Data homogenisation requirements for extremes

Blair Trewin

Climate Monitoring and Prediction

Australian Bureau of Meteorology

Why is data homogeneity an issue?

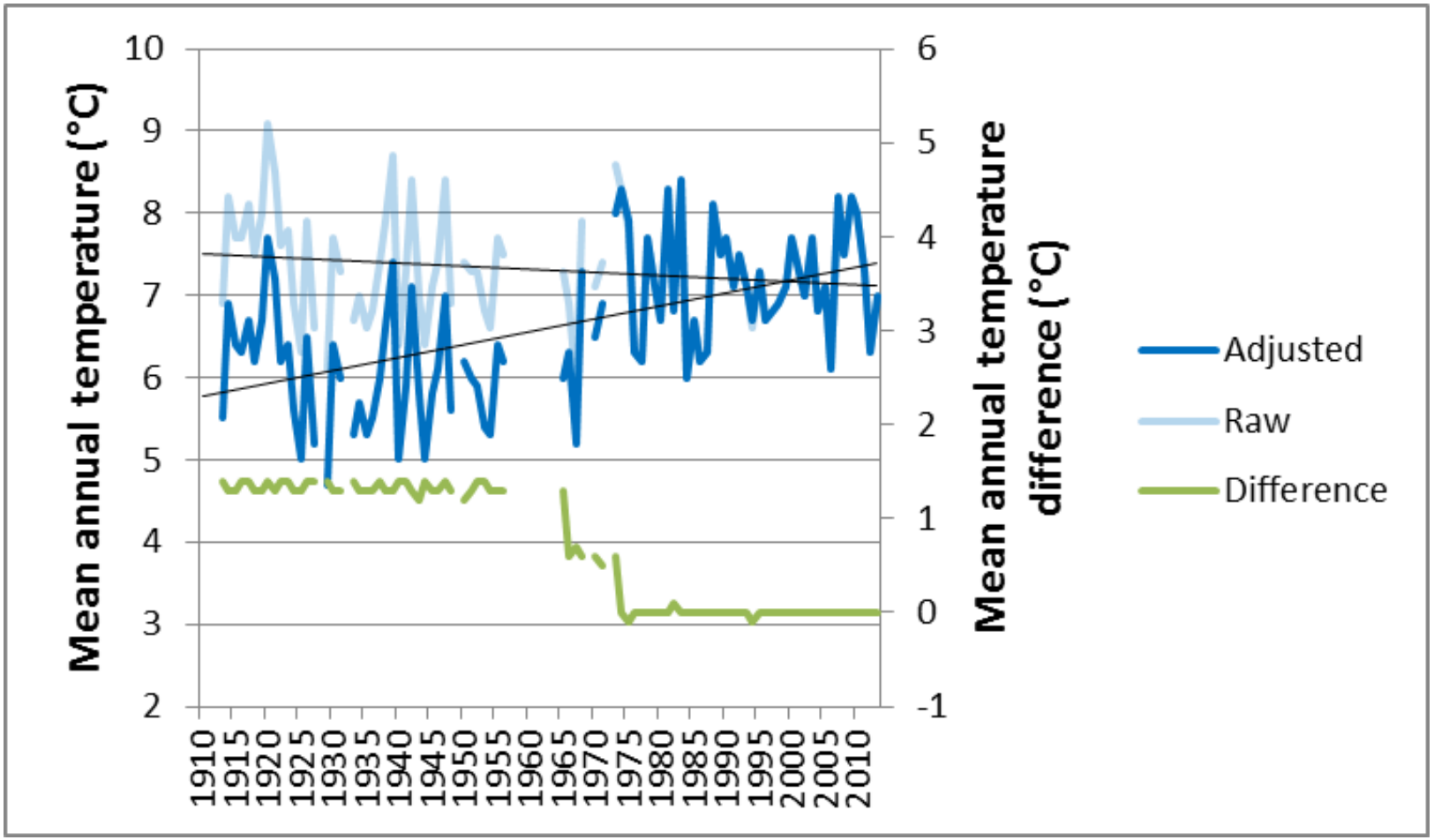
- We want to be sure that changes we are detecting reflect changes in the climate, not changes in the way the observations were made
- Non-climatic inhomogeneities can arise from numerous causes (e.g. instrument changes, site moves, observing practice changes, urbanisation/changes in site condition)
- Some issues will be more important for some variables than others

In this presentation, we will focus on daily maximum and minimum temperature.

Metadata issues

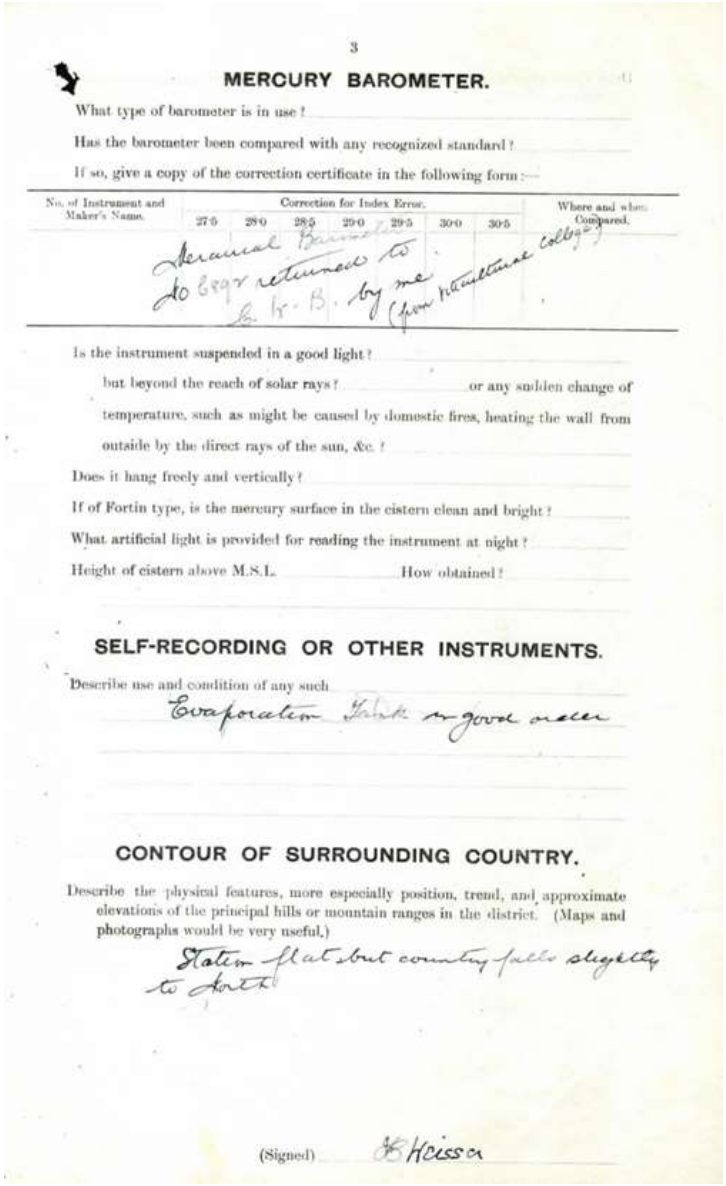
- Much metadata has not been digitised
- Some metadata (e.g. coordinates) can be readily put in a standard form, but others (e.g. local site environment) cannot
- Important metadata often does not exist
- Sometimes documentation will show that a change has happened between date X and date Y, but not exactly when

Rutherglen minimum temperatures – a contentious adjustment



The Rutherglen site in 1975

Information indicating pre-1958 location is inconsistent with 1975 site



3

MERCURY BAROMETER.

What type of barometer is in use? Aneroid Barometer

Has the barometer been compared with any recognized standard? No

If so, give a copy of the correction certificate in the following form:—

No. of Instrument and Maker's Name.	Correction for Index Error.						Where and when Compared.
	27·5	28·0	28·5	29·0	29·5	30·0	30·5

Aneroid Barometer returned to Mr. B. by me (from Agricultural College)

Is the instrument suspended in a good light?

but beyond the reach of solar rays? or any sudden change of temperature, such as might be caused by domestic fires, heating the wall from outside by the direct rays of the sun, &c.?

Does it hang freely and vertically?

If of Fortin type, is the mercury surface in the cistern clean and bright?

What artificial light is provided for reading the instrument at night?

Height of cistern above M.S.L. How obtained?

SELF-RECORDING OR OTHER INSTRUMENTS.

Describe use and condition of any such

Evaporation Tank in good order

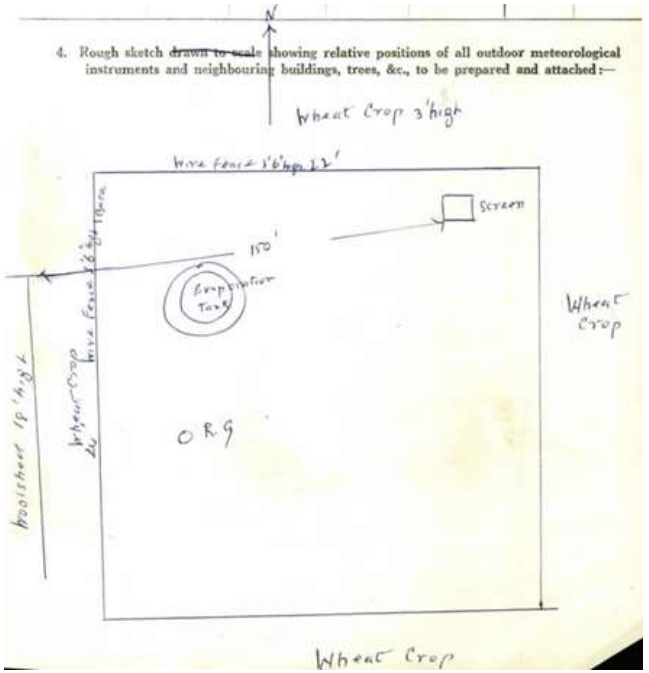
CONTOUR OF SURROUNDING COUNTRY.

Describe the physical features, more especially position, trend, and approximate elevations of the principal hills or mountain ranges in the district. (Maps and photographs would be very useful.)

Slopes flat but country falls slightly to South

(Signed) H Hicks

1939 inspection report



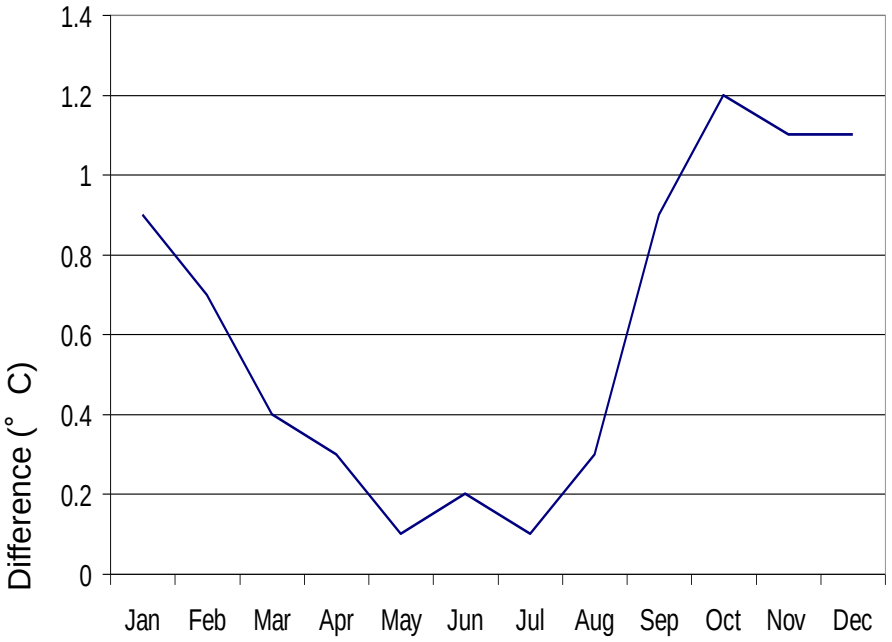
1958 site sketch

Why are extremes different?

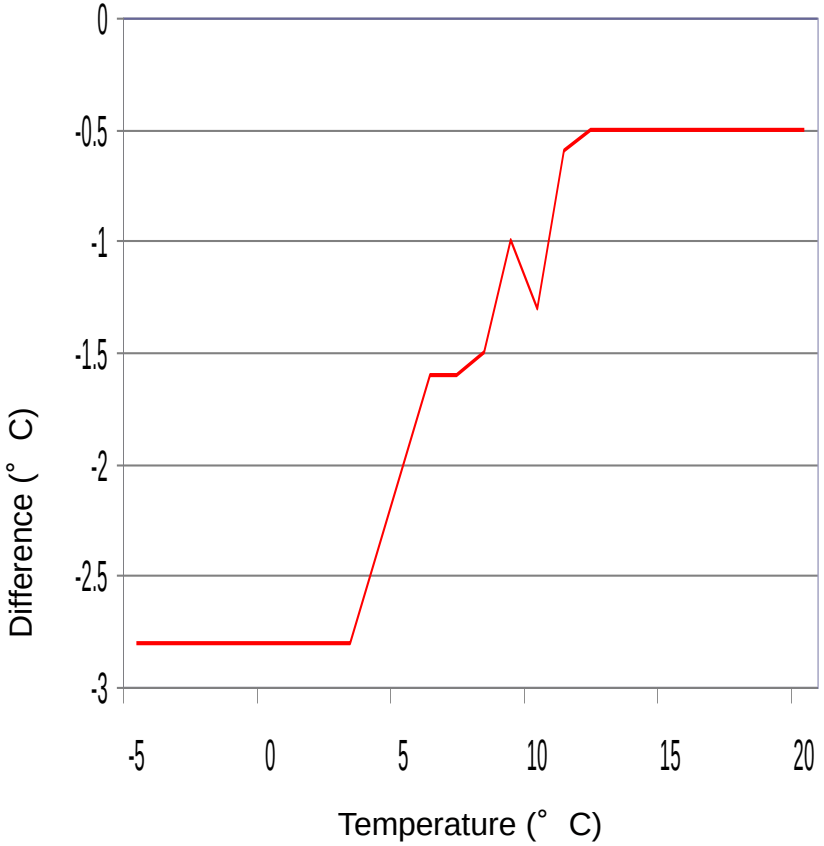
- Sometimes a mean signal manifests differently in extremes
- Local temperature differences (e.g. through urbanisation, topography, proximity to water) often maximised under clear, calm conditions

The examples in this talk are the 'hard cases' – most of the time it's not so complex!

Adjustment – why it isn't always as easy as it looks



Monthly mean maximum temperature differences, Port Macquarie

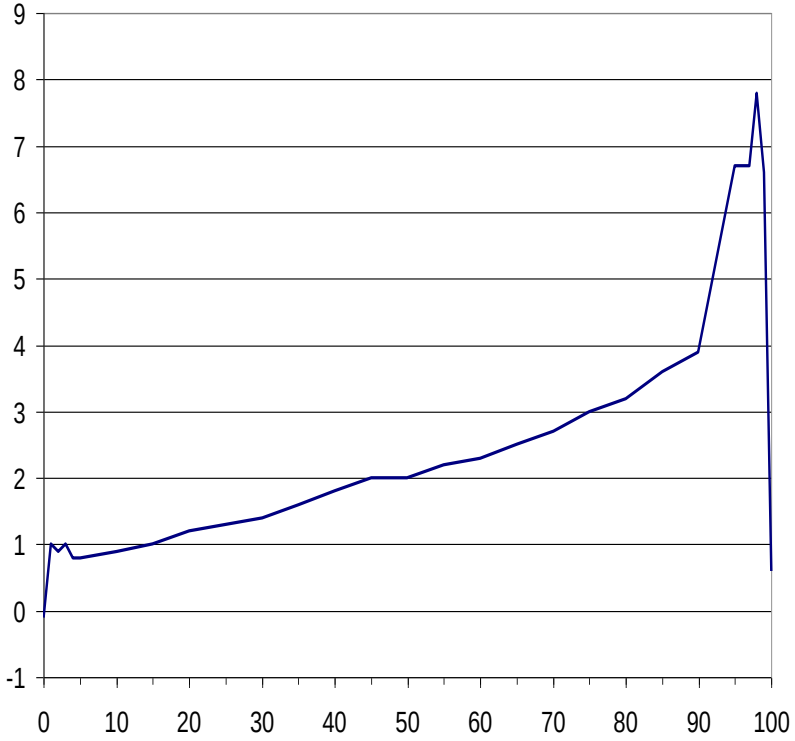


Winter minimum temperature differences , Port Macquarie

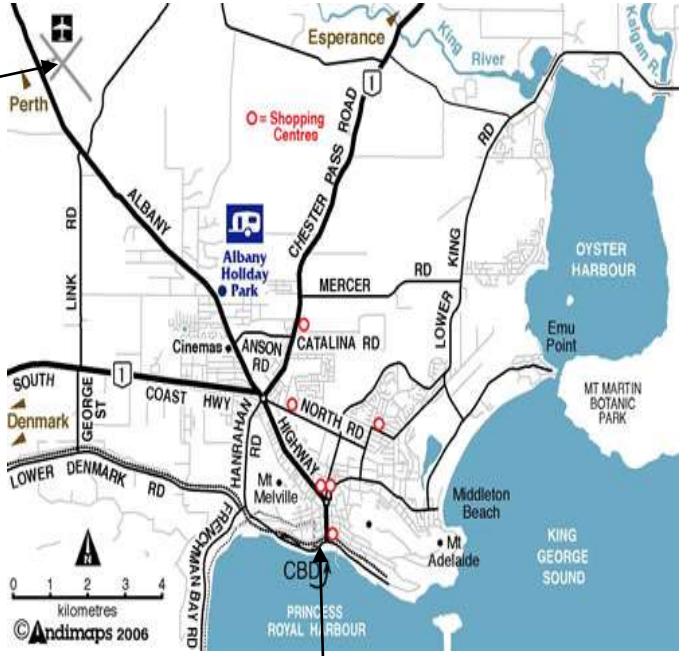
Port Macquarie is on east coast of Australia (~500km north of Sydney)
Old site was ~800m from coast, new site 4km inland

Albany – an example of an especially challenging homogenisation

Difference between percentile points, summer maximum temperatures

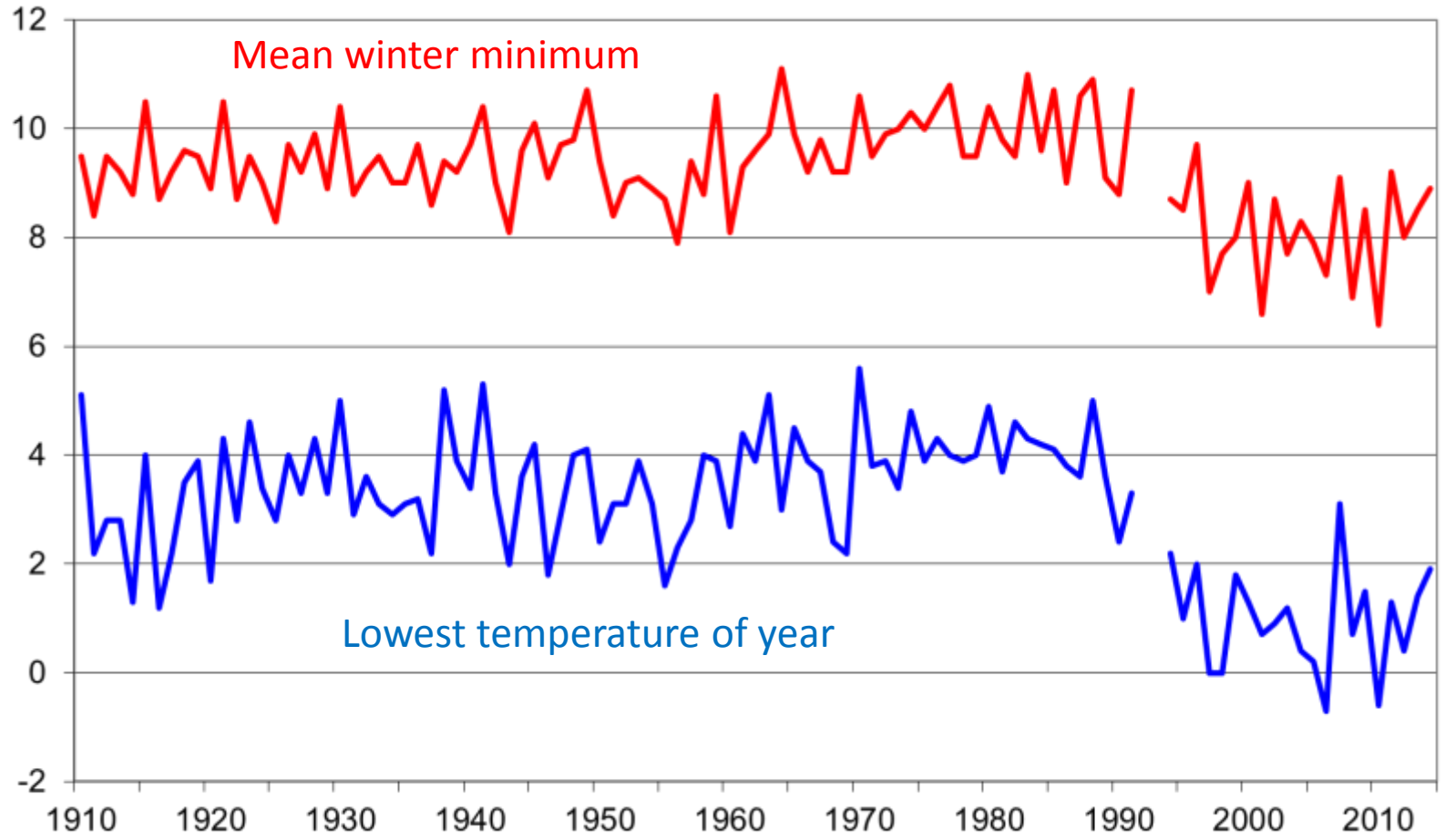


Post-1965 site: at airport



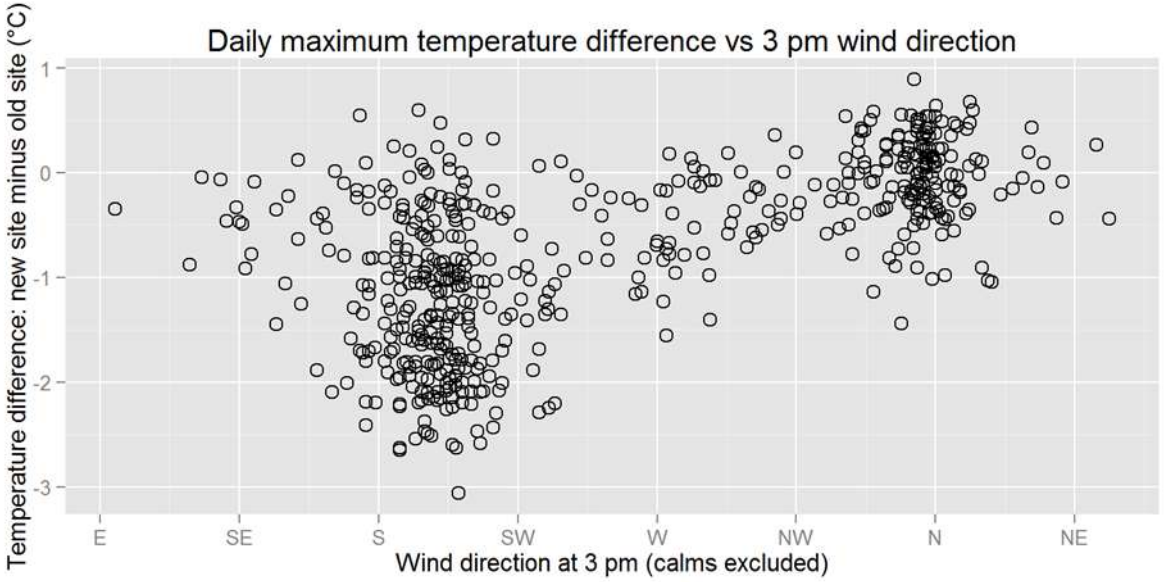
Pre-1965 site: in town

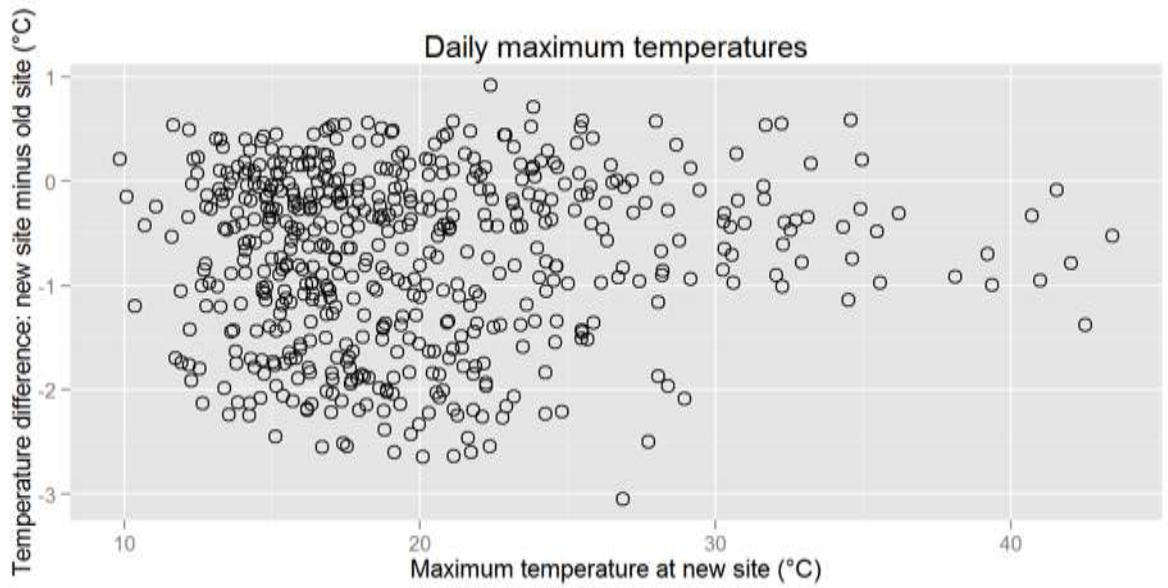
Perth – a very visible minimum temperature inhomogeneity



Site moved in 1992-93 from inner city to park north of city centre (earlier moves 1963, 1967)
Post-1993 site about 2°C cooler than 1967-92 site for means, 3°C cooler for extremes
1910-1992 record low (1.2°C) surpassed 38 times in 21 years since 1994

Melbourne: another example of a complex relationship between sites



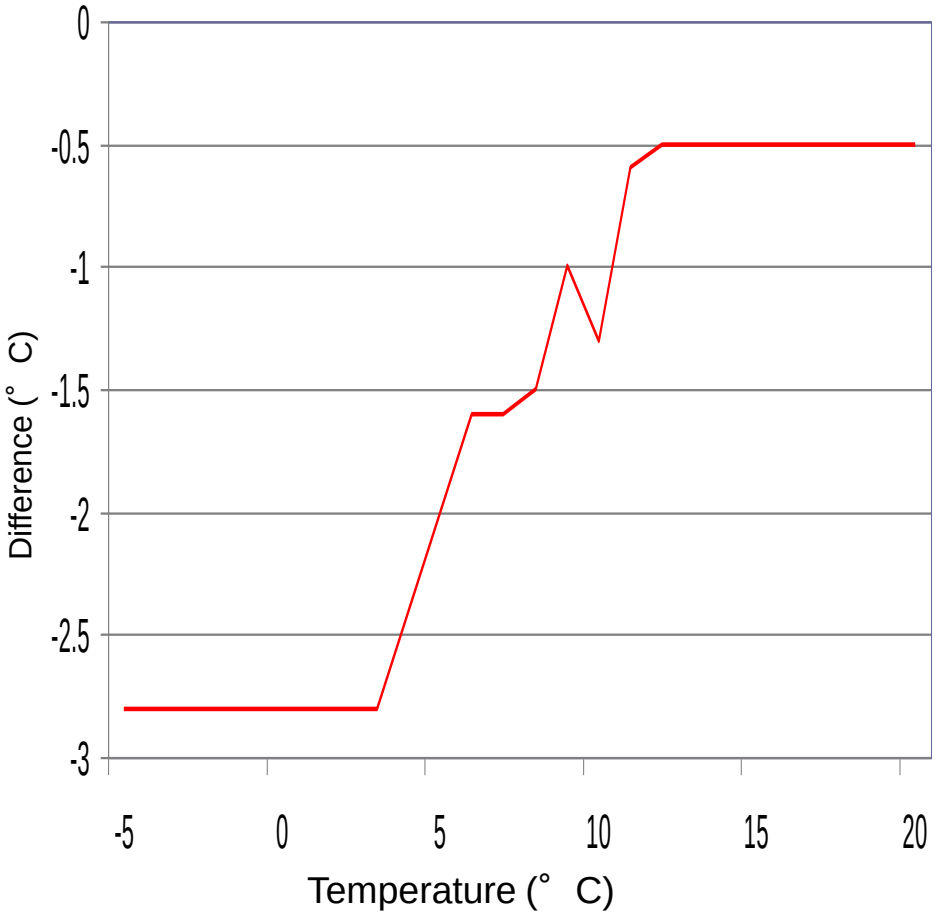


Images: plots Ian Barnes-Keoghan (BoM), map Wikipedia Commons

Clarifying what we mean by 'daily homogenisation', and methods proposed to do it

- The term 'daily homogenisation' (or similar) has been used in various places in the literature
- In some cases, this refers to adjustments which are based on monthly data, but are interpolated to a date
- For the purposes of this talk, the term will be used to refer to methods which apply differential adjustments to different parts of the frequency distribution
- Numerous methods in the literature (e.g. Della-Marta and Wanner 2006, Brandsma and Können 2006, Wang et al. 2010, Mestre et al. 2011, Trewin 2013)
- Most methods in literature still use monthly/annual data for detection.

The percentile-matching (PM) algorithm – overlap case

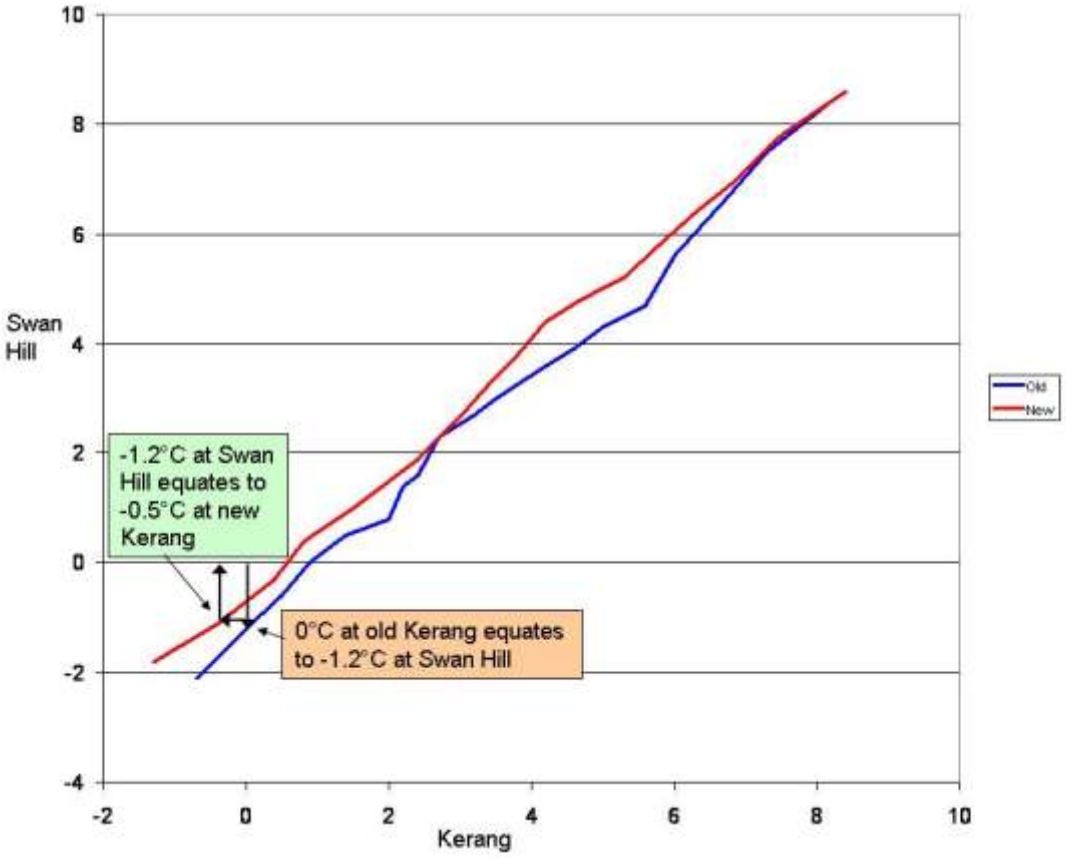


- 5th, 10th, …, 95th percentile points calculated for each site in overlap period for 3-month season
- These points used to develop transfer function
- Constant inter-site differences assumed below 5th, above 95th percentile

From Trewin (2013)

The PM algorithm – non-overlap case

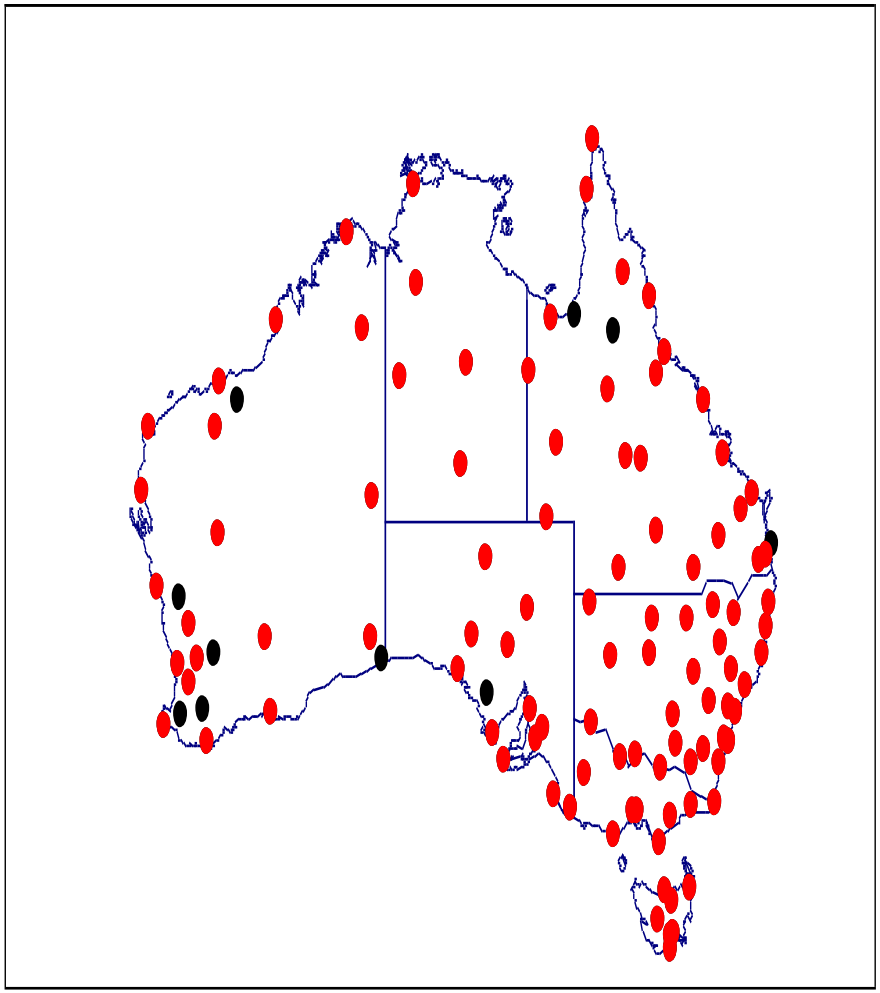
- Two-stage process using neighbouring sites:



Uses median of outcomes from 10 reference stations, if available

Evaluation shows modest improvements on monthly methods for mean-based metrics (e.g. RMS error), but much better performance for extremes-based metrics (e.g. highest/lowest values, values of 10^{th}/90^{th} percentile)

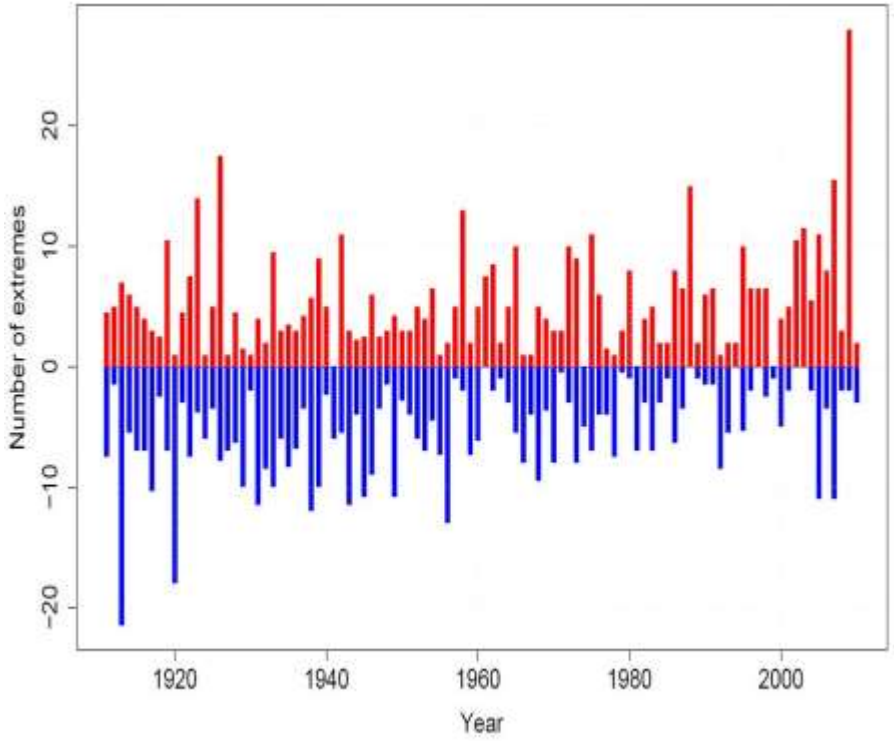
The Australian daily homogenised data temperature data set (ACORN-SAT)



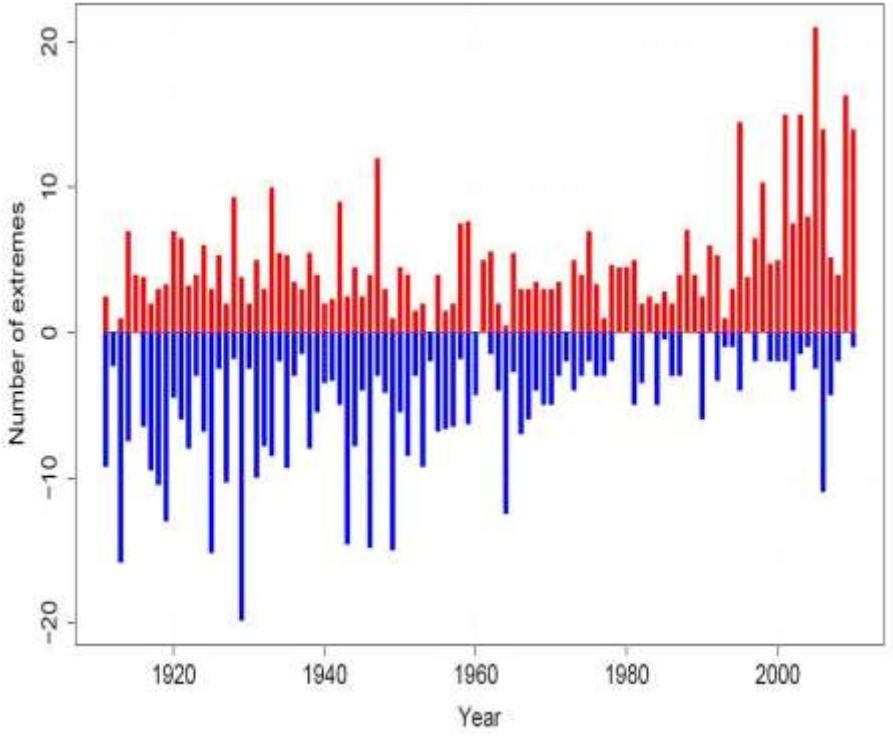
- 112 stations
- Many records go back to 1910
- First national year-round data set of this type

Red – stations in 2001 daily network
Black – stations added in this version

Extremes analysis supported by the ACORN-SAT data set

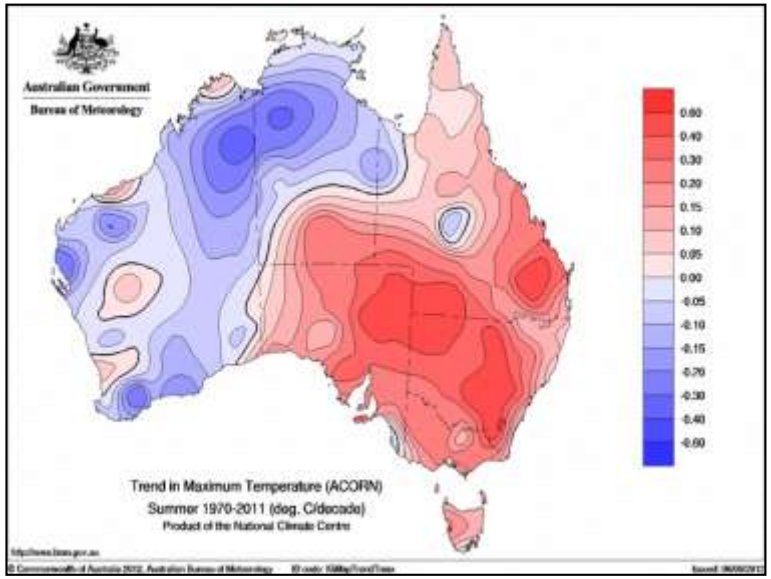


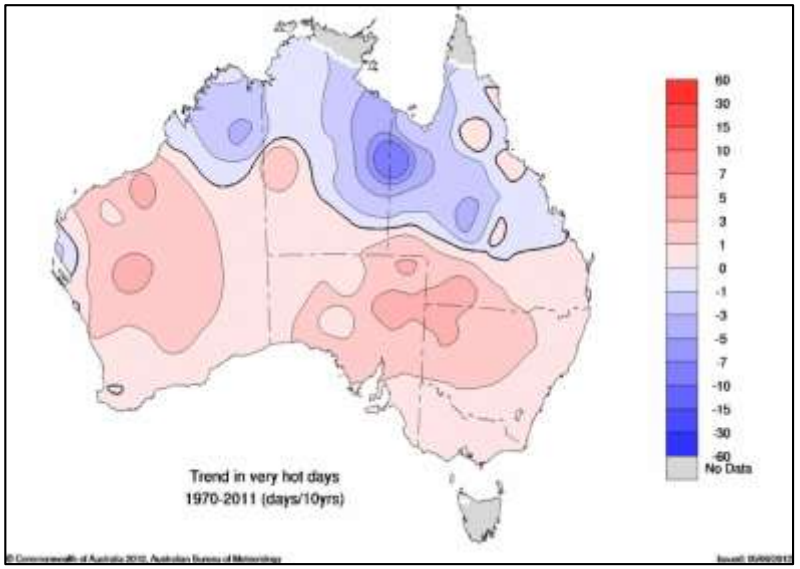
Maximum



Minimum

Some interesting disconnects between extremes and means





Summer mean maximum temperatures have declined over much of WA since 1970, but number of days over 40° C has risen