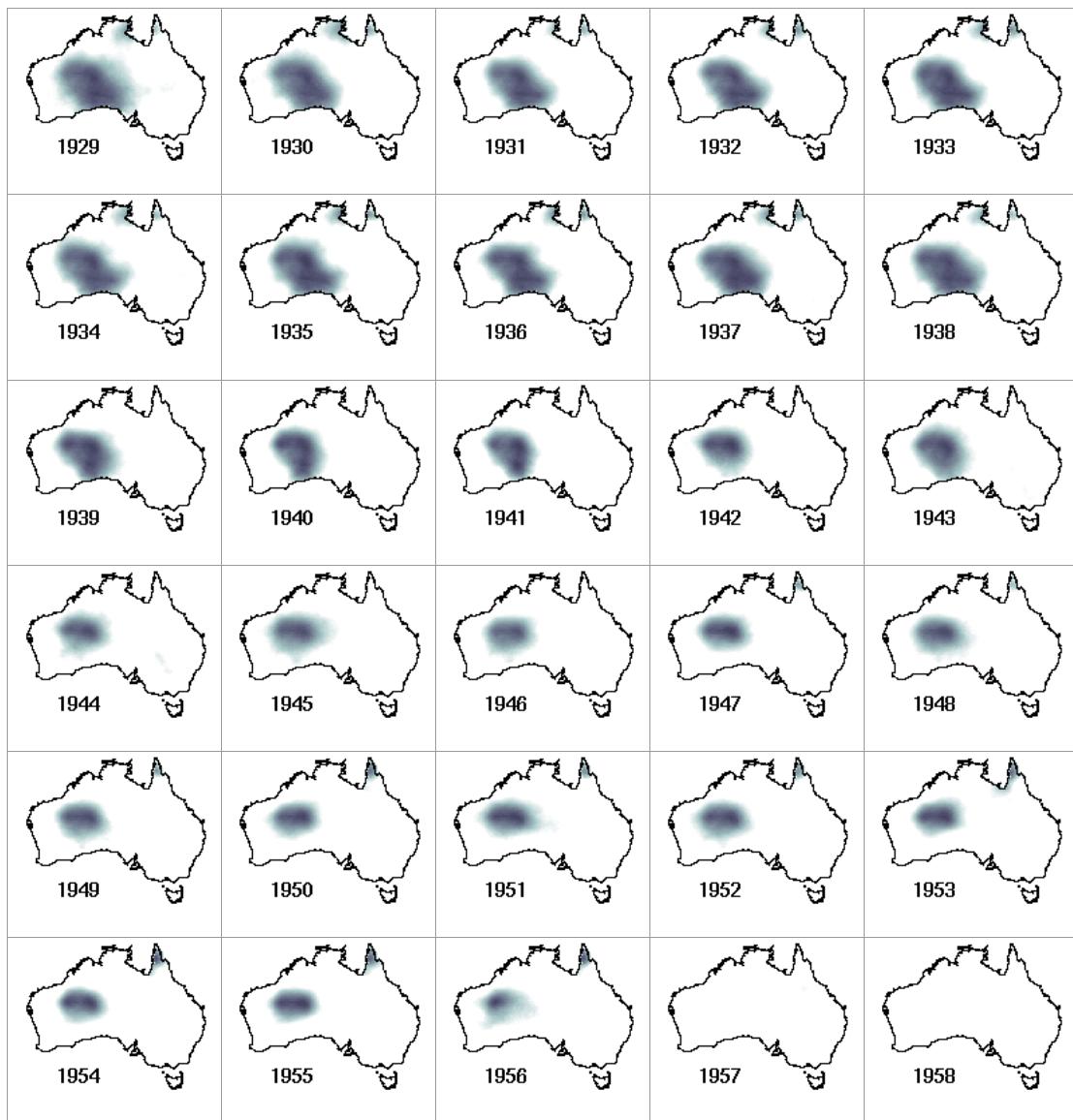


Figure 18.4: Day-to-day variations in the vapour pressure surfaces. (cont.)

19 Appendix F: The discontinuity in the NRM&E solar radiation surfaces due to the conversion of cloud observations from tenths to oktas.

19.1 Summary.

The conversion of pre-1949 cloud amounts from tenths to eighths (oktas) will cause a discontinuity in the NRM&E solar radiation surfaces. This “conversion discontinuity” contributes to, but does not fully account for, observed differences between the pre-1949 and post-1949 NRM&E solar radiation surfaces. This systematic error could be removed from future revisions of the NRM&E solar radiation surfaces, by calculating the pre-1949 surfaces using the original observations in cloud tenths.

Note that the conversion discontinuity is a mathematical artifact of the way in which the NRM&E solar radiation surfaces are derived from cloud amount observations, and exists independently of whether any changes occurred in the real or recorded average cloud amounts at this time.

19.2 Introduction.

Cloud amounts are currently recorded as oktas, where an okta is the number of eighths of the sky obscured by cloud. However, prior to 1949, observations were made in units of *tenths* of the sky obscured by cloud (Jones and Henderson-Sellers, 1992). In the CLIMARC data supplied by the Bureau of Meteorology, the pre-1949 cloud amounts have been converted from tenths to eighths by multiplying by 0.8, and then rounding. The effect of this conversion is given in Table 1 (reproduced below). Because of this conversion, cloud okta values of 2 and 6 are over-represented in the pre-1949 observations.

Table 3: Table used for converting cloud amount observations made in tenths (pre-1949) to cloud amounts in eighths (oktas). Because of this conversion, cloud okta values of 2 and 6 are over-represented in the pre-1949 observations.

Cloud amount in tenths	0	1	2	3	4	5	6	7	8	9	10
Cloud amount in eighths	0	1	2	2	3	4	5	6	6	7	8

The NRM&E solar radiation surfaces are calculated from a lookup-table of 9am/3pm cloud okta observations (Jeffrey *et al.*, 2001). The effect of the conversion of cloud amounts from tenths to oktas on the NRM&E solar radiation surfaces was investigated using simulated cloud amounts. In most locations, the conversion would cause a systematic difference between the pre-1949 and post-1949 solar radiation surfaces. This “conversion discontinuity” contributes to, but does not fully account for, the apparent discontinuity in the NRM&E solar radiation surfaces which occurred around 1949 (see Figure 9.14).

19.3 Simulation methodology.

The NRM&E solar radiation surfaces are calculated from a lookup-table of 9am/3pm cloud okta observations (Jeffrey *et al.*, 2001). The lookup-table is used to estimate daily percentage extra-terrestrial radiation, which is then converted to horizontal, ground level, daily solar radiation. So, the ideal way to determine the effect of the conversion from tenths to oktas would be to start with pre-1949 cloud amount observations in tenths, convert them to oktas using Table 1, and then use the oktas to derive solar radiation using the standard procedure. Comparing the derived solar radiation with the actual solar radiation would give the effect of the conversion. Unfortunately, this calculation cannot be performed, because there is no readily available estimate of the pre-1949 actual solar radiation.

The approach used in this analysis was to use post-1949 cloud okta observations to simulate cloud amount observations in tenths. These simulated observations were then converted back to oktas using Table 1, and then daily solar radiation calculated. Comparing the derived solar radiation from the simulated and real okta observations gave an estimate of the effect of converting from tenths to oktas.

Clearly, it was not possible to use Table 1 to convert individual cloud oktas values to cloud tenths, because there would be no way to tell, for example, whether 2 oktas should be converted to 2 tenths or 3 tenths. It was possible, however, to estimate the *distribution* of 9am/3pm cloud tenths from the distribution of 9am/3pm cloud oktas. A simulated time-series of cloud tenths was then constructed based on the estimated distribution of 9am/3pm cloud tenths, and the comparison could proceed as above.

19.3.1 Estimating distribution of 9am/3pm cloud tenths.

For each CLIMARC-network station where cloud amounts were recorded, the distributions of 9am/3pm cloud oktas were calculated using the period 1960-1979 inclusive. Separate distributions were calculated for each month. The distribution of recorded 9am/3pm cloud oktas for Longreach Post Office (36030) are shown in Figure 19.1 (*left*) as an example.

To convert the distribution functions from 9am/3pm cloud oktas to 9am/3pm cloud tenths, each okta distribution was considered as a set of (x, y, z) points, where:

- x was the 9am cloud amount, 0, 0.125, 0.25, ... 1,
- y was the 3pm cloud amount, 0, 0.125, 0.25, ... 1, and,
- z was the frequency of occurrence of (x, y) cloud oktas.

The estimated distribution of cloud tenths was obtained from these points using a conventional, non-smoothing, thin-plate spline interpolation algorithm to generate a set of points (x_{10}, y_{10}, z_{10}) , where:

- x_{10} was the 9am cloud amount, 0, 0.1, 0.2, ... 1,
- y_{10} was the 3pm cloud amount, 0, 0.1, 0.2, ... 1, and,
- z_{10} was the frequency of occurrence of (x_{10}, y_{10}) cloud tenths.

The estimated distributions of cloud tenths for Longreach Post Office are shown in Figure 19.1 (*right*). The estimated distributions of cloud tenths clearly replicate the overall properties of the observed distributions of cloud oktas.

Note the high relative occurrence of days with 9am/3pm okta records of 0/0 and 8/8 in Figure 19.1 (*left*). This was a common feature of the okta distribution functions for most stations. Because these values were often so high that they could not be sensibly interpolated, the values for 0/0 and 8/8 oktas were not used in the spline interpolation, but instead were inserted directly into the cloud tenths distribution as the as the values for 0/0 tenths and 10/10 tenths. Figure 19.1 (*right*) was calculated using this method.

For each station, the estimated distributions of cloud tenths were distributed to form simulated 20 year time-series of cloud amount observations in tenths. These time-series were converted back to cloud oktas using Table 1, and then the average monthly solar radiation was calculated using the technique of Jeffrey *et al.* (2001).

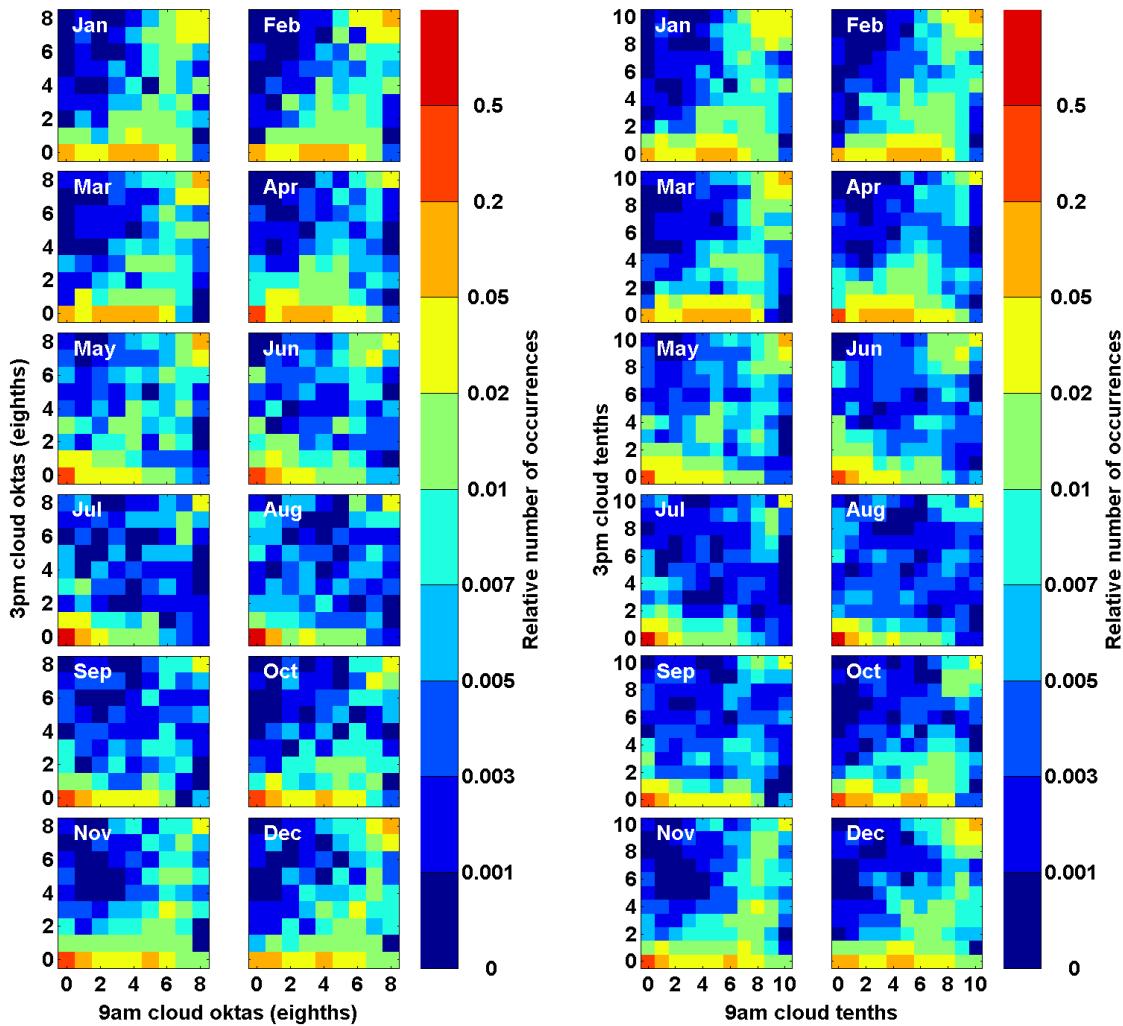


Figure 19.1: The distributions of recorded cloud amounts for Longreach Post Office (36030) in oktas (*left*), and the estimated distributions of cloud amounts for Longreach Post Office in tenths (*right*). The estimated distributions of cloud tenths clearly replicate the overall properties of the observed distributions of cloud oktas.

19.4 Results.

The differences between the solar radiation calculated from observed cloud oktas, and from simulated cloud tenths, are shown in Figure 19.2. The spatial correlation of the results between stations indicates the veracity of the simulation. For nearly all stations, the simulation indicated that solar radiation derived from cloud observations in tenths would be higher than, or equal to, the solar radiation derived from the same cloud amount observed in oktas. This result is consistent with the analyses in Sections 9.3.3 and 10.1.3, which indicated there is a decrease in the average NRM&E solar radiation surfaces around the time of the change in observing procedure from tenths to oktas.

Although the simulations indicated that the conversion of cloud records from tenths to oktas does contribute to the systematic differences between the NRM&E pre-1949 and post-1949 solar radiation, the conversion does not fully account for the differences. The differences between the average solar radiation for 1960-1979 inclusive, and 1928-1947 inclusive, calculated from the NRM&E daily solar radiation surfaces, are shown in Figure 19.3. For most stations and months, the observed differences are much larger than in Figure 19.2, indicating that these differences cannot be due to the conversion from cloud tenths to oktas alone. Secondly, in southwest Western Australia, the observed differences are actually positive in summer, which is the

opposite of Figure 19.2. If the differences in southwest Western Australia are due to climate change, then the NRM&E climate surfaces will in fact be underestimating the climate change.

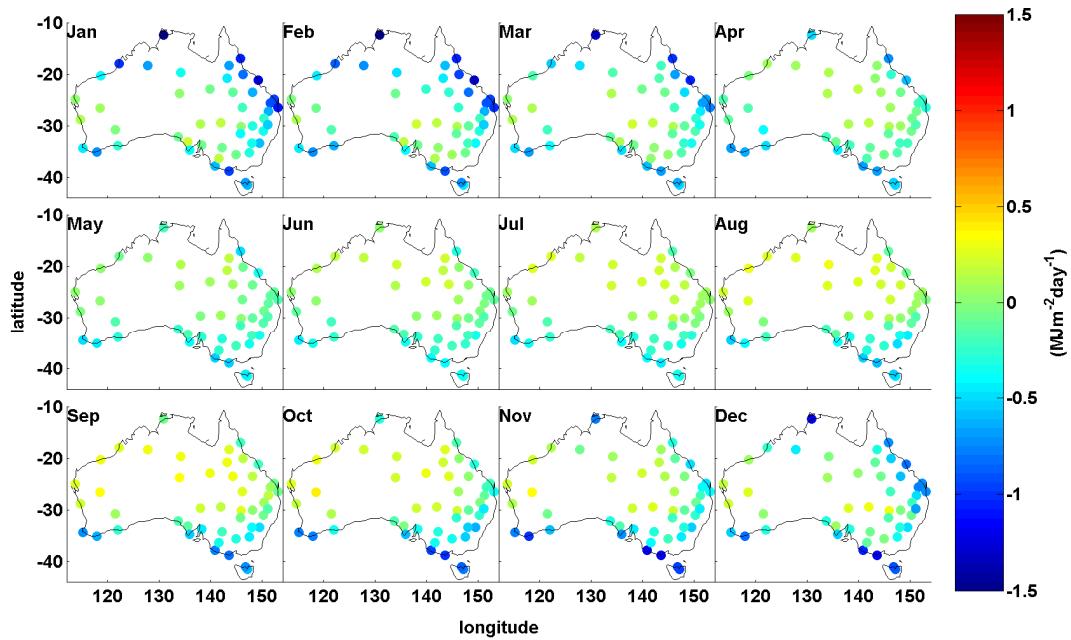


Figure 19.2: The differences between the average solar radiation calculated from observed cloud oktas and from simulated cloud tenths. For negative values, the solar radiation derived from cloud observations in tenths would be higher than the solar radiation derived from the same cloud amount observed in oktas.

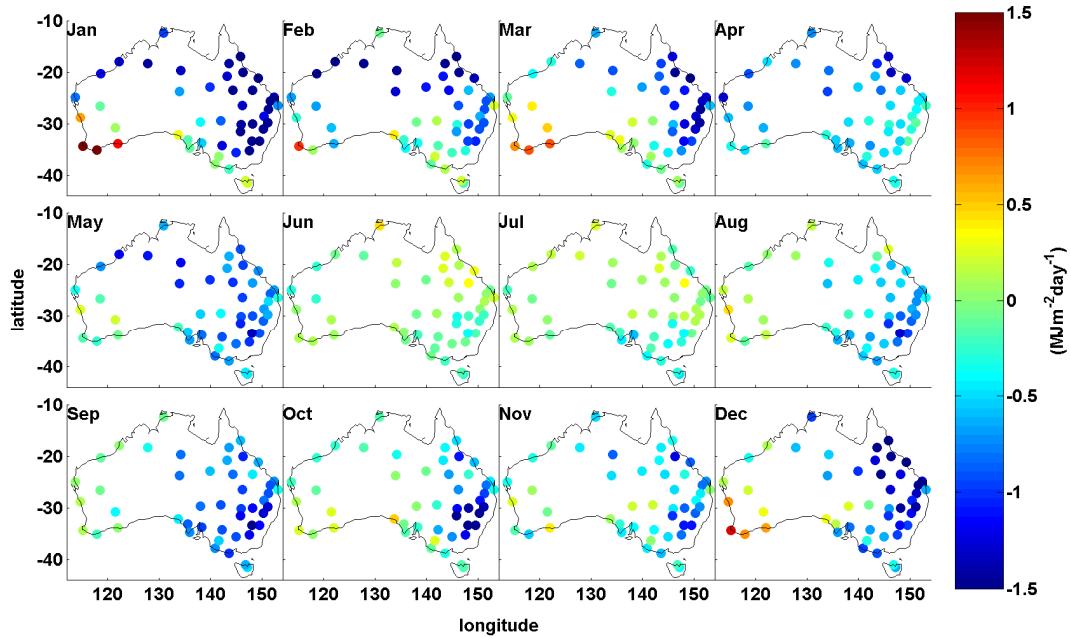


Figure 19.3 The differences between the average solar radiation for 1960-1979 inclusive, and 1928-1947 inclusive, calculated from the NRM&E daily solar radiation surfaces. For negative values, the solar radiation is higher in the period 1928-1947.

19.5 Further work.

The monthly differences shown in Figure 19.2 should not be used to correct the pre-1949 NRM&E daily solar radiation surfaces. Although deriving and applying daily corrections using a simple disaggregation algorithm would remove the conversion discontinuity in annual and mean-monthly solar-radiation surfaces, it would introduce other statistical discontinuities. For example, the pre-1949 solar radiation is calculated correctly for days for which the 9am/3pm cloud observations were 5/5 tenths, because this is exactly 4/4 oktas. Applying a simple disaggregated monthly correction to all days in the month would incorrectly bias the solar radiation for such days.

The preferred technique for “correcting” the pre-1949 solar radiation surfaces would be to derive the solar radiation from observations of 9pm/3pm cloud tenths in the first place. The original cloud observations (i.e. in tenths) have been requested from the Bureau of Meteorology. To use 9am/3pm observations of cloud tenths to estimate daily percentage extra-terrestrial radiation, the standard lookup-table for oktas described in Jeffrey *et al.* (2001) would need to be interpolated to a lookup-table for cloud tenths. The oktas lookup-table does not contain any abrupt changes between entries, so using an interpolated table would not significantly increase the error in the calculated daily solar radiation.

19.6 Conclusion.

Analysis of simulated cloud amount time-series demonstrated that the conversion of pre-1949 cloud amounts from tenths to eighths (oktas) will cause a temporal discontinuity in the NRM&E solar radiation surfaces. However, the magnitude of the simulated discontinuity is lower than the observed differences between the pre-1949 and post-1949 NRM&E solar radiation surfaces, so the conversion of cloud amounts from tenths to oktas cannot be the sole cause of the observed differences. Further, for some stations the sense of the simulated discontinuity is the opposite of the observed change, so the NRM&E climate surfaces may be underestimating long-term changes in solar radiation in these regions.