Australian Government
Bureau of Meteorology

A new gridded daily temperature analysis scheme for Australia

Wahid Zaman, Alex Evans, Andrew Frost, Blair Trewin, David Jones

October 2024

Bureau Research Report - 101

A new gridded daily temperature analysis scheme for Australia

Wahid Zaman, Alex Evans, Andrew Frost, Blair Trewin, David Jones

Bureau of Meteorology

Bureau Research Report No. 101

October 2024

National Library of Australia Cataloguing-in-Publication entry

Authors: Wahid Zaman, Alex Evans, Andrew Frost, Blair Trewin, David Jones

Title: A new gridded daily temperature analysis scheme for Australia

ISBN: 978-1-925738-87-2

ISSN: 2206-3366

Series: Bureau Research Report – BRR101

Enquiries should be addressed to:

Lead Author: Wahid Zaman

Bureau of Meteorology

GPO Box 1289, Melbourne

Victoria 3001, Australia

wahid.zaman@bom.gov.au

Copyright and Disclaimer

Contents

List of Figures

List of Tables

Executive Summary

The Bureau of Meteorology (the Bureau) provides a range of gridded national, daily climate data series generated from surface-based observations, including precipitation, maximum and minimum temperature and vapour pressure. Data for temperature is available from 1910 until the present – see www.bom.gov.au/climate/maps. The Australian Gridded Climate Data (AGCD) is currently provided at 0.05° resolution (approx. 5 x 5 km) and is based on the method described in Jones et al. (2009). The data is used for a range of applications including supporting historical analysis, inputs to modelling and for downscaling, calibration, verification of forecast products, and climate change studies.

The Bureau is updating the methods supporting this service including increasing the resolution of the product to 0.01° resolution (approx. 1 x 1 km) and improving the interpolation using a statistical interpolation approach, to produce the next version of the dataset. The monthly rainfall product described in Evans et al. (2020) is already operational and publicly available, and the daily precipitation is undergoing transition to operations. The work described herein presents the development and evaluation of the updated gridded data of daily maximum and daily minimum temperature, using an extension of the method of Evans et al. (2020). The new approach is evaluated with cross-validation analysis and compared to the performance of the existing product, showing a ~2% improvement in RMSE. In addition to the verification analysis, performance in reproducing observed national and state based extreme minimum and maximum temperature events is evaluated, demonstrating the improved performance of the newly applied method in such events compared to the existing method.

1. Introduction

The Bureau of Meteorology maintains an extensive network of high-quality observational climate and hydrological stations across Australia to monitor current conditions, analyse climatic variability and understand climate change. From these observations, a range of gridded data products are produced. The Bureau's official data for operational climate analysis, the Australian Gridded Climate Data (AGCD)[1] , includes daily and monthly values for rainfall, temperature (daily maximum and minimum), and vapour pressure at 0900 and 1500 local time. The datasets are also used widely within the Bureau and externally for operational and research purposes and are available from www.bom.gov.au/climate/maps and as a Data Collection through National Computation Infrastructure and other means such as FTP for near-real time access.

AGCD data plays an important role in climate, water and agriculture sectors. It is used for various purposes, including qualifying farmers for grant support programmes, monitoring extreme weather events, and forecast calibration and verification. These data are used by the Bureau as input to: various climate monitoring purposes, long-term science assessments of Australian climate such as the biennial State of the Climate report (Bureau of Meteorology and CSIRO, 2022), broader climate science (e.g. Dey et al., 2019; Pepler et al., 2021), water balance and hydrological modelling (e.g. awo.bom.gov.au; see Frost and Shokri, 2021, Vogel et al., 2021, Wilson et al., 2022) and verification of reanalysis (Su et al., 2023), amongst other applications. In sparsely gauged areas with few weather stations and sporadic historical meteorological observations, gridded products serve as crucial historical references to support ecological, agricultural, water resources management, bushfire, drought and climate change studies, allowing estimates of conditions to be obtained for any location.

The methodologies used in the initial AGCD data are described in Jones et al., (2009), and were developed as part of the Australian Water Availability Project (AWAP). The next generation of data is under development within the Bureau, targeting higher resolution (~1km x 1km) and higher accuracy through implementation of statistical interpolation (SI), rather than the existing Barnes analysis approach applied by Jones et al., (2009). The first updated AGCD data released was monthly precipitation (Evans et al., 2020) with a higher resolution (~1km grids) and a higher accuracy and reduced bias compared to the preceding analysis developed by Jones et al., (2009). Further work conducted at the Bureau has recently upgraded the daily precipitation analysis and this data is being transitioned to operations. This report details development and evaluation of new methodologies for minimum and maximum temperature data, for incorporation into ACGD extending on the prior work.

1.1. National and international climate datasets

There are currently three widely used publicly available gridded daily climate datasets within Australia based on Bureau in situ data (Fu et al., 2022).

- Bureau of Meteorology AGCD: www.bom.gov.au/climate/maps
- Queensland Government SILO: www.longpaddock.qld.gov.au/silo
- ANUClimate 2.0: https://dx.doi.org/10.25914/60a10aa56dd1b

[1] AGCD incorporates the suite of datasets previously known as the Australian Water Availability Project (AWAP) datasets. Unless otherwise stated, the former AWAP datasets are considered to be version 1 of AGCD, and the new datasets discussed in this report version 2. AGCD version 1 and AWAP are both used in this text to refer to the older dataset, depending on context.

All of these datasets are available for a range of climate variables, including rainfall and maximum and minimum temperature. The first two are updated daily while ANUClimate is not updated publicly in near real time. All products also seek, in their own ways, to incorporate the influence of topography on climate variables when creating spatially interpolated surfaces of observed daily climate data.

1.1.1. AGCD

The original operational gridded temperature developed by the Bureau (Jones et al., 2009) uses a procedure that combines a three-dimensional spline-based analysis of climatological mean values with an analysis of anomalies using the Barnes successive-correction gridding method as described by Koch et al., (1983), Barnes (1964) and Hutchinson (1995). The climatologies are generated with the use of a trivariate spline function that takes into account a continuously shifting reliance on elevation, as defined by a Digital Elevation Model-(DEM) (Hutchinson 1998). In subsequent advancements, a SI method was implemented to provide a grid spacing of 1 km for monthly rainfall (Evans et al., 2020), which has been applied to daily rainfall. This was found to perform substantially better than the original Barnes analysis approach by Jones et al., (2009) across a range of metrics, due to the optimisation of spatial covariances in the SI method through a choice of length scale which is 0.1 in this case.

1.1.2. SILO

The Queensland Government produces the SILO data, which covers all of Australia and is documented in Jeffrey et al., (2001). It is commercially available as patched point records, which means that missing data at rain-gauge locations are being 'patched' through interpolation, and it is synthetic data generated over a set of evenly spaced points. Many Australian hydrological studies have relied on this product (e.g., Chiew et al., 2008). It is currently produced using a three-step process: a) interpolation of monthly rainfall climatology normalisation parameters using a thin plate smoothing spline (Hutchinson 1993;1998a;1998b); b) ordinary kriging applied to the normalised monthly rainfall; and c) disaggregation of the monthly values per day during each month that were observed.

1.1.3. ANUClimate 2.0

ANUClimate 2.0 consists of gridded daily and monthly climate variables across the terrestrial landmass of Australia from 1970 (or earlier, depending on the variable) to the present. Rainfall grids are generated from 1900 to the present. The underpinning spatial models have been developed at the Fenner School of Environment and Society of the Australian National University. The methods generate climate variables on a 0.01° longitude/latitude grid. Most climate variables have been modelled by expressing each value as a normalised anomaly with respect to gridded 1976-2005 monthly means. These means and the anomalies were all interpolated by trivariate thin plate smoothing spline functions of longitude, latitude and vertically exaggerated elevation using ANUSPLIN Version 4.6, with additional dependences on proximity to the coast for the temperature and vapour pressure variables (Hutchinson and Xu, 2015).

1.1.4. International methods

Internationally, gridded climate data using in situ observations have been created using a wide variety of techniques. For instance, Thornton et al., (2021) created the Daymet V4 based on a combination of interpolation and extrapolation using gaussian kernels. This is a 40-year daily meteorological data on a 1 km grid for North America, Hawaii, and Puerto Rico, including information on temperature, precipitation, shortwave radiation, vapour pressure, snow water equivalent, and day length. The Met Office's Hadley Centre for Climate Science and Services has released a new data of monthly and daily gridded climate observations (1 km grid) for the United Kingdom (Hollis et al., 2019) using inverse distance weighted (IDW) averaging. Daily rainfall and temperature data from Japan and Asia were both gridded using the same IDW approach (Yatagai et al., 2020), although at different resolutions for example, 1 km (25 km) scale was applied to rainfall in Japan (Asia)). A 1 km grid monthly climate database for China was created using spline interpolation techniques (Hong et al., 2005). The Global Precipitation Climatology Centre's (GPCC) suite of gridded global products uses an angular distance weighting scheme (Becker et al., 2013). Gridded climate datasets have also been derived wholly or partly from reanalysis or satellite data, but these can have systematic differences with sets derived from in situ data (Alexander et al., 2020; Chua et al., 2022). Sub-daily interpolations are outside the scope of this report, noting that there is likely to be a point in future where sub daily and reanalysis approaches will merge.

The methods applied within Australia and internationally are typically based only on weighted values of the surrounding observed values (e.g. inverse distance weighted) or on specified mathematical formulas that determine the smoothness of the resulting surface (e.g. thin plate Spline) whereas the SI method is based on formal statistical models that include specification of spatial correlation information. Because of this, SI not only has the capability of producing a prediction surface but also providing some measure of error at each location. This method was adopted for AGCD monthly rainfall analysis due to improved performance (according to standard error metrics including bias and mean error) compared to other methods. The purpose of this report is to assess the performance of the application of the SI methodology previously used for precipitation to daily minimum and maximum temperature.

While AWAP has been the Bureau's primary operational product since 2009 for monitoring climate across a wide range of time scales and a number of different variables, it is not without known issues, including under-representation of extremes and over-smoothing, especially for rainfall (e.g., Beesley et al., 2009; King et al., 2013; Chubb et al., 2016). While reproduction of extremes is important, extremes are clipped as part of the analysis process to some extent for all interpolation methods. Nonetheless, here we examine the reproduction of extreme minimum and maximum temperature events as part of a set of case studies.

1.2. Aim and structure of report

The objective of this report is to document the development of the application of SI using daily temperature observations and assess its performance at producing gridded representations of temperature on a national scale and test its accuracy relative to the current operational temperature analysis across a range of metrics, including examination of extremes.

The structure of the report is:

- **Section 2**. Data and methods: description of observational data, interpolation technique and evaluation and comparison to current methods
- **Section 3.** Results and discussion: evaluation according to cross-validation, case studies of maximum and minimum temperatures
- **Section 4**. Summary and conclusion: key findings

2. Data and Methods

2.1. Data Sources

This research uses daily maximum and minimum temperature data from Australian meteorological stations stored in the Bureau of Meteorology's observational database - Australian Data Archive for Meteorology (ADAM).

Since 1965, the overall size of the temperature network has been relatively stable, between 600 and 800 operating stations. Prior to 1965, and especially prior to 1957, large quantities of observed daily temperature data have not been digitised and are therefore unavailable for use in analyses. As a result, the underlying station network for the temperature analyses is much more sparse in the pre-1965 period, and hence analyses can be expected to be less accurate as represented by traditional verification metrics. There are only minor changes in the availability of underlying station data between AWAP and AGCD, so differences reported here are mostly due to the analysis methodologies.

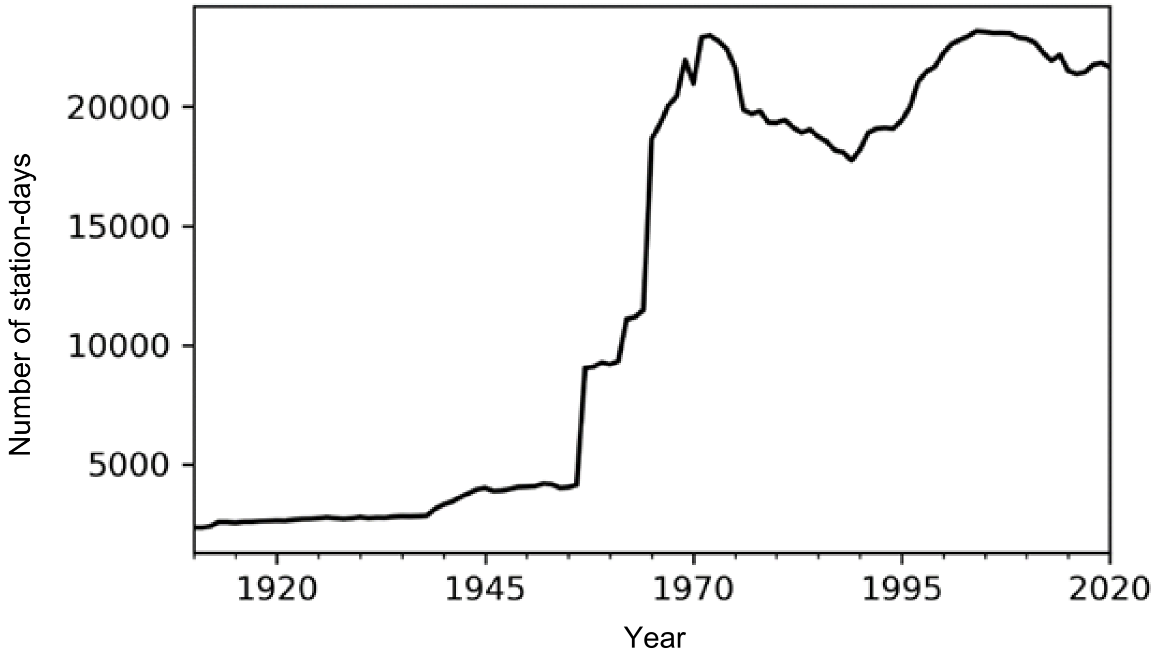

Figure 1: The total station days available in each year for the temperature analyses (1 station reporting every day for one year will be 365 for a non-leap year).

The AGCD temperature data starts in 1910. Prior to 1910, data coverage in many parts of Australia was very limited and the Stevenson screen had not yet been adopted nationally as a standard, leading to substantial temporal and spatial inhomogeneity (Trewin et al., 2020).

2.2 Analysis Methods

2.2.1 Statistical Interpolation

Statistical interpolation is a reliable technique for integrating and interpolating observational data, as demonstrated by Reynolds and Smith (Reynolds, R.W.; Smith, 1994). It involves combining a background field with observations and is commonly used in weather data and climate analyses. The technique produces grid-point values by applying a weighted sum of observation increments to a background field. The background estimate, commonly called the "first-guess" background field in meteorological analysis is a field against which observational increments and the associated correlations structures are defined. The closer the "first guess" field is to the truth, generally the better the analyses.

While SI yields similar results to Barnes analysis (as used in AWAP), we use the SI technique to avoid oscillation problems inherent to successive correction methods and to employ more explicit means of utilising data of differing quality. Evans et al., (2020) successfully applied the approach to generate monthly precipitation using climatology as the background field. Chua et al., (2022) further applied this method using satellite data as the background field for monthly precipitation across Australia and Papua New Guinea (Chua et al., 2023). Here, we have applied SI to minimum and maximum daily temperature data.

Following the annotation used in Chua et al.,(2023), the estimate of the value (A_i) for a particular location i is determined from the weighted sum of observation increments to background values (BG_i). This is written as:

$$A_i = BG_i + \boldsymbol{W}_i \times [\boldsymbol{F}_O - \boldsymbol{F}_B] \tag{1}$$

with

$$\boldsymbol{W}_i = \boldsymbol{B}_i \times (\boldsymbol{B} \times \boldsymbol{O})^{-1} \tag{2}$$

Here, $\boldsymbol{W}_i$ is the vector of weights, $\boldsymbol{F}_O$ is the vector of observations, $\boldsymbol{F}_B$ is the vector of background values at the observation coordinates, $\boldsymbol{B}$ is the background error covariance matrix, $\boldsymbol{O}$ is the observation error covariance matrix, and $\boldsymbol{B}_i$ is the vector of background error covariance between the analysis grid point and the observations.

For the most part, the column vector $\boldsymbol{B}_i$ and the background error covariance $\boldsymbol{B}$ form the most important elements in SI and determine the resulting analysis (Daley 1993).

As with the original AGCD monthly analyses described in Jones et al., (2009), the anomaly difference approach is used in the generation of the gridded spatial analyses. We adopt a zero-mean centred field, consistent with that used for rainfall, where the anomaly field $T'(t)$ is zero-mean centred and defined as:

$$T'(t) = T(t) - \bar{T}_t \tag{3}$$

Equation (3) expresses the temperature as an anomaly, a difference of the temperature observation $T(t)$ from the respective monthly climatology average $\bar{T}_t$. This representation is applicable for both station observations and the gridded analysis and provides an approximately unbiased background estimate. The monthly climatological averages are generated using the smoothing thin-plate spline approach based on the work of Hutchinson (1995).

Hutchinson's (1995) thin-plate spline method is applied to the station climate averages to produce the gridded climate averages with a resolution of 0.01×0.01 degrees (about 1 km × 1 km). It has been shown that smoothing thin-plate splines can effectively reduce error where stations lack longer term observations (Hutchinson and Bischof, 1983), making it a popular tool for analysing climatological surfaces in three-dimensional space in Australia and elsewhere (e.g., Tait et al., 2006; Jones and Weymouth, 1997; Jones and Trewin, 2000,2002; Jones et al., 2006, 2009). From these spline coefficients, an estimated network-wide climatological mean can be calculated for 'every' station that has ever reported an observation.

The SI algorithm fits a mathematical function to a specified number of points, or all points within a specified radius, to determine the output value for each location. The scheme considers overlapping sector domains parameterised for computational efficiency, considering the dimensions of the spatial correlation matrix and associated matrix inversion. Each observation is cross-validated against an interpolated field value derived from all observations in each sector domain (Blomley et al., 1989). To ensure smoothly varying boundaries between sector domains, adjoining domains are extended to include an overlap sector (Glowacki and Seaman, 1987). The sector size, overlap and extension into adjoining domains is determined through adjustable parameter settings (eg. length scale parameter*)* in the algorithm. To optimise the analyses, we trialled several length scale values and sector domain size and evaluated the performance according to cross-validated error metrics. This analysis is detailed in the results.

2.2.2 Climatology

The anomaly-based approach described above (Eq 3) requires input and estimation of the climatology. The climatology can be calculated over the entire series, or a subset of the record.

Jones et al., (2009) divided the climatology into three distinct periods: 1911–1940, 1941–1970, and 1971–2000, with the first and third periods being applied to data before 1911 and after 2000, respectively. Evans et al., (2020) used the years 1981-2010 for climatology analysis. Here we use the years 1991 through to 2020, corresponding to the period for WMO standard climatological normals.

In principle, a climatology that changes through time (1911–1940, 1941–1970, etc.) may be preferable if the network is relatively stable within periods and if there are marked shifts in the mean between periods. Unfortunately, network changes, particularly over inland parts and in regions near the coast with strong climatic gradients, trigger unreasonably large differences in the climatology between adjacent 30-year periods. It was largely for this reason that we have chosen to use the period that had highest density of data for the climatology for all years of analysis.

The anomaly approach used here requires the use of a climatology applied for each month (taking the days in each month's average value), based on interpolated monthly values over sites with adequate records. Any given 30-year climatology period will only allow climatologies to be calculated for a proportion of stations, as some stations will have only been operating for part (or none) of that 30-year period, and others may have missing data. The updated climatology period (1991-2020) has more available stations and therefore, forms a more complete dataset. Although Hopkinson et al., (2012) demonstrated that 30-year mean estimates can be calculated using regression for stations with shorter observation records using data from nearby stations with longer term averages, and WMO (2007) also found that 10-12 years of data were sufficient to simulate a 30-year average with reasonable skill, for this work, in general, we have only used stations with a minimum of 21 valid months of data for each calendar month between 1991 and 2020.

Due to the scarcity of data from stations at higher elevations, at locations above 1000 m elevation a more lenient guideline requiring only 14 valid months of data was implemented. Stations located at high altitudes contribute significantly to the overall analysis by quantifying the vertical gradient of climate averages. In their final analysis, Jones et al., (2009) highlight that the incorporation of extra high elevation stations improved the representation of high elevation climate by a small but significant amount.

2.2.3 Evaluation and comparison to previous method

We evaluate the accuracy of AGCD compared to the current product (labelled AWAP) through the calculation of error statistics based on data generated through leave-one-out cross-validation. Cross-validation error statistics (root mean squared error and bias) are generated:

- **Globally** – an overall summary measure across the entire series.
- **Temporally** – month-based statistics of the nationally averaged errors.
- **Spatially** – site-based statistics to show locational biases and errors.

Cross-validated error statistics are available for AGCD for the full post-1910 period and from 1980 onwards for AWAP.

We calculate the Root Mean Square Error (RMSE), mean bias (BIAS) and Mean Absolute Error (MAE) of the analysis at the observation stations (see Appendix 1 details in Evans et al.,2020). It is noted that there are differing equations depending on the error being calculated globally, on a site basis (spatially) or temporally.

Consider a cross-validated estimate of a station value T at station k and time t, denoted by $T(x_k, y_k, z_k, t)$. The cross-validated analysis error is given by:

$$E_k(t) = T(x_k, y_k, z_k, t) - T_k(t) \quad (4)$$

Noting that the error will differ depending on the station and whether the analysis is for daily (global/site) or monthly (temporal) statistics. The observation consists of a "true" and an "observational error" component. The "true" component is what would be measured if the observation at the station were perfectly accurate, whereas the "observational" error component is the error introduced due to factors such as instrument miscalibration, misreading by the observer, errors in spatial representativeness resulting from specific factors at the observation site, etc. On the assumption that the observational errors are statistically independent of the interpolated and actual values (Daley 1993), this can be simplified further (see Evans et al., 2020 for the specifics). Clearly, even a faultless analysis will have a non-zero error because observations contain error. To acquire a cross-validation error of zero, the observations must be "perfect," while acknowledging that analysis techniques are not perfect. Although the cross-validated differences between independent observations and analyses are commonly referred to as "analysis errors" and are defined as such here, they include observational error. Daley (1993) and Jones and Trewin (2000) describe the estimation of observational errors, at least at the large scale. Additional measures of bias are computed when defining the averaged time- and station-based analysis errors.

3. Results and Discussion

3.1 Evaluation according to cross-validation

3.1.1. Overall summary statistics

Table 1 presents the summary daily analysis cross-validation error statistics for the standard climatological reference period covering 30 years (1991-2020) for AWAP vs AGCD. AGCD improves marginally for both Tmax and Tmin (by 0.017° /0.016° or ~1%) compared to that of AWAP according to RMSE, and to a greater extent according to bias (reducing by 0.008°/0.015° or 50%). This indicates an improvement overall according to the summary metric.

It is noted that various correlation length scales were trialled (0.1, 0.5, and 0.8) and cross-validation statistics were calculated (see Appendix A Table S1). The length scale of 0.1 was found to have the optimal site and temporal statistics, so was chosen for application and presentation here.

Table 1: Overall average error statistics of Tmax and Tmin during 1991-2020 (using length scale = 0.1)

	RMSE (°C)		Bias (°C)	
	AWAP	AGCD	AWAP	AGCD
Tmax	1.289	1.272	0.016	0.008
Tmin	1.800	1.784	0.030	0.015

RMSE values for Tmin are higher than those for Tmax in both AWAP and AGCD, reflecting the fact that Tmin is more influenced than Tmax by small-scale processes such as cold air drainage that are unlikely to be fully resolved by the observing network.

3.1.2. Temporal error analyses

Using AGCD (1910-2020) and AWAP (1980-2020) data, we calculate the Tmax and Tmin error statistics on a monthly basis to track how errors change over time and according to station density. Summary statistics are for the common period 1991-2020 provided in Table 2 as this period has the greatest density of stations, and greatest consistency of observational practices.

Table 2: Temporal statistics of Tmax and Tmin during 1991-2020 (correlation length scale – 0.1)

Variable	RMSE (°C)		BIAS (°C)	
	AWAP	AGCD	AWAP	AGCD
Tmax	1.265	1.245	-0.0123	-0.0023
Tmin	1.774	1.759	0.026	0.012

The AGCD data RMSE is 1.6% for Tmax and ~1% for Tmin lower than AWAP, and bias is at least half that of AWAP also, demonstrating improved temporal aggregation properties in addition to overall characteristics.

It is noted that the correlation length scale of 0.1 was found to be optimal for temporal statistics (similar to spatial statistics) as presented in Appendix Table S2, so the results presented here relate to correlation length scale 0.1.

Figure 2(a) and (b) presents a boxplot of the monthly RMSE values for Tmax and Tmin. Compared to AWAP, there is marginal improvement in RMSE. Figure 3(a) and (b) presents a boxplot of the monthly bias values for Tmax and Tmin. Compared to AWAP, a considerable reduction in BIAS in AGCD relative to AWAP.

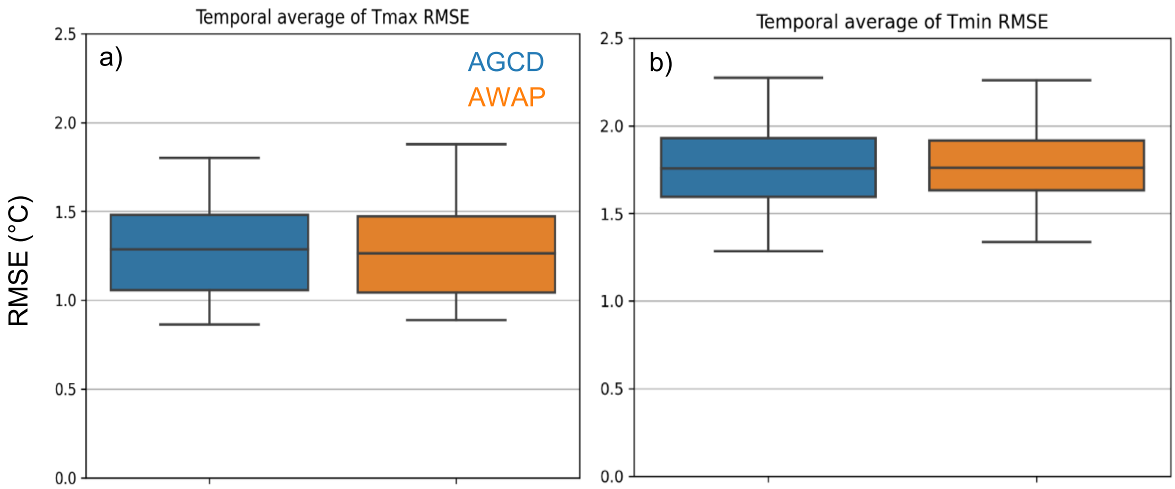

Figure 2: Temporal monthly average of RMSE (a) Tmax and (b) Tmin for AGCD (blue) and AWAP (orange) during 1991-2020. Box plots show the five-number summary of RMSE: including the minimum score, first (lower) quartile, median, third (upper) quartile, and maximum score.

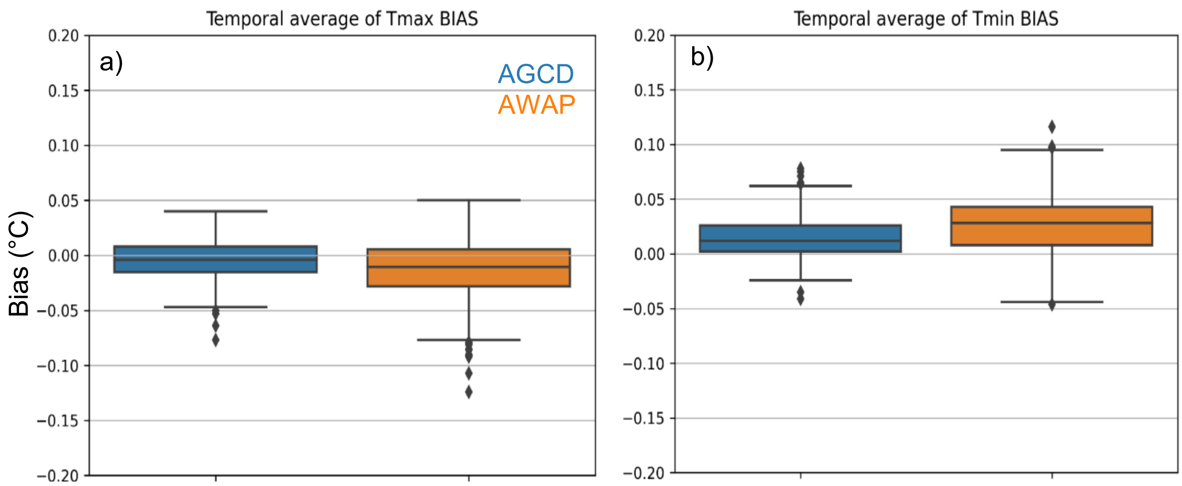

Figure 3: Temporal monthly average of bias (a) Tmax and (b) bias for AGCD (blue) and AWAP (orange) during 1991-2020. Box plots show the five-number summary of Bias: including the minimum score, first (lower) quartile, median, third (upper) quartile, and maximum score. Outlier (shown as point) is defined as a data point that is located outside the whiskers of the box plot.

Towards understanding the interannual variability of the metrics, Figure 4 and Figure 5 plot the RMSE (a), bias (b), and MAE (c) from one month to the next for Tmax and Tmin respectively.

Particularly intriguing is the comparison of the Tmax RMSE and bias time series over the post-1980 period (Fig 4d-f); the AGCD bias is very small in magnitude, slightly negative, and generally constant over time, whereas the AWAP bias is larger in magnitude and displays more variability, including a clear seasonal cycle. When the calculated value is greater than the observed value, the station is said to be biased in a positive direction. The AWAP study is based on a 3-epoch, rolling climatology, which may account for the observed change in the AWAP bias.

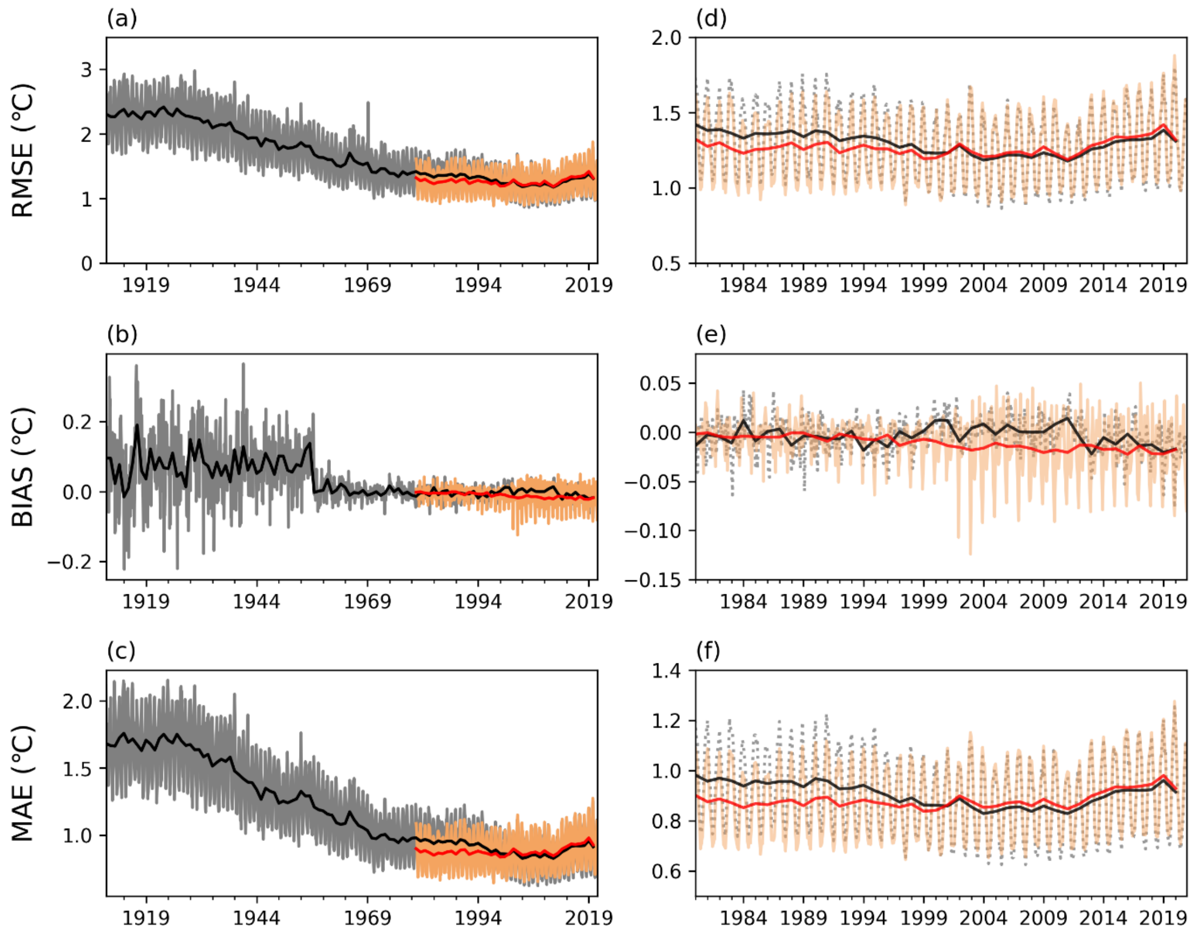

Figure 4: Temporal distribution of Tmax error statistics (°C) during 1910-2020 (left) and enlarged for 1980-2020 (right). Gray and orange line shows the monthly distribution for AGCD and AWAP respectively whereas the black and red line shows their yearly mean.

For both maximum and minimum temperatures, the RMSE and MAE metrics show a fairly steady decreasing trend over the period from 1910 to about 1970, followed by a stabilisation after 1970. This is consistent with the increasing density of the available network up until the 1960s. However, the bias metrics show an abrupt change in behaviour around 1957; prior to that date there is much higher month-to-month variability of bias, as well as a noticeable step change in mean bias, particularly for maximum temperature (see also Appendix B Figure S1).

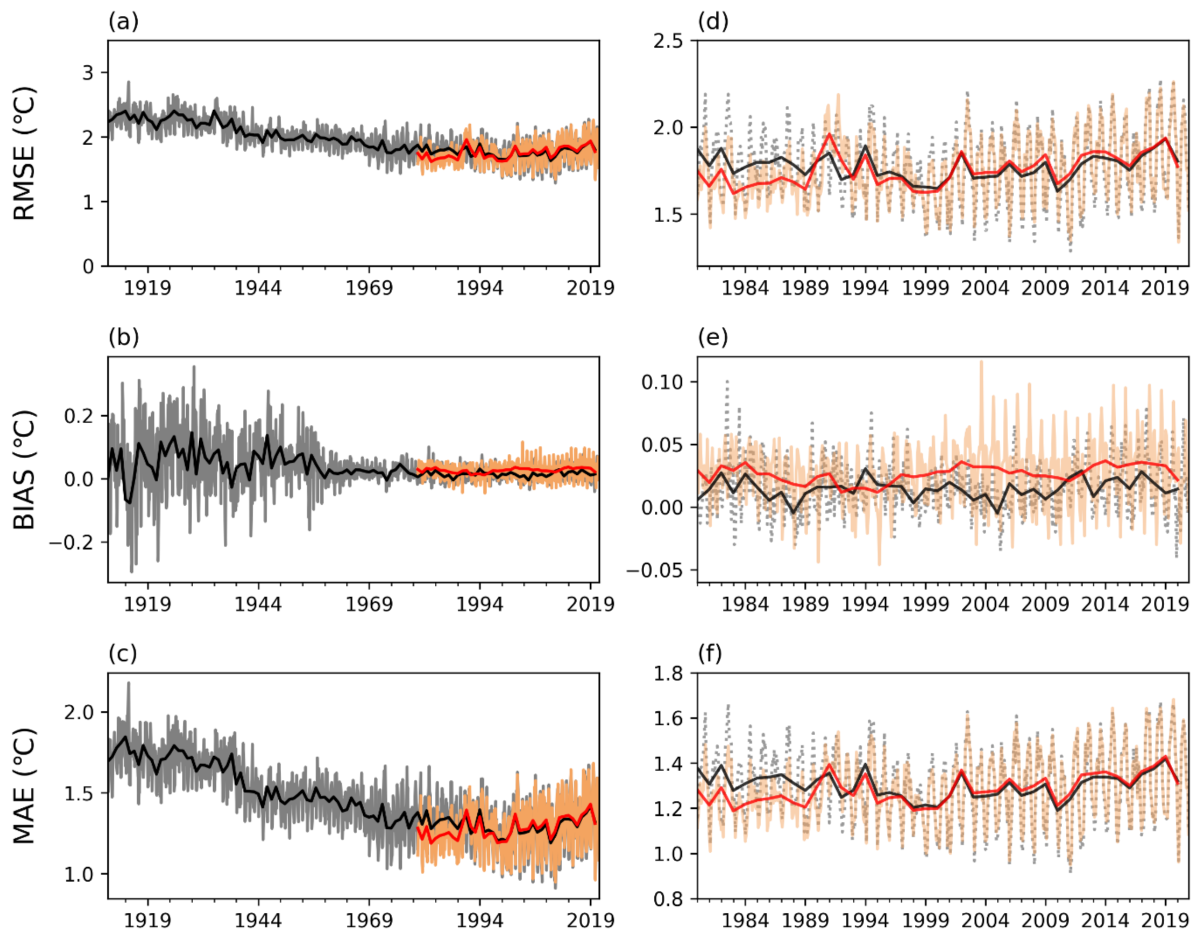

Figure 5: Temporal distribution of Tmin error statistics (°C) during 1910-2020 (left) and enlarged for 1980-2020 (right). Gray and orange line shows the monthly distribution for AGCD and AWAP respectively, whereas the black and red line shows their yearly mean.

Exploring the change in the behaviour of bias statistics over time in more detail, the amount of month-to-month variability in bias, and the absolute values (particularly for maximum temperature), shifts sharply in 1957. This coincides with a large increase in data availability, as substantial quantities of pre-1957 daily temperature data have not been digitised and are therefore not available to the analysis. Considering the date 1 February (a weekday in both years and hence not affected by potential missing weekend data, a common issue in that era) as an example, there are 135 stations nationally with maximum and/or minimum temperature available in 1956, but 304 available in 1957. There is a further large increase in the availability of digitised data in 1965 (on 1 February 1965 there are 606 stations available), but this has only a modest additional impact on the behaviour of bias statistics at national scale, indicating that there is a critical level below which network density is insufficient to maintain stable bias statistics.

We now focus on a specific example with max and min temperature temporal error statistics calculated for the Northern Territory (Figure 6). The Northern Territory has a relatively sparse network, particularly historically. Prior to 1957, on most days there are only 3 or 4 stations within the Northern Territory with available data (Darwin, Tennant Creek, Alice Springs and sometimes Daly Waters), although the analysis is also influenced by data from stations in border areas of Western Australia and Queensland such as Halls Creek, Burketown, Camooweal and Boulia. On most days in the 1957-1964 period there were fewer than 10 stations, increasing to over 30 stations from 1965 onwards.

As would be expected, both bias and RMSE show marked changes in behaviour going backwards in time, particularly in 1957 and in 1939 (data from both Daly Waters and the Queensland border station of Camooweal start in 1939). However, an interesting aspect of the Northern Territory error statistics is that there is a noticeable change in minimum temperature bias around 1969, when the level of noise in monthly bias statistics is still relatively low. There were two noteworthy changes in the Northern Territory network at this time. One was the opening of a station at Rabbit Flat in what was previously an extremely data-sparse area, and the other was the replacement of the Tennant Creek Post Office site with the Tennant Creek Meteorological Office. Both locations have large footprints in the analysis because of local data sparsity, and it is known from the data homogenisation carried out as part of the development of the ACORN-SAT data (Trewin et al., 2020) that mean annual minimum temperatures pre-1969 at the Tennant Creek Post Office site were approximately 0.7 °C lower than those post-1969 at the Meteorological Office (most likely largely driven by local topographic differences between the two sites). This indicates the potential for changes in the underlying network to affect the longer-term characteristics of the AGCD temperature analysis, particularly in data-sparse areas.

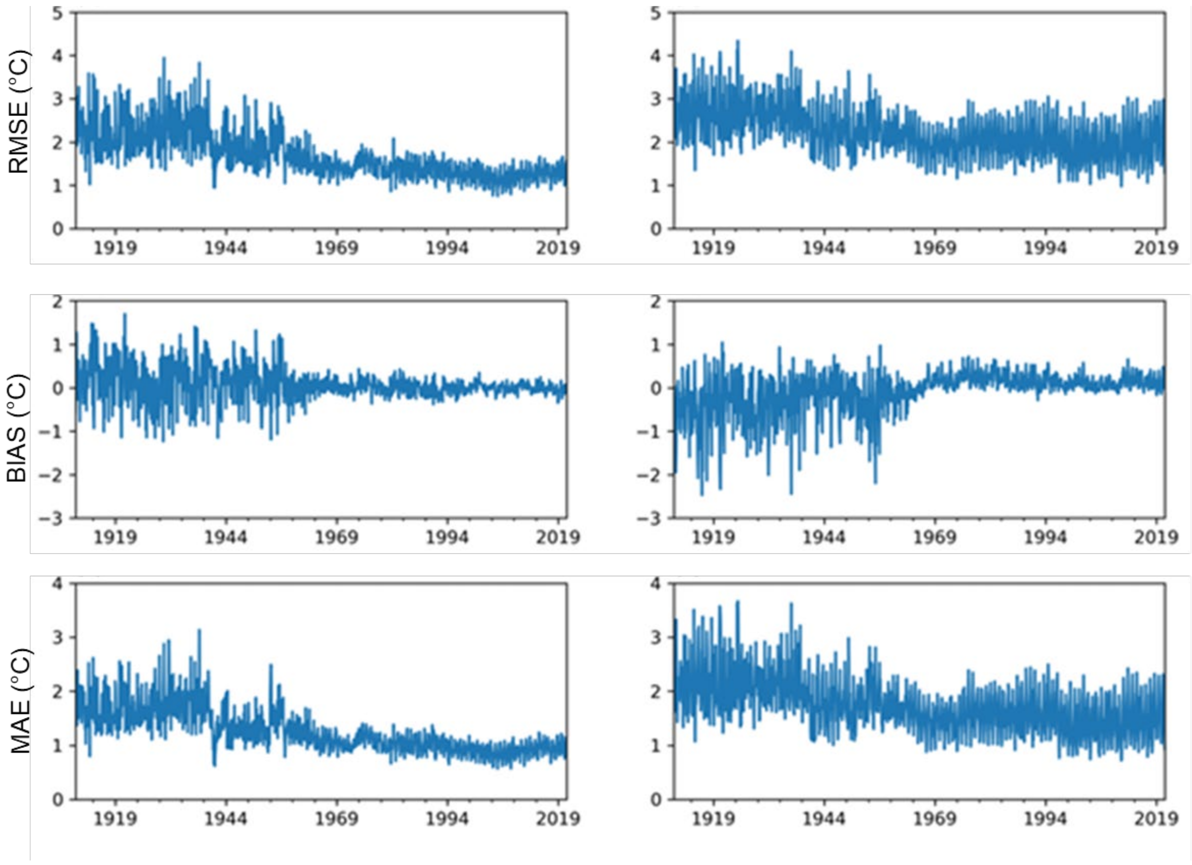

Figure 6: Temporal distribution of RMS, bias and MAE statistics for the Northern Territory, for (left) maximum and (right) minimum temperature.

3.1.3. Spatial error analyses

The Tmax and Tmin error statistics were calculated on a site basis for the period 1991-2020 to understand local biases and errors and are summarised in box-plot form in Figure 7. Where the calculated value is larger than the observed value, there is a positive bias at that station. The two figures show broadly similar properties although, as expected from the nationally aggregated results shown in Table 1, absolute RMSE values are typically larger for Tmin than Tmax.

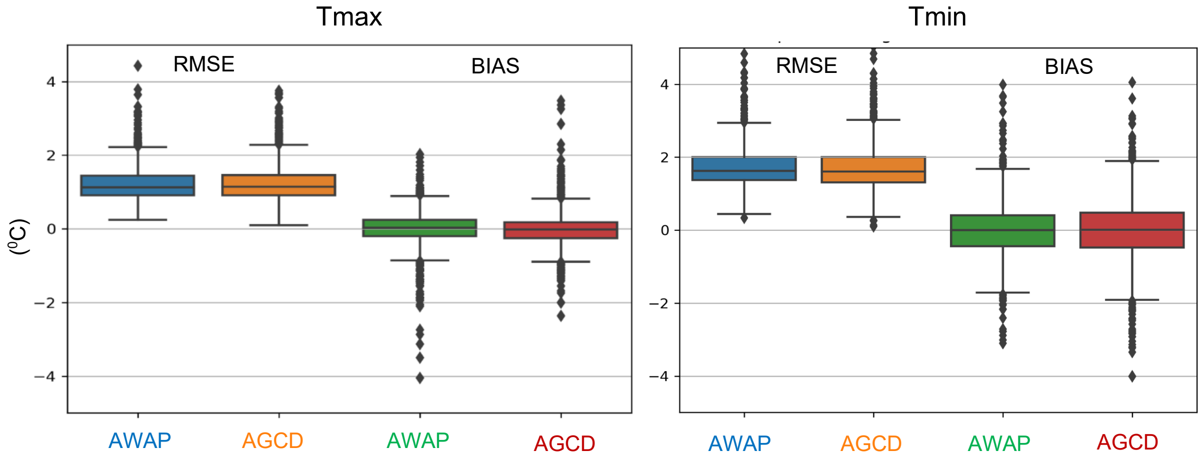

Figure 7: Distribution of station based Tmax and Tmin error statistics, calculated over the period 1991-2020.

In Figs. 8-9, we show the spatial distribution of the RMSE and BIAS of AWAP and AGCD (containing the absolute value AWAP minus the absolute value of AGCD) at stations and mapped nationally. Tmax RMSE (shown in green) is in the lower range (0.5-1.5 °C) across Australia. When the absolute value of AGCD minus the absolute value of AWAP is negative (shown in blue in Fig. 8c,f), AGCD is performing better than AWAP. Compared to Tmax, the RMSE and bias for Tmin are larger (Figs. 8 and 9), but even so, they are still quite small. While there are few clear spatial patterns, AGCD tends to perform better than AWAP for Tmax at coastal locations in southern Australia (Fig. 8c), and overall, it is clear that errors are typically largest where topography is variable and in the vicinity of the coastline, where strong gradients exist in temperatures and the networks are often insufficient to properly constrain the temperature variations in the analysis. It is likely that these patterns are evident nationwide, including in those areas without stations which can constrain the analysis of errors.

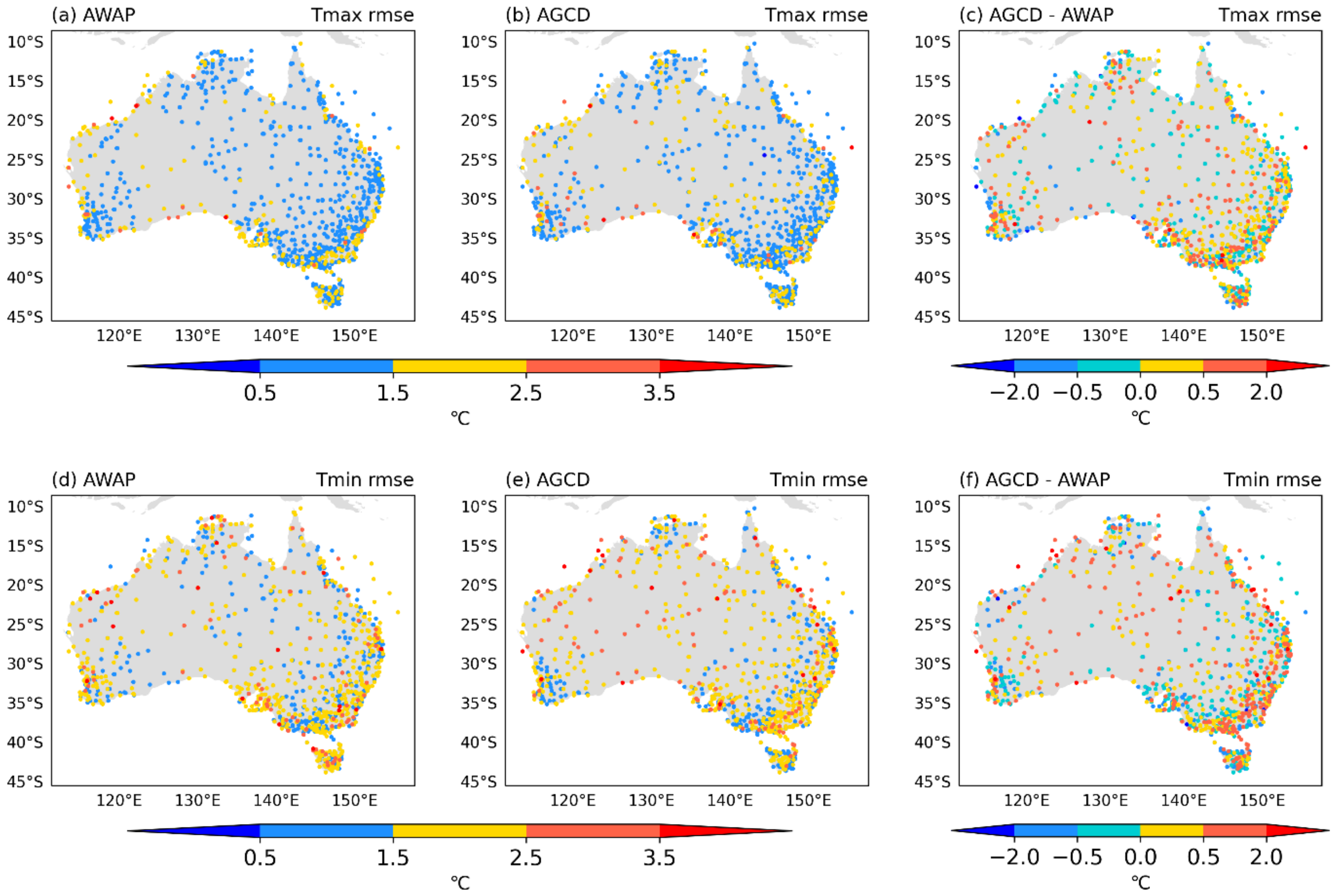

Figure 8: Spatial distribution of Tmax and Tmin RMSE across Australia for AWAP and AGCD, and the difference between them.

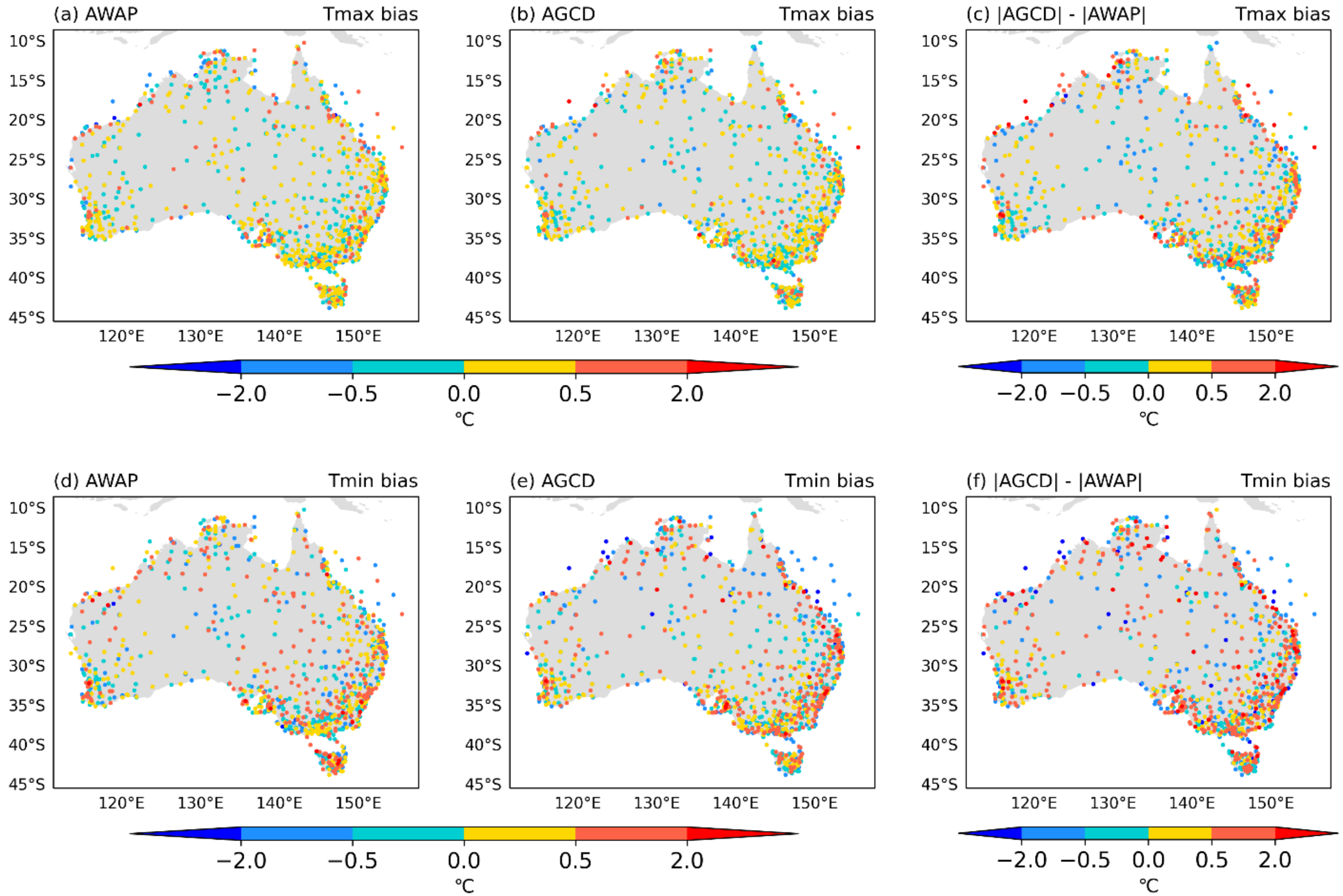

Figure 9: Spatial distribution of Tmax and Tmin bias across Australia for AWAP and AGCD, and the difference between absolute biases in the two analyses.

3.2. Extreme minimum/maximum temperatures

3.2.1 Case studies

One area of specific interest is how well the new AGCD dataset resolves significant extremes. As part of this, events which resulted in the highest recorded maximum (Table 3) and minimum (Table 4) temperatures in each state and territory were considered. Four of these are presented below as case studies.

Table 3: Recorded maximum temperature for each state and territory.

Site No.	Location	State/ Territory	Date	Observed record (°C)	Temp AWAP (°C)	Temp AGCD (°C)
17043	Oodnadatta	SA	02 Jan 1960	50.7	49.7	50.3
5017	Onslow	WA	13 Jan 2022	50.7	49.4	50.3
46043	Wilcannia	NSW	11 Jan 1939	50	49.7	49.8
38002	Birdsville	Qld.	24 Dec 1972	49.5	48.7	49.3
77010	Hopetoun	Vic.	07 Feb 2009	48.8	47.9	48.5
15526	Finke	NT	02 Jan 1960	48.3	48.2	47.6
92094	Scamander	TAS	30 Jan 2009	42.2	40.2	41.7

Table 4: Recorded minimum temperature for each state and territory.

Site No	Location	State/ Territory	Date	Observed record (°C)	Temp AWAP (°C)	Temp AGCD (°C)
71003	Charlotte Pass	NSW	29 Jun 1994	-23	-14.8	-17.3
96033	Liawenee	TAS	07 Aug 2020	-14.2	-11.6	-12.8
83025	Omeo	VIC	15 June 1965	-11.7	-8.8	-10.2
41095	Stanthorpe	Qld	23 Jun 1961	-10.6	-9.1	-9.7
19062	Yongala	SA	20 July 1976	-8.2	-6.7	-6.9
15590	Alice Springs	NT	17 July 1976	-7.5	-6.3	-8.4
11019	Eyre Bird Observatory	WA	17 Aug 2008	-7.2	-2.5	-3.6

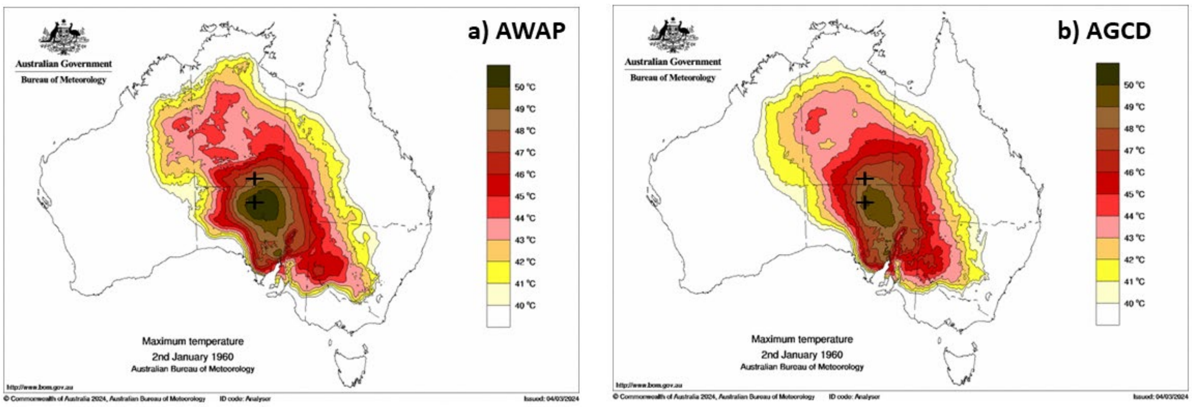

Figure 10: Distribution of maximum temperature in South Australia and the Northern Territory using a) AWAP and b) AGCD on 2 January 1960. Plus (+) sign indicates the location of the highest observations in South Australia (Oodnadatta, 50.7 °C) and the Northern Territory (Finke, 48.3 °C).

In general, the state and territory high extremes are well resolved by both AWAP and AGCD (though the latter has smaller errors overall), reflecting the fact that these extremes mostly occurred in relatively flat terrain where maximum temperature fields are expected to be relatively smooth. The largest differences between the two analyses occur for the two extremes which occurred nearer the coast, being those for Western Australia and Tasmania.

The first case study, see Figure 10, focusses on 2 January 1960, when record high temperatures were observed for South Australia and the Northern Territory (Table 3). The highest temperature analysed by AGCD at the nearest point to Oodnadatta is 50.3 °C, while the corresponding value in the AWAP analysis is 49.7 °C. In general, the AWAP analysis is hotter than AGCD in the areas with the most extreme heat, with an extensive area of temperatures greater than 50 °C analysed east of Oodnadatta in AWAP which is not replicated in AGCD. As there are no observations in this region, it cannot be determined definitively which of the two has resolved the event more effectively.

The second case study for extreme heat focusses on 30 January 2009 (Figure 11), when a state record was set for Tasmania, of 42.2 °C at Scamander. On this day on the northern east coast of Tasmania, there was a mix of extreme high temperatures (with 42.2 °C at Scamander and 41.8 °C at St. Helens Airport) and lower temperatures, with only 26-28 °C at the more exposed locations of Larapuna (Eddystone Point), Bicheno and Friendly Beaches. Neither analysis fully resolves the extent to which the extreme heat approaches the east coast with peak values along the coast in the 36-39 °C range, but AGCD analyses an area exceeding 39 °C just inland from the east coast, which is too small to be visible in AWAP, and also better resolves the relatively cool conditions from Bicheno southwards. At Scamander specifically, the analysed AGCD temperature (41.7 °C) is 0.5 °C lower than observed, whereas AWAP (40.2 °C) is 2.0 °C lower than observed.

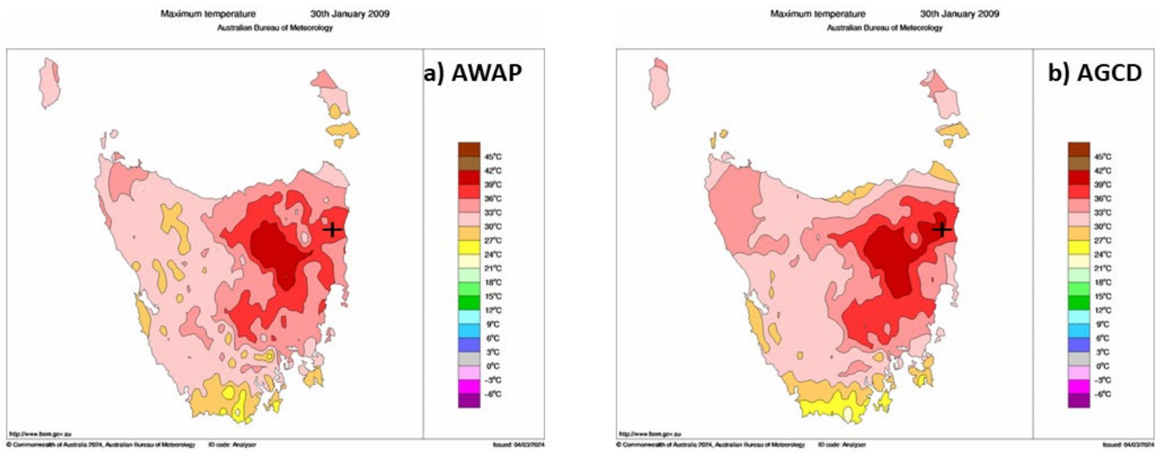

Figure 11: Distribution of maximum temperature in Tasmania on 30 January 2009. Plus (+) sign indicates the location of the highest observations in Tasmania.

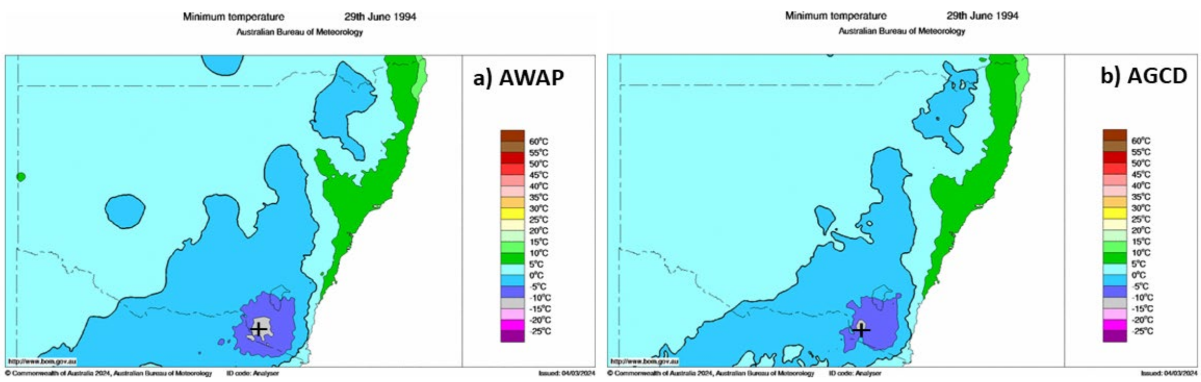

Figure 12: Distribution of minimum temperature in New South Wales on 29 June 1994. Plus (+) sign indicates the location of the lowest observations in New South Wales.

Turning to case studies of extreme cold, we find that for six of the seven listed State and Territory extremes, both AWAP and AGCD are warmer than observed in the listed events. This reflects the influence that local-scale topography has on extreme low minimum temperatures, with variations near or above 10 °C within a few kilometres not unknown under some conditions. It is not surprising that gridded analyses fail to fully resolve such localised effects. In general, AGCD more closely approaches the observations than AWAP, doing so in five of the six cases where the observation is colder than both.

First, we consider 29 June 1994 (Fig. 12), when the Australian record low temperature of −23.0 °C was set at Charlotte Pass. Charlotte Pass is topographically favoured for extreme minimum temperature, being located in the bottom of a valley with regular winter snow cover. Of its nearest neighbours with broadly similar elevations (within 200 m), Perisher Valley (−18.0 °C) has largely similar topography, but Thredbo Top Station (−9.5 °C) is a much more exposed location, and it is not unusual for Charlotte Pass to be 10 °C or more colder than Thredbo Top Station under conditions favourable to cold air drainage. We would not expect a gridded analysis to fully resolve these sharp local topographically-driven minimum temperature gradients. AGCD approaches the observed value more closely, with the analysed value of −17.3 °C being 5.7 °C warmer than observed, compared with an 8.2 °C difference for AWAP.

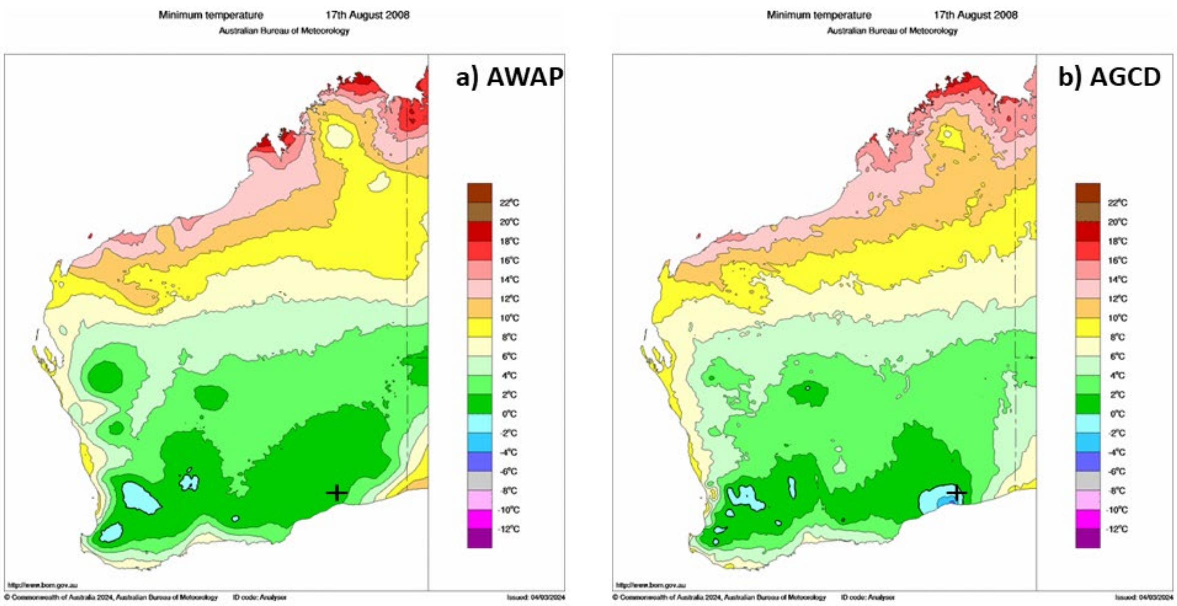

Figure 13: Distribution of minimum temperature in Western Australia on 17 August 2008. Plus (+) sign indicates the location of the lowest minimum observations in Western Australia

The final case study is from 17 August 2008 (Fig. 13), where a record low of −7.2 °C for Western Australia was set at Eyre, on the Nullarbor coast. Eyre is a location favoured for extreme low minimum temperatures; although the site is less than 1 km from the coast, the instruments are located in an enclosed depression in sand dunes which regularly experiences cold air drainage under clear, calm conditions. The Nullarbor is a very challenging environment for minimum temperature analyses because of the sparseness of the observation network and the large topographic contrasts between sites; in particular, the Eucla site (at the top of an escarpment) is in a highly unfavourable location for very low minimum temperatures. On 17 August 2008, Eucla and Nullarbor Roadhouse, to the east, had minimum temperatures above 10 °C, with other Nullarbor sites further west (such as Forrest and Balladonia) having minimum temperatures between 0 and 2 °C.

3.2.2 Treatment of areas with spurious extremes in AWAP

A known issue with the previous AWAP temperature analysis is that it is susceptible to producing spurious extremes in regions with sharp temperature gradients and significant network changes. One noteworthy example of this is in the Nullarbor region of far western South Australia, around the (former) settlement of Cook. Observations at Cook ceased in 1998, and following that date, the nearest observations are at Nullarbor Roadhouse (104 km south and much closer to the coast), with the nearest inland observations being at Forrest (222 km west) and Tarcoola (399 km east).

Nullarbor has similar extreme high temperatures to Cook but much cooler mean summer maximum temperatures (over the 1987-1997 period of common observations at the two sites, the highest temperature of the year on average is 0.3 °C hotter at Nullarbor, but the mean summer maximum is 4.8 °C cooler), and hence has much higher temperature anomalies on extreme hot days than Cook. Without observations at Cook to constrain the analysis, the very high anomalies at Nullarbor are analysed too far inland, and these are added in AWAP to a climatology which includes Cook data to produce spurious high values in the Cook area. This is visible in Figure 14, which shows daily maximum temperature anomalies from extreme hot days in 1990 (when Cook was open) and 2007 (when it was not), with the area of anomalies above +18 °C analysed substantially further inland in the latter case.

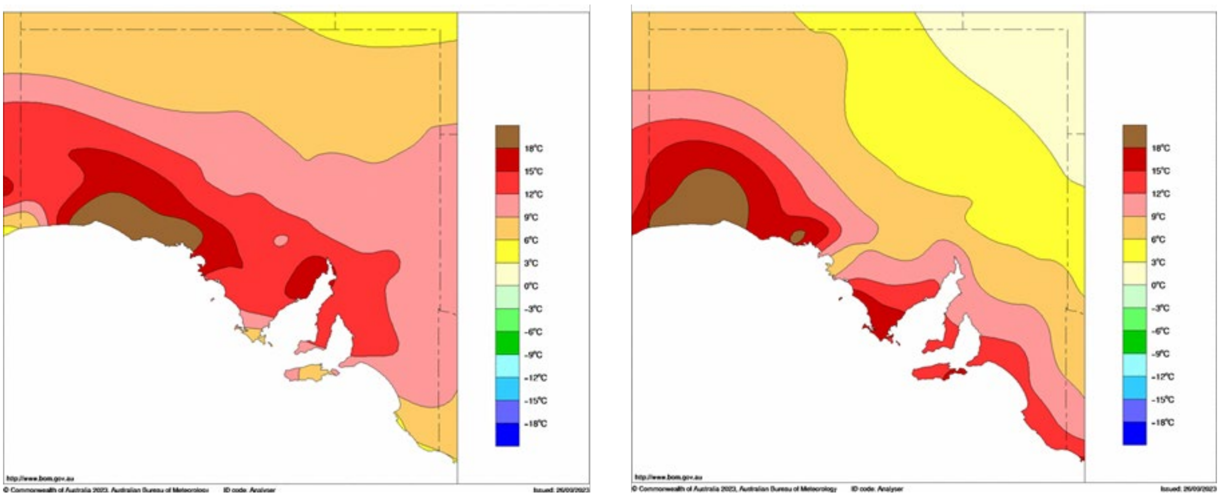

Figure 14: Maximum temperature anomalies (from 1961-1990 baseline) on 2 January 1990 (left) and 17 February 2007 (right). On 2 January 1990, 48.5 °C was observed at Nullarbor Roadhouse and 48.0 °C at Cook. On 17 February 2007, 47.6 °C was observed at Nullarbor Roadhouse with no observations at Cook.

To assess how the two analyses perform in this area, we consider the highest temperature of each year at the approximate location of Cook (30.6 °S, 130.4 °E) in the two analyses, alongside the ACORN-SAT station data at the geographically comparable locations at Forrest and Tarcoola (Figure 15). Prior to the closure of Cook in 1998, the four series compare very closely, with the 1962-1997 mean being 45.1 °C for both the AWAP and AGCD gridded analyses, 45.4 °C at Forrest and 44.5 °C at Tarcoola. However, there is a sharp change in behaviour after 1998, with AWAP consistently substantially warmer than Forrest and Tarcoola, and AGCD generally cooler. The post-1998 AWAP analysis also has numerous physically unrealistic values, with 52 °C exceeded on numerous occasions and a highest value of 54.99 °C on 19 December 2019 (when Nullarbor Roadhouse reached 49.9 °C, an Australian record for December).

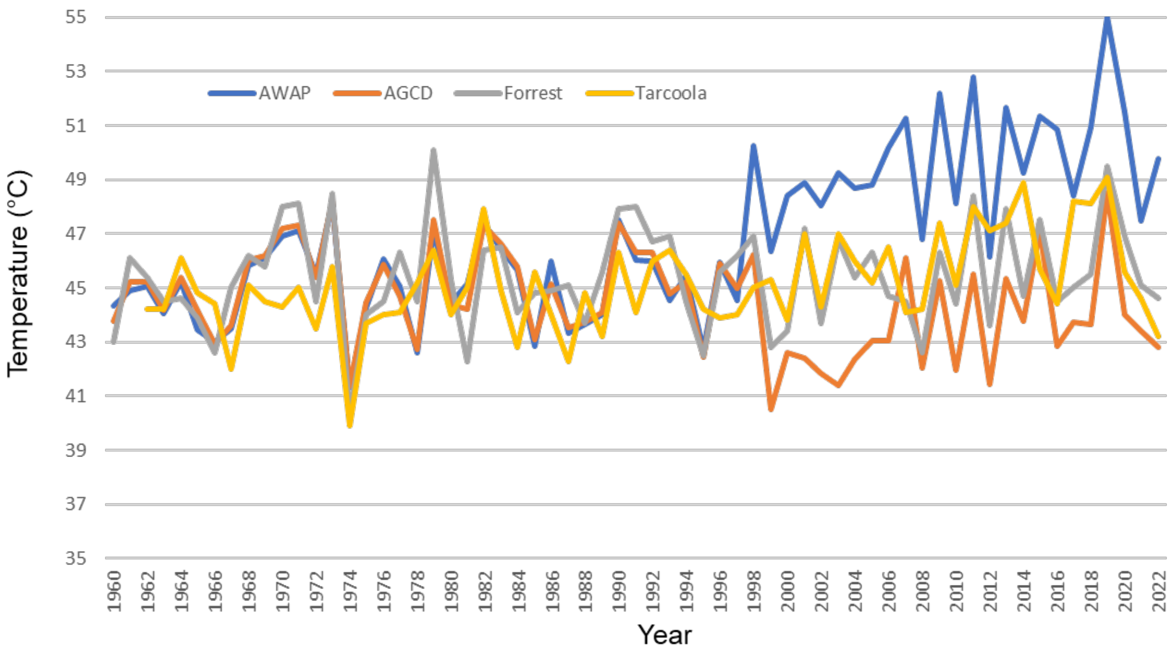

Figure 15: The highest temperature of each year at Forrest and Tarcoola, and at 30.6 °S 130.4 °E in AWAP and AGCD.

Averaged over the 1999-2022 period, AWAP is 6.2 °C warmer than AGCD. Using the Forrest and Tarcoola station data as a baseline indicates a likely bias in annual heat extremes in the order of +4 °C in AWAP and −2 °C in AGCD. This indicates that the new AGCD analysis eliminates the spurious heat extremes (positively biased) at this analysis location, although it produces a opposite, but smaller bias. It is not surprising that AGCD does lead to a slight dampening of the extremes as the sum of weights applied in the algorithm is less than one implying slightly reduced variances even in data rich environments.

4. Summary and Conclusions

In this report, we describe a new national daily temperature analysis product- AGCD developed by the Australian Bureau of Meteorology. The overall summary error statistics are better relative to its predecessor approach, AWAP. While the performance gains made for AGCD temperature relative to AWAP are modest compared with those for the equivalent transition for rainfall, there are operational benefits in having a consistent analysis method in use across all variables. In future studies, we plan to apply this kind of analysis to even more variables (ie. vapour pressure and monthly analyses). The approach is reliable, as it maintains the baseline climatology over time, and it is computationally efficient, allowing whole historical analyses to be re-run in a relatively short amount of time. One of the many advantages of using the SI technique is it is simple to amend the approach to include information from new datasets. For instance, for a rainfall analysis, including radar, satellite rainfall estimates, or projections from a short-range forecast model is relatively straightforward, with good progress already having been achieved in this regard. Our findings indicate that the analysis errors reduce only slightly with an increasing total number of stations once that number has reached the values which were characteristic of the Australian network after 1957. A substantial expansion of the network would therefore be unlikely to result in a substantial improvement to the analysis at a national scale, though it might still have value at a more localised or regional scale, such as for targeting specific observation gaps.

References

Alexander, L. V, Bador, M., Roca, R., Contractor, S., Donat, M. G., & Nguyen, P. L. (2020). Intercomparison of annual precipitation indices and extremes over global land areas from in situ, space-based and reanalysis products. *Environmental Research Letters*, *15*(5), 55002. https://doi.org/10.1088/1748-9326/ab79e2

Barnes, S. L. (1964). A Technique for Maximizing Details in Numerical Weather Map Analysis. *Journal of Applied Meteorology and Climatology*, *3*(4), 396–409. https://doi.org/https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., & Ziese, M. (2013). A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, *5*(1), 71–99. https://doi.org/10.5194/essd-5-71-2013

Beesly, C.A., Frost, A. J., & Zajaczkowski, J. (2009). A comparison of the BAWAP and SILO spatially interpolated daily rainfall datasets. *18th World IMACS/MODSIM Congress*.Cairns, Australia.

Blomley, J. E., Smith, N. R., & Meyers, G. (1989). *An oceanic subsurface thermal analysis scheme*. *BMRC Research Report No. 18*, Bur. Met. Australia.

Bureau of Meteorology and CSIRO. (2022). *State of the Climate 2022*.

Chiew, F., Vaze, J., Viney, N., Jordon, P., Perraud, J-M., Zang, L., Teng, J., Arancibia, J.P., Morden, R., Freebairn, A., Austin, J., Hill, P., & Wiesemfeld, C.M. R. (2008). *Rainfall-runoff modelling across the Murray-Darling Basin: A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project.*

Chua, Z.-W., Evans, A., Kuleshov, Y., Watkins, A., Choy, S., & Sun, C. (2022). Enhancing the Australian Gridded Climate Dataset rainfall analysis using satellite data. *Scientific Reports* |, *12*(20691 (2022)). https://doi.org/10.1038/s41598-022-25255-6

Chua, Z.-W., Kuleshov, Y., Watkins, A. B., Choy, S., & Sun, C. (2023). A Statistical Interpolation of Satellite Data with Rain Gauge Data over Papua New Guinea. *Journal of Hydrometeorology*, *24*(12), 2369–2387. https://doi.org/https://doi.org/10.1175/JHM-D-23-0035.1

Chubb, T., Manton, M., & Siems, S. P. A. (2016). Evaluation of the AWAP daily precipitation spatial analysis with an independent gauge network in the Snowy Mountains. *Journal of Southern Hemisphere Earth Systems Science*, *66*(1), 55–67.

Daley, R. (1993). *Atmospheric data analysis.* Cambridge University Press.

Dey, R., Lewis, S.C., Arblaster J.M, & Abram, N.J. (2019). A review of past and projected changes in Australia's rainfall. *WIRES Climate Change*, *10(3)*(e577).

Evans, A., Jones, D., Smalley, R., & Lellyett, S. (2020). An enhanced gridded rainfall analysis scheme for Australia. *Bureau Research Report-41*, Mellbourne, Australia.

Frost, A. J., & Shokri. A. (2021). *The Australian Landscape Water Balance model (AWRA-L v7).* Technical Description of the Australian Water Resources Assessment Landscape model version 7.*Bureau of Meteorology Technical Resport.*

Fu, G., Barron, O., Charles, S.P, Donn, M.J., Van Niel, T.J., & Hodgson, G. (2022). Uncertainty of Gridded Precipitation at Local and Continent Scales: A Direct Comparison of Rainfall from SILO and AWAP in Australia. *Asia-Pacific Journal of Atmospheric Sciences*, *58*, 471–488.

Glowacki, T. J., & Seaman, R. (1987). *A general purpose package for univariate two-dimensional data checking and analysis. BMRC Research Rep. 5. Bureau of Meteorology, Australia.*

Hollis, D., McCarthy, M., Kendon, M., Legg, T., & Simpson, I. (2019). HadUK-Grid—A new UK dataset of gridded climate, Observations. *Geoscience Data Journal*, 152–161.

Hong, Y., Nix, H. A., Hutchinson, M. F., & Booth, T. H. (2005). Spatial interpolation of monthly mean climate data for China. *International Journal of Climatology*, *25*(10), 1369–1379. https://doi.org/https://doi.org/10.1002/joc.1187

Hopkinson, R. F., Hutchinson, M. F., McKenney, D. W., Milewska, E. J., & Papadopol, P. (2012). Optimizing Input Data for Gridding Climate Normals for Canada. *Journal of Applied Meteorology and Climatology*, *51*(8), 1508–1518. https://doi.org/https://doi.org/10.1175/JAMC-D-12-018.1

Hutchinson, M. F. (1995). Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems*, *9*(4), 385–403. https://doi.org/10.1080/02693799508902045

Hutchinson, M. F., & Xu, T. (2015). Methodology for Generating Australia-wide Surfaces and Associated Grids for Monthly Mean Daily Maximum and Minimum Temperature , Rainfall , Pan Evaporation and Solar Radiation for the Periods. *NARCliM Report to the NSW OEH*.

Hutchinson, M.F. (1993). *On thin plate splines and kriging. In: Tarter ME and Lock MD.(eds). Computing and Science in Statistics*.

Hutchinson, M.F. (1998). Interpolation of Rainfall Data with Thin Plate Smoothing Splines - Part I: Two Dimensional Smoothing of Data with Short Range Correlation. *Journal of Geographic Information and Decision Analysis*, *2*(153–167).

Hutchinson, M.F., & Bischof, R. (1983). A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales. *Australian Meteorological Magazine*, *31*, 179–184.

Jeffrey, S.J., Carter, J.O., Moodie, K.B., & Beswick, A.R. (2001). Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Env. Model. and Software, 66*, 309–330.

Jones, D., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Aust. Meteorol. Oceanogr. J., 58*, 233–248.

Jones, D. A., & Trewin, B. (2000). The spatial structure of monthly temperature anomalies over Australia. *Australian Meteorological Magazine*, *49*(4), 261–276.

Jones, D.A., Wang, W., Fawcett, R., & Grant, I. (2006). The generation and delivery of Level-1 historical climate data sets. *Australian Water Availability Project Milestone Report.*, 32.

Jones, D.A., & Trewin, B. (2000). On the description of monthly temperature anomalies over Australia. *Australian Meteorological Magazine*, *49*, 261–276.

Jones, D.A., & Trewin, B. (2002). On the adequacy of historical Australian daily temperature data for climate monitoring. *Aust. Met. Mag., 51*, 237-50.

Jones, D.A., & Weymouth, G. (1997). An Australian monthly rainfall data set. In *Technical Report No. 70*.

King, A.D., Alexander, L.V., & Donat. M. (2013). The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia. *International Journal of Climatology*, *33*(10), 2376–2387.

Koch, S.E., DesJardins, M., & Kochin, P. (1983). An interactive Barnes objective map analysis scheme for use with satellite and conventional data. *Jnl Clim. Appl. Met., 22*, 1487-1503.

WMO (2007). *The role of climatological normals in a changing climate*.

Pepler, A.S., Dowdy, A.J, & Hope. P. (2021). The differing role of weather systems in southern Australian rainfall between 1979–1996 and 1997–2015. *Clim Dyn*, *56*, 2289–2302.

Reynolds, R.W.; & Smith, T. M. (1994). Improved global sea surface temperature analyses using optimum interpolation. *Journal of Climate*, *6*, 12.

Su, C.-H., Rennie, S., Torrance, J., Dharssi, I., Tian, S., Howard, E., Pepler, A., Stassen, C., & Steinle, P. (2023). *Preliminary assessment of regional moderate-resolution atmospheric reanalysis for Australia*.

Tait, A.B, Henderson, R., Turner, R., & Zheng. X. (2006). Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *International Journal of Climatology*, *26*(14), 2097–2114.

Thornton, P.E., Shrestha, R., Thornton, M., Kao, S.C, Wei, Y., & Wilson. B. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data*, *8*, 190.

Trewin, B., Braganza, K., Fawcett, R., Grainger, S., Jovanovic, B., Jones, D., Martin, D., Smalley, R., & Webb, V. (2020). An updated long-term homogenized daily temperature data set for Australia. *Geoscience Data Journal*, *7*(2), 149–169. https://doi.org/https://doi.org/10.1002/gdj3.95

Vogel, E., Lerat, J., Pipunic, R., Frost, A. J., Donnelly, C., Griffiths, M., Hudson, D., & Loh, S. (2021). Seasonal ensemble forecasts for soil moisture, evapotranspiration and runoff across Australia. *Journal of Hydrology*, *601*, 126620. https://doi.org/https://doi.org/10.1016/j.jhydrol.2021.126620

Wilson, L., Bende-Michl, U., Sharples, W., Vogel, E., Peter, J., Srikanthan, S., Khan, Z., Matic, V., Oke, A., Turner, M., Co Duong, V., Loh, S., Baron-Hay, S., Roussis, J., Kociuba, G., Hope, P., Dowdy, A., Donnelly, C., Argent, R., … Bellhouse, J. (2022). A national hydrological projections service for Australia. *Climate Services*, *28*, 100331. https://doi.org/https://doi.org/10.1016/j.cliser.2022.100331

Yatagai, A., Maeda, M., Khadgarai, S., Masuda, M., & Xie. P. (2020). End of the Day (EOD) Judgment for Daily Rain-Gauge Data. *Atmosphere*, *11*(8), 772.

Appendix A: Summary site and temporal statistics for correlation scale trials

Table S1: Summary spatial average error statistics of Tmax and Tmin during 1991-2020 using different correlation scales. Bold indicates the optimal selection.

	CS=0.1	CS=0.5	CS=0.8	CS=0.1	CS=0.5	CS=0.8
Variable	RMSE	RMSE	RMSE	BIAS	BIAS	BIAS
Tmax	**1.261**	1.263	1.263	**-0.002**	-0.002	-0.002
Tmin	**1.774**	1.775	1.775	**0.012**	0.015	0.015

Table S2: Temporal average error statistics of Tmax and Tmin during 1991-2020 using different correlation scales (CS). Bold indicates the optimal selection.

	CS=0.1	CS=0.5	CS=0.8	CS=0.1	CS=0.5	CS=0.8
Variable	RMSE	RMSE	RMSE	BIAS	BIAS	BIAS
Tmax	**1.245**	1.247	1.247	**-0.0023**	-0.0024	-0.0024
Tmin	**1.759**	1.760	1.760	**0.012**	0.014	0.014

Appendix B: Temporal error differences

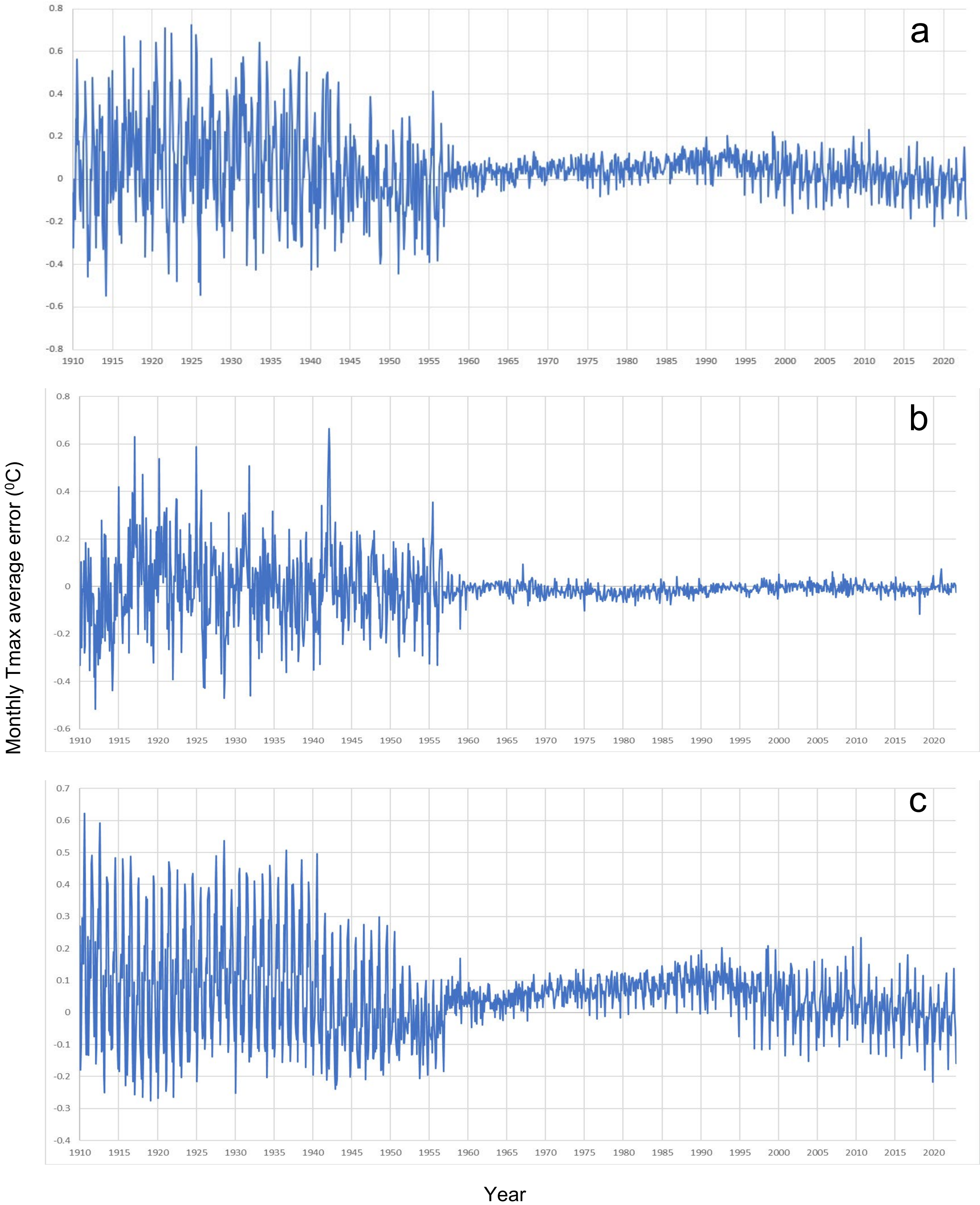

Figure S1: Difference in monthly mean maximum temperature a) AGCD (monthly average of daily) - AWAP (monthly average), b) AWAP (monthly average of daily) - AWAP (monthly) and c) AGCD (monthly average of daily) - AWAP (monthly average of daily) during 1910-2020.