Australian Government
Bureau of Meteorology

Victoria State Office
Bureau of Meteorology
GPO Box 1289, Melbourne VIC 3001 Australia

In reply please quote

13MD341921_2022

Mr Geoffrey H. Sherrington
1/1137 Doncaster Rd
DONVALE VIC 3111
T: 03 9842 1824
M: 0466 561 783
E: sherro01@outlook.com

31 March 2022

Dear Mr Sherrington

Thank you for your correspondence dated 17 January 2022, and continued interest in the Bureau of Meteorology's (the Bureau's) climate data and statistical techniques. I apologise for my delay in responding. I will endeavour to respond to your queries.

The full description and quantification of the uncertainties associated with temperature measurement technologies are included in *Instrument Test Report 716 – Near Surface Air Temperature Measurement Uncertainty V1.4* (2022) which is attached. Measurement uncertainties are calculated and described in three categories, depending upon the volume of data available:

1. Isolated measurements

   These are measurements taken with little supporting evidence or experience of the performance of the observation system. Uncertainty in these cases is heavily affected by the inspection method tolerance, which reduces as the history of the site increases. Similarly, these measurements reflect a lack of supporting information from repeat inspections, correlation with multiple sensors/AWS, and other key data.

   As a result, this categorisation is typically relevant for new stations with less than a year's operation, or situations where there is limited data available from nearby locations to verify the quality of observations.

2. Typical measurements

   Typical measurements form the majority of those taken across the automatic weather station network. These refer to data from stations where there is supporting evidence or a history of reliable operation from the site. Typically, this will include 5-10 inspections of the site, alongside available data from neighbouring stations with similar climatology to confirm the consistency of performance.

3. Long-term measurements

   This applies to aggregated data sets across multiple stations and over extended periods. This aggregation mitigates random errors and is suitable for use in determining changes in trends over extended periods. These measurements typically are constructed from aggregating data for stations with supporting evidence or experience from neighbouring stations and other supporting data sources, all with several years of operation.

The uncertainties, with a 95% confidence interval for each measurement technology and data usage, are listed below. Sources that have been considered in contributing to this uncertainty include, but are not limited to, field and inspection instruments, calibration traceability, measurement electronics or observer error, comparison methods, screen size and aging.

Measurement Technology	**Ordinary Dry Bulb Thermometer**	**PRT Probe and Electronics**
Isolated single measurement – No nearby station or supporting evidence	±0.45 °C	±0.51 °C
Typical measurement – Station with 5+ years of operation with 10+ years of operation with at least 5 verification checks.	 ±0.23 °C ±0.18 °C	 ±0.23 °C ±0.16 °C
Long-term measurement – Station with 30+ years of aggregated records with 100+ years of aggregated record	 ±0.14 °C ±0.13 °C	 ±0.11 °C ±0.09 °C

I would stress that in answer to your specific question of "*If a person seeks to know the separation of two daily temperatures in degrees C that allows a confident claim that the two temperatures are different statistically by how much would the two values be separated*", the 'Typical measurement' Uncertainty for the appropriate measurement technology would be the most suitable value. This value is not appropriate for wider application to assess long-term climate trends, given typical measurements are more prone to measurement, random, and calibration error than verified long-term datasets.

The impacts of these uncertainties on the long-term climate record are described in the peer-reviewed publication *Estimating the uncertainty of Australian area-average temperature anomalies* (Grainger et al., 2021), available at https://doi.org/10.1002/joc.7392.

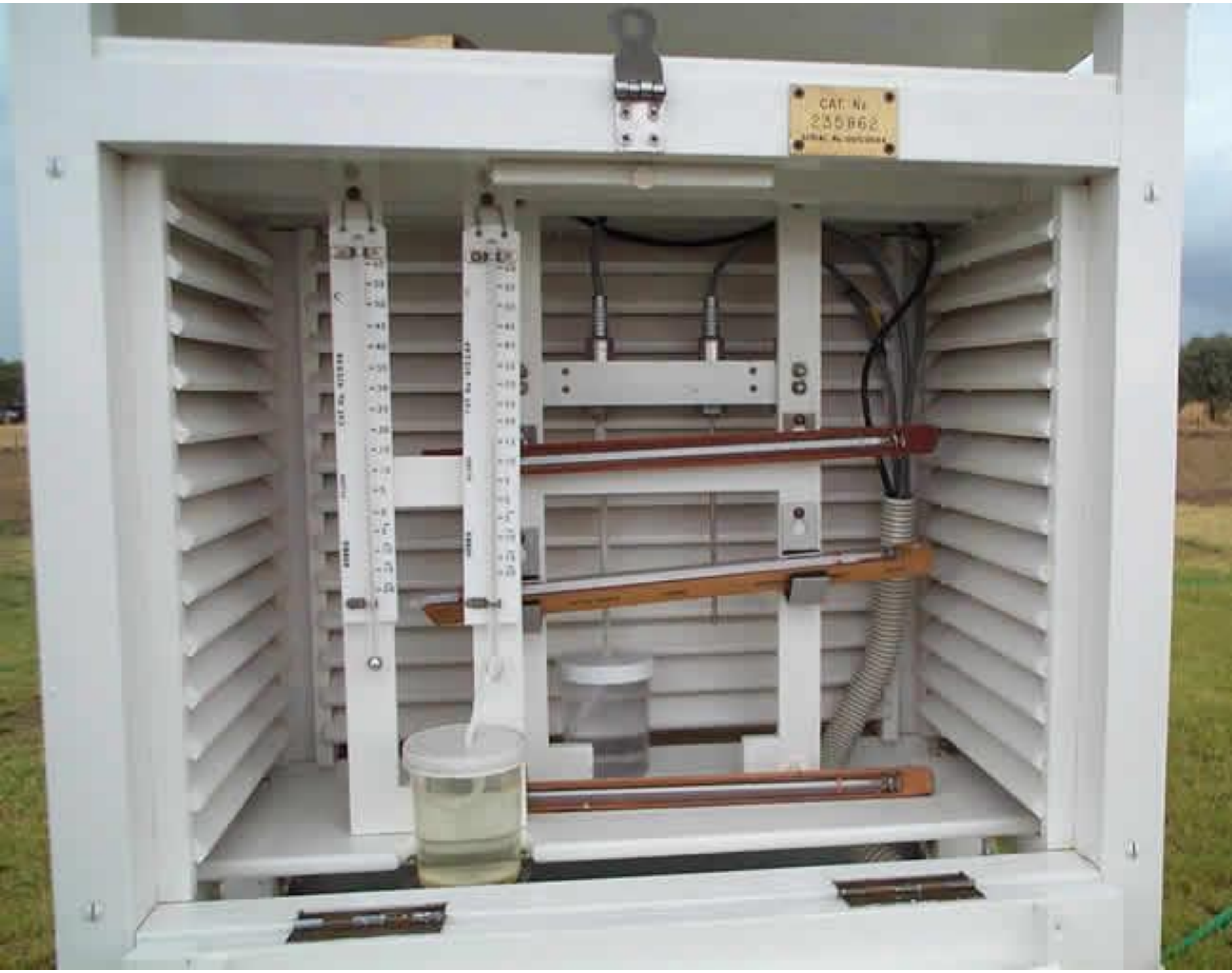
I trust that this answers your query.

Kind regards

Dr Boris Kelly-Gerreyn
Gen Manager, Data Program & Chief Data Officer
Data & Digital Group
Bureau of Meteorology
700 Collins Street, Docklands, VIC 3008

Instrument Test Report 716

Near Surface Air Temperature Measurement Uncertainty V1.4_E



Bureau of Meteorology
GPO Box 1289 Melbourne Vic 3001

Table of Contents

Aim

To provide an estimate of the uncertainty of the near surface[1] temperature measurements achievable by the Bureau's surface observation instruments, installed and operated at both staffed stations (OBS) and Automatic Weather Stations (AWS).

Scope

The uncertainty estimates in this report apply to routine manual 3-hourly near surface temperature measurements from staffed stations and automated 1-minute near surface temperature measurements at AWS. It precludes maximum and minimum temperatures at staffed stations as these are taken using minimum and maximum thermometers.

Background

Temperature, what is it?

Temperature is one of the fundamental SI (International System of Units) measurable quantities, along with length, mass, time, current, substance or mole and light intensity. From these seven quantities, all other measurable quantities can be derived [Bentley 1999].

At its core, temperature is a measure of the kinetic energy (energy of movement) present in the atoms and molecules of the substance being measured[2]. There are multiple sources of energy and processes involved in the generation and measurement of temperature. These include thermal energy in the form of radiation (sunlight), convection (wind, which can both increase or remove energy from the system), conduction through contact with nearby materials and re-irradiation of stored energy from nearby objects.

In an ideal world the atmospheric temperature could be perfectly represented as a mathematical model. All contributions to a measurement would be identified or controlled and quantified. While in a laboratory environment this may be possible, in the real world there is much that cannot be controlled. As a result, the method of measurement becomes a part of the definition of the measurement itself.

Near Surface Air Temperature and Meteorology

Historically, meteorologists were interested in the large-scale average kinetic energy of the air molecules, independent of radiation and convective sources. This is commonly known as "shade temperature".

Early in the history of the meteorology of atmospheric temperature measurement, scientists recognised that measuring temperature in the natural environment was difficult. Small changes in cloud cover (radiation), gusty wind (convection) and rain (conductive evaporation), all cause significant variability in the measurement.

In the late 1700s and early 1800s a process for developing a standard (and consistent)

---

[1] Throughout the document surface temperature is considered the equivalent of near surface temperature and refers to the air temperature at typical height of 1.2m above the ground.

[2] Temperature is related to energy through the specific heat capacity and amount of a material.

$$\Delta T = \frac{\Delta Q}{m\, c_p(p,T)}$$

Where ΔQ is the change in energy of the material, m is the mass of the material, $c_p(p,T)$ is the specific heat capacity at a particular pressure and temperature ΔT is the change in temperature resulting from the change in energy.

means of measuring atmospheric temperature was being developed. Early mounts for thermometers required the observer to turn the stand so the thermometer was not in direct sunlight. Later the "stands" were incorporated into huts and screens.

> In 1873 the Royal Meteorological Society (UK) agreed on the following criteria [Parker 1994]: -
>
> 1. *that the stand be protected from direct rays of the sun,*
> 2. *that the temperature of the stand should not affect the thermometers,*
> 3. *that there should be no reflected heat from the ground or other objects,*
> 4. *that no radiation to the sky should be allowed,*
> 5. *that the stand must be independent of all other objects,*
> 6. *that there should be free circulation of air,*
> 7. *that the thermometers should not be touched by rain or snow,*
> 8. *that there should be no need for attention between readings,*
> 9. *that ample room for duplicate instruments was preferred,*
> 10. *that the screen should not be costly, and*
> 11. *that the stand should be easily moved*
>
> Additionally, they also stated that it be made of yellow pine, painted white; that the box be 4 feet above the ground and set in the middle of a turfed area 15-feet square. [Parker 1994]

This is the description of the Stevenson Screen, and its positioning has persisted, with minor modifications, to this day. It aims to minimise radiation sources, control airflow, avoid contamination of readings by rain or snow and encourages good exposure. The concepts behind the description can be considered the basis of the measurand, atmospheric temperature, even if methods have evolved over time.

Temperature Uncertainty

The temperature uncertainties estimated in this report are evaluated according to the methods described in the International Guide to the Expression of Uncertainty in Measurement (GUM) [JCGM 100:2008, ISO 2009]. This guide provides methods to combine the known sources of error associated with the measurement to generate an uncertainty evaluation. This is driven by the concept of a well understood "measurand" or the information needed from the measurement. The GUM aims to validate the measurement process and identify sources of error and process improvements. It is not limited to just the prominent elements such as calibration but encourages an analysis of the way the measurement is taken. It encourages the identification of additional sources of uncertainty, which, in the case of atmospheric temperature, can include instrument radiation screens, timing effects, maintenance and quantities of influence etc. It also states that where biases (errors) are known, these should be eliminated rather than being included in the estimated uncertainty.

The purpose of the uncertainties determined by this method is to report, with a high confidence level, that the “true value of the measured temperature” lies within these bounds.

An uncertainty evaluation provides an indication of the level of reliability of these measurements. However, it is crucial to understand how the estimations are determined and what considerations are included in the analysis, as this strongly influences the utility of measurements in observation, forecasting and weather products.

The following section (Measurand) describes three different uses of a temperature measurement and what is considered within the estimations of uncertainty.

Measurand

Within the context of metrology, the description of what is being measured and the influences on that measurement are referred to as the measurand, and the uncertainty evaluation is uniquely associated with that measurand.

> For example, a metre-long steel ruler that has been calibrated at 20 °C ± 2 °C could have an uncertainty of ±0.01mm. If the ruler is now used outdoors on a day when the temperature is 35 °C, the temperature induces an expansion of the ruler, changing the scale and introduces a new and unaccounted for error.

Features commonly considered when defining a measurand are:

- the conditions under which the instrument was calibrated,
- the type of instrument and how it is used, and
- how the information is gathered.

These are all important when it comes to understanding uncertainty estimates. As a corollary, the uncertainty is also influenced by the usage of the measurement. The simplest example of this is the situation where repeated measurements of the same value are taken. This can, in some circumstances, allow the user to reduce the uncertainty.

It is also important to note, that as understanding of the measurement process improves, so does the estimate of the uncertainty for related measurands.

Field Measurement Instruments

The instruments considered in this report are those that provide the measurement of temperature for surface observations sites (both manual and automatic). These measurements and their associated instruments are listed in Table 1.

Table 1. Field Near Surface Temperature Instruments

Instrument Name	Measurement Description	Measurement Method
Ordinary Dry Bulb Thermometer	Three-hourly instantaneous (SYNOP[3]), or hourly instantaneous temperatures (METAR[4]), Stevenson screen temperature	Human read mercury-in-glass thermometer
Air temperature probe	One-minute average Stevenson screen temperature	Automated electronically read platinum resistance temperature probe (PRT)

Near Surface Air Temperature Measurand Definitions

The uncertainties outlined in this report cover the performance of the:

[3] SYNOP refers to the alphanumeric code (FM12) used by members of the World Meteorological Organization for the communication of surface observations from fixed land stations. (WMO 2019)

[4] METAR refers to the alphanumeric code (FM15) used by members of the World Meteorological Organization for the communication of Aerodrome routine meteorological observations. (WMO 2019)

- Standard Bureau ordinary mercury-in-glass thermometer or electronic platinum resistance temperature (PRT) probe[5], with a:
- Large[6] or Small[7] Stevenson screen,
- maintained at a consistent site[8] and
- operated according to existing documentation, including standard maintenance of the screen and instruments.

The uncertainty estimates of temperature measurands are computed for three different user scenarios.

- *Isolated Measurement*,
  This is the measurement uncertainty (95% confidence interval) of a single surface air temperature measurement (i.e. 1-minute average or single manual observation) from a newly established system under the conditions mentioned above, and/or with no correlated or related measurements from other nearby sites.

  Usage examples – examining data from a new system or site, when comparing Bureau data with measurements from an unknown source or in the initial stages of a new instrument or system design.

- *Typical Measurement*,
  This is the measurement uncertainty of a single surface air temperature measurement (i.e. 1-minute average or single manual observation) taken from an established network of temperature instruments within the Bureau system. Within this context, an established system has operated for more than a year and has supporting information including records of maintenance and verifications in line with Bureau guidelines.
  Such a measurement is considered to be part of a sequence of measurements in time or space that are assumed to have related parameters.

  Usage examples - determination of maximum or minimum temperature at automated weather stations, comparisons between one or more stations, data assimilation in numerical weather prediction models.

- *Long Term Measurement*
  This is the uncertainty related to long term surface air temperature measurement. Within this context, a long-term measurand has an operational history of more than

---

[5] The electronic platinum resistance air temperature probe is commonly abbreviated to PRT Probe. Wherever the terms, electronic, air temperature or platinum resistance are used they refer to the PRT probe

[6] Large Stevenson Screen (BoM 1990), Construction: Red cedar (Aust.), Western red cedar (Canadian), Radiata Pine, Californian Red Wood,
Dimensions: Internal - depth 537 mm, width 710 mm height 660 mm;
External - depth 637 mm, width 810 mm, height 797 mm (back) and 853mm;
Roof depth 760 mm, width 915 mm height 20 mm;
Louver – 20 sides and 16 door, external 83 mm by 12 mm, internal 47 mm by 12 mm.

[7] Small Stevenson Screen Construction (BoM 2017): Red cedar (Aust.)
Dimensions: Internal - depth 270 mm, width 521 mm height 395 mm;
External - depth 346 mm, width 597 mm, height 579 mm (back) and 555 mm;
Roof depth 425 mm, width 627 mm height 20 mm;
Louver – 13 sides and 11 door, external 55 mm by 8 mm, internal 32 mm by 8 mm.

[8] Complies to WMO Siting Classification 1 or 2 [WMO 2018].

five years with related metadata, maintenance, and verification in line with Bureau guidelines and nearby stations that the record can be validated against. Such a measurement is considered to be part of a long sequence of measurements in time or space that are assumed to have related parameters

Usage examples - monitoring of climate, aggregation of multiple individual measurements such as monthly, annual or multi-year means, identification of external impacts e.g. identification of artefacts caused by changes in the local environment.

Method

To determine the uncertainty of each measurand, the steps in the measurement chain are examined and the sources of error or uncertainty quantified. This verification process is critical to the determination of a valid estimate of uncertainty.

Sources of Uncertainty

Traceability and Calibration

To ensure the quality of temperature measurements, the Bureau of Meteorology provides a chain of traceability from field measurement to national and international standards.

Field instruments are either calibrated by the Bureau's Standards and Metrology Laboratory (SML) or purchased with calibration certificates. If calibrated by the Bureau, the traceability is through a reference 25 Ohm Standard Platinum Resistance Thermometer (SPRT) to define fixed points[9] of the International Temperature Scale (ITS-90).

Fixed Point references use the physical characteristics of materials such as their melting, freezing and triple points to generate and sustain known temperatures for calibration. The triple point of a substance is the stable temperature at which all three phases, (gas, liquid and solid) co-exist and are used in the definition of the temperature scale (ITS-90) [Preston-Thomas 1990]. Once every 3 to 5 years, the SPRT is calibrated by the National Measurement Institute of Australia as a verification of the internal calibration processes. Additionally, the laboratory takes part in Proficiency Testing Schemes in line with the National Association of Testing Authorities [NATA 2018] accreditation criteria with both domestic and international laboratories.

Field inspection instruments are calibrated by the SML and returned for recalibration every 12 months.

Measurement Interface

Measurement interface is either a person (human-read), in the case of ordinary dry bulb thermometers or a data logger, in the case of PRT probes.

To determine the contribution to the uncertainty from a human-read measurement interface, an experiment was undertaken in the laboratory which involved presenting eight staff who undertake inspections in the field with a range of thermometers immersed in water at different temperatures. These included ice points with pure water and salt water, room temperature and water slightly elevated above room temperature. Each person read the temperature of each thermometer and repeated the sequence of reading three times. From these data, estimates of measurement uncertainty by humans reading a thermometer were determined. No estimate of uncertainty was included for transcription errors. Typically, these

---

[9] The Bureau use triple point cells of Mercury, Water, Gallium, and Indium which cover the temperature range -38.8344 °C to 156.5985 °C

errors can be identified during quality control and the values flagged accordingly in the record.

For the electronic interface, calibrations were undertaken across the range of expected field temperatures -30 °C to 60 °C. The electronics were installed in a climate chamber (Weiss Technik C340/40) and tested at four specific temperatures -30, 0, 30 and 60 °C. At each temperature, a reference source (Fluke 5522A, S/N 4255904) was used to generate resistances equivalent to specific temperatures -30, 0 20 and 60 °C. The data for eight ALMOS MSI2 Sensor Cards, that have been tested four times each, was assessed to determine the uncertainty contributors from the electronics.

Field Verification and Replacement Process

Once installed in the field, the traceability of field instruments is maintained via verification tests performed each year against traceable references. Traceable field references are calibrated by the Bureau's Standards and Metrology Laboratory, listed as a World Meteorological Organisation Regional Instrument Centre[10] (RIC) in Australia. The field references are issued to regional staff who undertake field verification checks of the instruments every six months for climate sites, and twelve months for standard sites.

The verification tests during a station visit consists of an ‘on arrival' and ‘on departure' test. These tests are a set of comparisons or "inspections" against a reference sensor to confirm the performance of the field sensor. The reference sensor is mounted so that it touches the tip of the field sensor and is left to equilibrate for 30 mins. Three sets of comparison measurements are taken. The results are compared to the required field tolerance (see Table 2). If the difference between the reference and field sensor exceeds the allowed tolerance as defined by the test uncertainty ratio (TUR) [Bennett 2005], the instrument is replaced.

The use of test tolerance and associated Test Uncertainty Ratios is common within the calibration industry. The aim is to manage the likelihood of an instrument being outside its design specification. The choice of the value of the TUR is dependent on several factors but most typically driven by the risk of a false positive, where the sensor being tested may appear to pass when it is in fact faulty.

Most industries use a TUR of between 2 and 4 [Bennett 2005]. For the Bureau's electronic sensors, the design tolerance is 0.1 °C and for thermometers 0.2 °C. The design tolerance is multiplied by the TUR to give the Test Tolerance in Table 2.

The Bureau applies TURs in a manner that manages both false positive and false negatives. If the result is 0.3 °C or less, the TUR method tells us that the probability the unit under test is operating within specification is effectively 100%. As the difference increases toward the test tolerance the risk of a false positive increase, similarly the risk of replacing a good sensor also increase, false negative.

To manage these risks of false positives and false negatives, field staff track previous verifications stored in the Bureau's meta-database. They are recommended to change the sensor if sequential tests approach or exceed the test tolerances (Table 2). They are also trained to determine if the conditions of test, such as variable environmental conditions or instability have influenced the measurement. These modifications to the TUR process limit the impact of increased uncertainty through false negative results, replacing sensors that are not faulty and keeping sensors that appear good, but are actually faulty, false positives.

As an example, if the previous comparisons demonstrate a slow drift in the sensor, then the technical staff may replace the sensor even if it is passes according to the TUR, or alternately retain a sensor in the field that exhibits a bare failure, if there is no evidence of

[10] See https://community.wmo.int/activity-areas/imop/Regional_Instrument_Centres

drift in previous tests. Inspection tests that show a clear exceedance of the test tolerance are replaced.

Table 2. Test Tolerance and Test Uncertainty Ratio (TUR) [Bennett 2005] of field verifications [BoM 2010, BoM1989, BoM 1991]

Sensor	Comparison Method	TUR	Test Tolerance (°C)
Ordinary Dry Bulb Thermometer	Electronic Transfer Standard	4	0.4
	Certified Thermometer	2	0.4
Air Temperature Probe (Slim PRT probe)	Electronic Transfer Standard	4	0.4
	Certified Thermometer	2	0.4
Air temperature probe (Old PRT probe)	Electronic Transfer Standard	5	0.5
	Certified Thermometer	2	0.5

To determine the actual impact of this traceability and validation chain on the measurement's quality, a range of sources of calibration and verification data were assessed. These included: -

- calibration certificates and results for sensors and reference instruments,
- calibration results of AWS sensor boards,
- manufacturers' specifications, and
- statistical analysis of field verification data [Warne 2016], to determine,
  - sensor drift,
  - verification of uncertainty characteristics as a function of site,
  - verification of uncertainty characteristics as a function of sensor,
  - verification of uncertainty characteristics as a function of inspector
  - analysis of both before and after checks to determine the effectiveness of the TUR approach to verification,
  - frequency of verifications by individual sensor and site

Instrument Design

The design of the instrument contributes significantly to the associated uncertainty of the temperature measurement. For example, the size and shape of the mercury bulb impacts the response time of the thermometer, as do the diameter and packing material in an electronic PRT probe. By purchasing to a specification, the variability in manufacture and practise are managed.

Ordinary Dry Bulb Thermometers: - The purchase and measurement specifications for ordinary mercury-in-glass thermometers are covered by a Bureau specification A410 (issue 6) “Meteorology thermometers Ordinary - Maximum and Minimum” [BoM 1982], and the Australian Standard AS2819-1985[11] [AS 1985]. The measurement tolerances from the Australian Standard for the thermometer are provided in Appendix A. The other relevant performance specifications for ordinary thermometers are provided in Appendix B.

Platinum Resistance Temperature Probes: - The purchase and measurement specifications for thermometers are covered by a Bureau specification A3001 [BoM 2011] and it refers to the withdrawn Australian Standard AS2819 [AS 1985]. The measurement tolerances from the Australian Standard for the thermometer are provided in Appendix A. The other relevant performance specifications for PRT probes are provided in Appendix C.

Algorithm and Message Reporting

In the case of temperature measurements from electronic sensors, the measurement needs to be converted from a raw resistance to a value on the temperature scale. This is achieved through a mathematical algorithm that transforms the resistance to temperature on the International Temperature Scale 1990 (ITS90). This conversion is an approximation using Callendar-Van Dusen [BIPM 1997].

The resolution of the measurement also affects the uncertainty if it is larger compared to the overall uncertainty. For example, if the resolution of the measurement is 1 °C and all other contributions are 0.3 °C then the resolution will be a significant source of uncertainty.

Screen Type and Maintenance

The effect of the screen design and construction on the uncertainty of the temperature measurement. For example, small "beehive" screens where the sensor is positioned close (<100 mm) to the screen walls tend to display strong local warming and cooling due to irradiation from the screen surfaces. The larger screens, where the sensor was greater than 100mm from the north or west walls, experienced significantly less warming [Warne 1998]. Similarly, discolouration due to aging and dirt build up also causes warming. Regular maintenance reduces this impact.

Screen Siting

Screen siting is a separate consideration - topography and the surrounding environment can add biases to the temperature. In the methodology adopted in this study, the assumption has been made that the site is compliant with WMO Siting Classification 1 or 2[12] [13][WMO 2018] and [BoM 1997]. This implies the site is well maintained and local disturbances are minimised.

---

[11] The standard has since been labelled "withdrawn" from sale, although it is still accessible on the Standards Australia website at the time of publishing this report. In practice, this means that the standard is not being updated or amended nor, in this case, is there a replacement standard(s). The standard was first published in 1966 as ASR13 and renamed AS2819 in 1985. AS2819 is the standard against which the thermometers used in the Bureau network have been purchased and tested since 1966.

[12] If the site is not WMO Siting Classification 1 or 2 compliant this may result in systematic differences with other sites, but not necessarily inconsistencies within the sites own historical record. If significant changes have occurred at the site, such as construction of a large building near the site, then there may be systematic biases introduced into the record. In either situation, these will need to be investigated on a case-by-case basis. The analysis provided in this document will aid in identification and quantification of any impact.

[13] Recent studies in the open literature show that the impact of some site changes may be less significant than previously understood. See [Kinoshita 2014, Clark 2016 and Coppa 2019]

Results

Sources of Uncertainty Information

The process of identifying sources of uncertainty for near surface atmospheric temperature measurements was carried out in accordance with the International Vocabulary of Metrology [JCGM 200:2008]. This analysis of the measurement process established seven root causes and numerous contributing sources. These are described in Table 3 below. These sources of uncertainty correlate with categories used in the uncertainty budget provided in Appendix D.

Table 3 - Ordinary Dry Bulb Thermometer and Air Temperature PRT Probe Uncertainty Contributors. Definitions in accordance with the International Vocabulary of Metrology [JCGM 200:2008]

Uncertainty Parameters	*Mercury in Glass Thermometer (Ordinary Dry Bulb)*	*PRT Probe (Air Temperature Probe)*
Instrument Performance		
Calibration uncertainty	The uncertainty contribution from the reference instrument and system was taken from the laboratories ISO/IEC 17025 accreditation documentation. This was calculated from calibration data of instruments deployed to the field after calibration.	
Residual correction	The uncertainty contribution of the ordinary dry bulb thermometer and PRT probe determined via the analysis of past calibrations.	
Scatter	The uncertainty contribution of the ordinary thermometer or PRT probe determined via the analysis of past calibrations.	
Secular Change or Zero Drift	Zero drift is the irreversible elastic error resulting from the ordinary dry bulb thermometer mercury reservoir (bulb) changing shape over time to relieve internal stresses in the glass. The rate of change commonly reduces over time and has been estimated as 0.01 °C /year. [Bentley 1999]	Zero drift is aging of the sensor's materials over time resulting in small increases in resistance over time. The estimate of drift was determined from field performance data over a 30-year period. [Warne 2016]
Measurement Interface		
Residual Corrections	Scale marking and pointing errors of the ordinary thermometer [Bentley 1999]	The error introduced by the electronics card. [BoM 2021, BoM 1991]
Algorithm		The error introduced by the algorithm conversion of resistance to temperature. [BoM 2021]
Reading Subdivision	This is the expected ability of an observer to subdivide the minimum scale division of 0.5°C (by eye) to the required reading resolution. [Bentley 1999]	

Parallax	The error introduced when reading the indicator fluid against the engraved scale as the scale is offset from the indicator fluid by the thickness of the ordinary thermometer wall.	
Reproducibility	The error introduced by different people making the same reading. This was determined via experimentation.	The error introduced by different electronics cards making the same reading. [BoM 2021, BoM 1991]
Repeatability	The error introduced by the same person making the same reading multiple times. This was determined via experimentation.	The error introduced by the same electronics card making the same reading multiple times. [BoM 2021, BoM 1991]
Inspection Reference		
Residual Corrections	The error and scatter of the reference thermometer[14] determined during calibration. Determined from laboratory calibration. It also includes the residual bias of the inspection process.	
Resolution	The error resulting from the resolution of the reference thermometer.	
Calibration Uncertainty	The operational uncertainty of the inspector's calibrated reference thermometer, during the annual inspection. [Dollery 2008]	
Inspection Method		
Inspection Tolerance	The uncertainty influences of the verification tolerance for the annual check of the ordinary dry bulb thermometer or sensor against the inspector’s digital reference. [Warne 2016]	
Message Reporting Resolution		
Resolution	The uncertainty contribution resulting from the resolution of the reported value.	
Screen Type and Maintenance		
Type	The error resulting from differences in size of the screen, only applicable for trend estimates [Warne 1998]	
Screen Cleanliness	The error resulting from deterioration of the condition of the screen [Warne 1998]	
Screen Siting		
Site Impact	Based on a WMO Site Classification of 1 or 2, therefore no added uncertainty [WMO 2018]	

[14] The reference thermometer in the early 1990's was a high-quality mercury-in-glass thermometer but was replaced with a high-quality platinum resistance thermometer in the mid 2000's.

Uncertainty Estimates

The overall uncertainty of the mercury in glass ordinary dry bulb thermometer and PRT probe to measure atmospheric temperature is given in Table 4. This table is a summary of the full measurement uncertainty budget given in Appendix D.

Table 4 – Summary table of uncertainties and degrees of freedom (DoF) [JCGM 100:2008] for ordinary dry bulb thermometer and electronic air temperature probes also referred to as PRT probes.

	Ordinary Dry Bulb Thermometer		**Air Temperature Probe**	
	U_{95} (°C)	**DoF**	**U_{95} (°C)**	**DoF**
Isolated Measurement	0.45	40	0.51	6
Typical Measurement	0.18	86	0.16	12
Long-term Measurement Statistic	0.13	67	0.09	119

A detailed assessment of the estimate of least uncertainty for the ordinary dry bulb thermometer and air temperature probes is provided in Appendix D. This details the uncertainty contributors mentioned above in Table 3.

Discussion

Uncertainty Calculation

The detailed assessments of the combined expanded uncertainty provided in Appendix D have been undertaken in line with the ISO Guide to the Expression of Uncertainty in Measurement (GUM) [JCGM100:2008, ISO 2009]. This involved evaluation of the measurement method, the determination of the sources of uncertainty and derivation of the standard uncertainty components. These uncertainty components may be derived either from statistical analysis of experimentation results, Type A contributors; or, for example, from professional experience and specifications, which are classified as Type B contributors.

This expanded uncertainty value is converted to a standard uncertainty (typically equivalent to the standard deviation) by determining the coverage factor [15]. If the data is derived from a statistical analysis of experimental data, the source is a Type A contributor and has coverage factor determined from the analysis, but typically has the value of two. If the data has an unknown distribution because it is determined from professional experience or specification sheets, the contributors are classified as a Type B error source.

In this study, where the value is determined from prior experimentation or known to be normally distributed, then a coverage factor of 2 is used. Otherwise, these errors have been treated as square or triangular distributions and the expanded uncertainty is divided by the coverage factor 1.732 (or square root 3) or 2.449 (or square root 6) in line with the GUM. This results in a conservative estimate of the sources contribution to the overall uncertainty.

The combined Expanded Uncertainty is then calculated by combining the individual contributors and multiplying it by the coverage factor k which is determined using the overall degrees of freedom.

Sensitivity coefficient and Independent or Correlated Factor

Not all sources of error contribute equally to the final combined expanded uncertainty. To accommodate this, the GUM allows for a sensitivity coefficient. For example, the uncertainty components that have been previously discussed are assumed to be relevant to a single observation or experiment. However, if the contributing component is random, and multiple measurements or instances occur, then the GUM allows for this by dividing the sensitivity coefficient by the square root of the number of times the measurement is made or instances. For example, when multiple inspections have been undertaken at a single location, or multiple locations are used to generate an average temperature.

For clarity, in this study, the number of measurements or instances in the column labelled "Independent/Correlated Factor" have been separated out (See Tables in Appendix D). This allows for clear distinction between bias (correlated[16]) which cannot be removed by multiple measurements, and random uncertainties (independent[17]) which can.

---

[15] The coverage factor is a "number larger than one by which a combined standard measurement uncertainty is multiplied to obtain an expanded measurement uncertainty." [JCGM 100:2008] The combined standard measurement uncertainty is the combination of standard deviations, of the individual contributor's measurement uncertainty. Typically, in this report the coverage factor is approximately k=2 and typically represents 95% of the data uncertainty.

[16] "Correlated" in this context relates to sources of error that are common to most or all measurements across the network. For example, the bias of the calibration system which will impact every sensor calibration across the network.

[17] "Independent" in this context relates to sources of error that are randomised. For example, the uncertainty contribution due to reading a thermometer. Each site will have multiple inspectors checking the temperature overtime therefore the uncertainty due to these reading errors will be randomised.

Inspection Process and Uncertainty

These parameters were determined from the analysis of tens of thousands of field comparisons of electronic and dry bulb thermometers over a 30-year period. The analysis of inspections of electronic sensors is shown in Figure 1(a) and dry bulb thermometers in (b). In general, the distribution of all these errors was more triangular than normal. This type of curve is reminiscent of a Voigt curve which is a weighted combination of a Gaussian and Lorentzian (Cauchy-Lorentz) distribution. (Figure 1).

The idea that the data is well represented by the combination of two distributions was expected as there are two distinct processes influencing the generation of the data, the routine *Measurement Process* and *Inspection Process*.

The *Measurement Process* represents the expected measurement system performance uncertainties (i.e. sensor, AWS and calibration). There is good evidence to assume the sample averages will to be normally distributed and therefore a Gaussian distribution was applied.

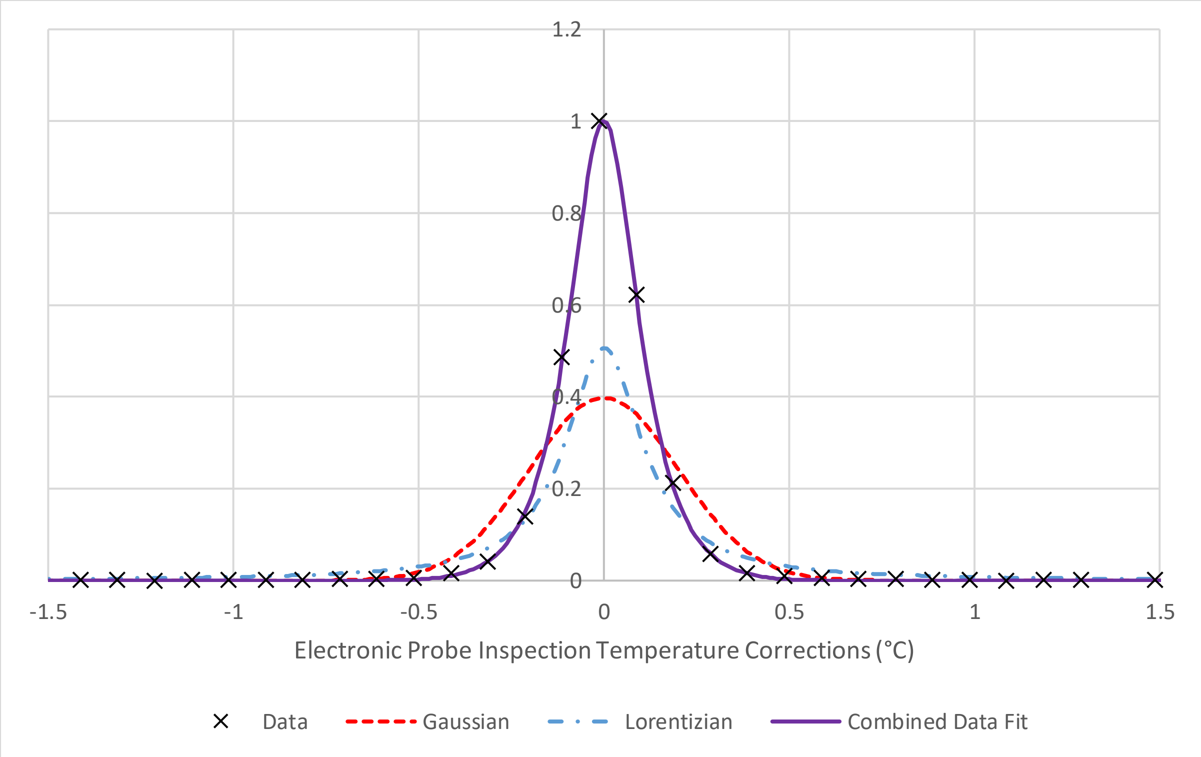
The *Inspection Process* aims to verify if the sensor and electronics are performing within specification. This involves the comparison of the field instrument with a transfer reference. The Inspector makes a judgement on whether the sensor is in good condition and reporting reliably based on this comparison and the physical condition of the sensor. If the inspection differences is less than or equal to 0.3°C, there is no statistically detectable change in the instrument and it will be left in place. A difference of greater than the tolerance of either 0.4 or 0.5°C (see Table 2) implies a high likelihood the sensor is faulty, and it is replaced. In the case of an observed difference of between 0.3°C and the tolerance, the inspector has the authority and training to decide if the instrument needs replacement. This discretion is to allow for false negatives due to the observing conditions.

As a result of the inspection thresholds, their contribution to the overall distribution of the inspection data was not expected to be normally distributed. The overall distribution displays the characteristics of a Voigt distribution; therefore, a Lorentzian distribution was chosen to represent the Inspection Process. This type of distribution tends to be sharper in the centre and with more data in the extended wings than a normal distribution.

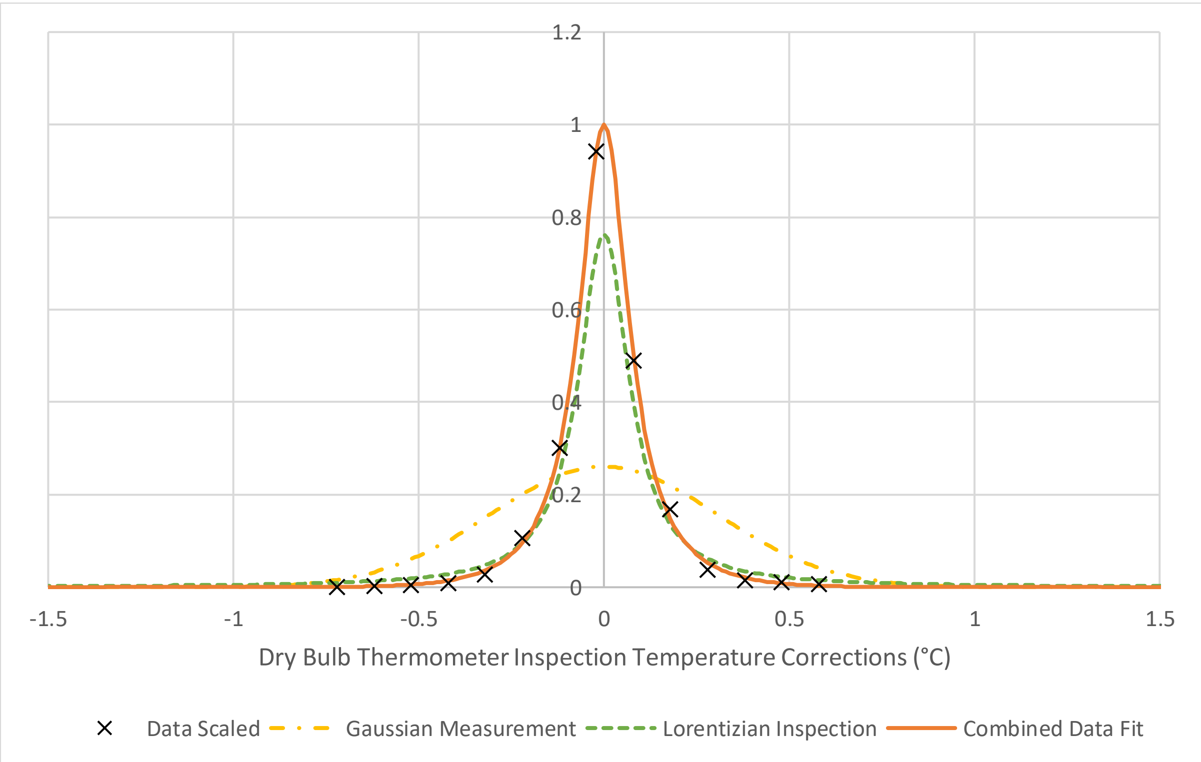
It contributes to the overall uncertainty only in the case that a faulty sensor is not removed. The two distributions have been combined in accordance with Equation 1 below using an estimate of equal probability (i.e. 50%). It is recognised that this is an overestimate of the contribution from the inspection process to the uncertainty.

Eq (1) $$f_{Obs}(x) = p . f_M(x) + (1-p) . f_I(x)$$

where	$f_{Obs}(x)$	*is the observed distribution*
	$f_M(x)$	*is the measured process distribution*
	$f_I(x)$	*is the inspection process*
	p	*is the probability fraction, 0.5*



(a) Electronic Probe Inspections



(b) Dry Bulb Inspection

Figure 1. Plots of the observed comparison data (X) between the inspection instrument and dry bulb field for 13128 inspections between June 1973 and June 2013. This was generated using a bin size of 0.1 °C with bins centred on 0.07 °C and 0.063 °C respectively. The red dashed line is a Gaussian distribution representative of the *Measurement Process*, and blue dashed/dotted line is a Lorentzian distribution representative of the *Inspection Process*. The Purple is the distribution for the combined measurement and Inspection processes.

An iterative approach was used to model the raw data and determine the 68th percentile of the Gaussian and the Lorentzian[18] distribution. Any bias detected in the fit was attributed to the inspection instrument and included in the uncertainty calculation separately. Table 5 provides a summary of the resultant model parameters. The overall fit accounts for 99.98% of the distribution for temperature probes and 99.79% for thermometers.

Table 5 – Results of modelling of temperature inspection data to determine the contribution to uncertainty from the measurement and inspection processes.

	Fit Parameters (°C)	Offset (°C)	Weighting
Ordinary Dry Bulb Thermometer			
Measurement Model	± 0.305°C	-	50%
Process Model	± 0.166°C	0.070°C	50%
PRT Probe			
Measurement Model	± 0.202°C	-	50%
Process Model	± 0.253°C	0.063°C	50%

The *Measurement Model* was identified as having a standard deviation of 0.20 °C for the air temperature probe and 0.31 °C for the dry bulb thermometer which is commensurate with the standard uncertainty of the measurement components of the uncertainty (i.e., 0.2 and 0.38 °C respectively for temperature probes and thermometers) from the component analysis given in Appendix D, when the inspection method component is removed.

The *Process Model* contribution was identified as having a full width half maximum of deviation of 0.25 °C for temperature probes and 0.17 °C for thermometers. This was converted to an expanded uncertainty of 0.51 °C and 0.33 °C which is consistent with the inspection process limits.

Proportionately, the measurement and process models contribute equally to the total area under the verifications curve.

Inspection Process Reliability and Sensitivity Coefficient

These two processes, the Measurement and Inspection, are distinct. This is demonstrated by the fact that in normal operation the Inspection Process is absent. As mentioned previously, the inspection only impinges on the delivered data if there is a false positive such as when the inspection indicates the sensor is good when it has actually failed or is found to be faulty and not replaced immediately because of logistic issues. Conversely, our analysis of the contribution of "faulty" sensors to the uncertainty of measurement relies on a low false negative rate. That is, a low number of inspections that indicate failures when in practice the sensor was sound.

Analysis of 11,992 individual inspections was undertaken to determine the false positive and false negative rate of inspections. Each sensor, with 4 or more inspections at an individual location, were analysed and the results are presented in Table 6. Inspections were assessed in banks of three inspections, the inspection of interest and the one before and after. If, for example, all three were the same, that is three "Pass" then the inspection of interest would be considered valid and count as a true pass. 95% of pass inspections were counted using this method as a true "Passes". However, of the individual "Fails", only 7.2% were considered true fails. Given all the failed inspections make up less than 2% of all inspections

[18] The full width half height of the Lorentzian distributions were used for the estimation as this was considered the most conservative estimate of the Expanded Uncertainty.

and only 7.2% of these were actual fails it indicates that the inspection process is successfully detecting genuinely faulty sensors, 0.03%. Similarly, if we examine the detection of false negatives, 61.2% of all inspection fails are passes.

Table 6. Analysis of False Positive and False Negative rate of Inspection Tests. The first column denotes the outcome of the inspection, the second records the counts of those outcomes. True is a count of three successive inspections with the expected outcome. The count and percentage of False Positives where an inspection was Passed when it should have Failed/Suspect, and False Negatives where the inspection was marked as a Fail/Suspect when it should have been a Pass.

	Count	True		False Positive		False Negative	
Pass	11563	10930	94.5%	18	0.2%	147	1.3%
Suspect	191	0	0.0%	4	2.1%	0	0.0%
Fail	237	17	7.2%	4	1.7%	145	61.2%

As such the sensitivity coefficient for the inspection method was reduced to 0.66 (from the typical value of 1). This determined the basis that the ratio of false negatives in suspect and failed inspections compared to the total number of suspect and failed inspections is 34%. This indicates a significant inflation of the inspection process model uncertainty.

It is noted that in real world measurements, the component of uncertainty that relates to the inspection method is not present during the routine measurement of temperature. It purely relates to the process of checking. As such the inspection method (or process model) artificially inflates the uncertainty evaluation and the use of the sensitivity coefficient mitigates this problem. For this study, based on the false negative and positive rate, a sensitivity coefficient of 0.66 has been applied.

When looking at data where only a single field inspection of a sensor in the field has been undertaken, the effect of the measurement model cannot be separated from the effect of the process model and hence the sensitivity factor should be applied as 1.

Three Measurands

Having taken into consideration the various contributions to uncertainty and the method by which they were determined; the uncertainty, the sensitivity, and independence for each of the components, the total measurement uncertainty under different measurand scenarios can be calculated.

Isolated Measurement

For an isolated or very limited number of measurements, or where there is little supporting evidence or experience of the performance of the observation system, the uncertainty in the measurement of a mercury-in-glass ordinary dry bulb thermometer is approximately 0.45 °C while for an air temperature PRT probe the figure is somewhat larger at 0.51 °C (Table 6).

These uncertainties reflect the consequences of a lack of key supporting information in isolated measurements. The impact of these values can be reduced statistically with repeated inspections, randomisation via the use of multiple sensors or multiple AWS over either time or space.

Isolated measurements are also heavily affected by the inspection method tolerance, which cannot be reduced if there are only a limited number of inspections in its history. In practice this is only relevant for new stations, with less than a year's operation, and/or where there are few nearby locations to verify the quality of the observations.

Table 7 – Table of uncertainties for the "isolated measurement measurand", from both ordinary dry bulb thermometer and air temperature PRT probe.

Isolated Measurement	I/C Factor[19]	U_{95}[20] (°C)	DoF[21]
Ordinary Dry Bulb Thermometer and Observer	1	± 0.45	40
PRT Probe and Electronics	1	± 0.51	6

Typical Measurement

This measurand refers to the data from stations where there is supporting evidence or experience of the performance of the observation system. These stations will have operated for multiple years, and therefore have sufficient supporting evidence such as inspection data, model comparison and overall reliability to engender a level of confidence in the performance and reproducibility of the data gathered.

Typically, 5 to 10 inspections have been undertaken. Additionally, the results can be compared with nearby stations and with stations with similar climatology to confirm the consistency of the performance. As such, the uncertainty rapidly reduces compared to isolated measurements, where there is a lack of understanding of random factors and inspection information.

This estimate of uncertainty is suitable for use in determining maximum and minimum temperatures and for understanding short term changes and trends. Two estimates of expanded uncertainty are provided in Table 8 reflecting the reduction due to longer records. For example, the uncertainty associated with an I/C Factor of 5 applies for a site with 5 to 10 years of operation, while 10+ is relevant for sites with 10 or more years of operation.

Table 8 - Table of uncertainties for the "typical measurement measurand", from both ordinary dry bulb thermometer and air temperature PRT probe.

Typical Measurement Statistics	I/C Factor	U_{95} (°C)	DoF
Ordinary Dry Bulb Thermometer and Observer	5	± 0.23	66
	10+	± 0.18	86
PRT Probe and Electronics	5	± 0.23	9
	10+	± 0.16	12

[19] I/C Factor is a measure of the amount of supporting evidence available to assure the stability of the system (See Section Sensitivity coefficient and Independent or Correlated Factor)

[20] U95 (°C) the expanded uncertainty

[21] DoF Degrees of Freedom

Long-term Measurement

The last measurand applies to aggregated data sets across many stations and over extended periods. This aggregation further mitigates random errors and is suitable for use in determining changes in trends and overall climatic effects.

These are measurements constructed from aggregating large data sets where there is supporting evidence or experience of the performance of the observation systems. Typically, these will be for groups of stations with several years operation and where there are nearby locations to verify the quality of the observations.

Two estimates of expanded uncertainty are provided in Table 9. For example, the uncertainty associated with an I/C Factor of 30 applies for records with a total of more than 30 years of aggregated records and where verification can be undertaken with at least one other verification record. Likewise, the uncertainty associated with an I/C factor of 100 applies when aggregating records greater than 10 years in length for 10 sites to an overall average or, for example a 25-year record with at least five verifying records[22,23]. This estimate of uncertainty is suitable for a range of purposes including determining climate extremes, analysis of trends, and identification of inconsistencies in long term records.

Table 9 - Table of uncertainties of the aggregated “long-term measurement measurand” from ordinary dry bulb thermometer and air temperature PRT probe.

Long-term Measurement Statistic	**I/C Factor**	**U_{95} (°C)**	**DoF**
Ordinary Dry Bulb Thermometer and Observer	30	± 0.14	86
	100	± 0.13	67
PRT Probe and Electronics	30	± 0.11	65
	100	± 0.09	119

[22] I/C Factor = Years of Temperature Record x (No of Temperature Records Sites + No. of Verification Sites -1) the maximum number of Verification Sites used is five.

[23] Note for an I/C Factor of 30 the original record needs to be at least 5 years long and for an I/C Factor of 100 the original record needs to be at least 10 years long.

Conclusion

The temperature measurement uncertainty attributable to the temperature measurement systems installed in the Bureau's surface observations was determined. The importance of understanding the use of the measurement and its relationship to the uncertainty has been discussed and analysed in detail.

Three key measurands were identified for this study: isolated, typical, and long-term measurements. Of these, the typical and long-term estimates of uncertainty are the most useful for data users.

A summary of the uncertainty of the surface temperature parameters measured by field instruments is provided below, using a 95% confidence interval.

Ordinary Dry Bulb Thermometer,

Isolated	± 0.45 °C with a coverage factor of 2.0
Typical	± 0.23 to 0.18 °C with a coverage factor of 2.0
Long-Term	± 0.14 to 0.13 °C with a coverage factor of 2.0

PRT Probe,

Isolated	± 0.51 °C with a coverage factor of 2.6
Typical	± 0.23 to 0.16 °C with a coverage factor of 2.2
Long-Term	± 0.11 to 0.09 °C with a coverage factor of 2.0

When using these estimates of uncertainties, the user needs to consider the environment of the measurements at the time of observation. These values are the averages for the three described measurands. If the differences are less than these values, then the user may be confident that there is no measurable difference. However, if the measured difference is greater than these estimates of uncertainty then difference is likely to be significant, if the meteorological conditions are similar.

Future Directions

This process of analysis has identified a number of interesting avenues to improve our understand of field uncertainty. While currently the use of the inspection data is useful for determination of uncertainties related to Isolated observations, it is not optimal for Typical or Long-Term climatic estimates. In future editions of this report, incorporation of the actual determined drift of sensors should be considered.

Also, the examination of false positives and negatives in the inspection process identified the opportunity to recalculate the contribution of the inspection data, taking into account the known false negatives and laboratory validation of the failure. This will reduce the magnitude of the inspection process uncertainty and improve the overall confidence in isolated uncertainty estimated.

It is also expected in the next 12 months, after implementation of an improved field inspection process, the "inspection process" uncertainties will reduce the uncertainty further. It is recommended that these uncertainties are recalculated in the next 3 to 5 years.

References

AS 1985	Australian Standards, Thermometers - Meteorological - Maximum and Minimum, AS2819, 1985.
Bennett 2005	Bennett, K., Calibrate with Confidence, Metrology Concepts: Understanding Test Uncertainty Ratio (TUR); Transact Calibration and Repair Services, May 2005
Bentley 1999	Bentley, R., Handbook of Temperature Measurement, Vol 1, p23, Springer-Verlag Singapore Pte, 1999
BIPM 1997	Techniques for Approximating the International Temperature Scale of 1990 (ITS-90), Bureau International des Poids et Measures, 1997.
BoM 1989	Bureau of Meteorology, Guidance Specification (Functional) for a General-Purpose Automatic Weather Station, Equipment Specification No. A2659, 1989.
BoM 1990	Bureau of Meteorology, Drawing Instrument Shelter Meteorological, Large; DRG No.MI0401-01 – 09, 9/7/1990.
BoM 1991	Bureau of Meteorology, Specification for Provision of an Automatic Weather Station Type I (High Performance Version), Equipment Specification No. A2671, 1991.
BoM 1997	Bureau of Meteorology, Guidelines for the Siting and Exposure of Meteorological Instruments and Observing Facilities, Observation Specification No. 2013.1, Inspection Handbook Appendix 4.6. 1997.
BoM 2001	Bureau of Meteorology, Surface Observational Handbook, 2001.
BoM 2003	Bureau of Meteorology, Specification for a Digital Telemetry System Equipment Data Formats Document, Equipment Specification A2669,5.1, 2003.
BoM 2010	Bureau of Meteorology, Inspection Handbook, 2010.
BoM 2011	Bureau of Meteorology, Specification for an Electronic air temperature probe, Equipment Specification No. A3001, 2011.
BoM 2017	Bureau of Meteorology, Drawing Instrument Shelter Meteorological, Small; DRG No.MI0201-01 – 13, 24/3/2017.
BoM 2021	Warne, J.O., Benbow, D., Dollery, I., Whitehead, H, and Lam, T., Performance characteristic and Uncertainty analysis of ALMOS Electronics, In preparation.
BoM 1982	Bureau of Meteorology, Meteorology Thermometers, Ordinary, Maximum and Minimum, Bureau specification A410 (Issue 6), 1982
Clark 2016	Clark, M., Dent-Jones, T., and Moore, D., Investigating the impact on screen temperature of concrete surfaces within the meteorological enclosure, WMO TECO 2016, P1(7)
Coppa 2021	Coppa, G., Quarello, A., Steeneveld, G-J., Jandric, N. and Merlone, A.; Metrological Evaluation of the Effect of the Presence of a Road on Near-Surface Air Temperatures, International Journal of Climatology, 2021, p 1-20. DOI: 10.1002/joc.7044.

Dollery 2008 — Dolley, I., and Storm, J., Calculation of Uncertainty for Working References in Temperature, Instrument Test Report 686, Bureau of Meteorology, January 2008.

IEC 2008 — International Electrotechnical Commission, Industrial platinum resistance thermometers and platinum temperature sensors, BS EN 60751, 2008.

ISO 2005 — International Organization for Standardization, General requirements for the competence of testing and calibration laboratories, ISO/IEC 17025, 2005.

ISO 2009 — Uncertainty of Measurement – Part:1 Introduction to the Expression of Uncertainty in Measurement, ISO/IEC Guide 98-1:2009, 2009.

JCGM100:2008 — Joint Committee for Guides in Metrology, Guide to the expression of uncertainty in measurement, (GUM) https://www.bipm.org/en/publications/guides/

JCGM200:2008 — Joint Committee for Guides in Metrology, International vocabulary of metrology — Basic and general concepts and associated terms (VIM), https://www.bipm.org/en/publications/guides/

Kinoshita 2014 — An Evaluation Method of the Effect of Observation Environment on Air Temperature Measurements, Boundary-Layer Meteorology, 2014, 154:91-105, DOI 10.1007/s10546-014-9918-2

NATA 2018 — National Association of Testing Authorities, Specific Accreditation Criteria, ISO/IEC 17025 Application Document Calibration – Appendix, July 2018

Parker 1994 — Parker, D. E., Effects of Changing Exposure of Thermometers at Land Stations, International Journal of Climatology, Vol 14, 1-31, 1994.

Preston-Tomas 1990 — Preston-Thomas, H., The International Temperature Scale of 1990 (ITS-90), Metrologia, Vol 27, pp 3-10, 1990.

Warne 1998 — Warne, J.O., A Preliminary Investigation of Temperature Screen Design and their Impacts on Temperature Measurements, Bureau of Meteorology, ITR649 Report, 1998

Warne 2016 — Warne, J.O., and Dollery, I, Real World Performance of Automatic Weather Stations in the Field, WMO CIMO-TECO Conference, Madrid, 2016 https://community.wmo.int/teco-2016session1

WMO 2018 — World Meteorological Organization, Guide to Meteorological Instruments and Methods of Observation, WMO Guide No. 8, 2018.

WMO 2019 — World Meteorological Organization, Manual on Codes, International Codes, Vol I.1, Annex II to the WMO Technical Regulations, Part A, 2019

Appendix A

Measurement Tolerance

Table 10. Tolerances for Purchased Meteorological Thermometers in Celsius from AS2819-1985.

Type of thermometer	Permissible error at any graduation mark				Permissible algebraic difference in errors at each end of any 10 °C interval		
	< 0 °C	0 °C	0 to 25 °C	> 25 °C	< 0 °C	0 to 25 °C	> 25 °C
PRT probe	± 0.08	± 0.05	± 0.08	± 0.08			
Ordinary dry bulb thermometer	+0.15 -0.30	-	+0.05 -0.15	+0.05 -0.15	0.20	0.10	0.10

Appendix B

Ordinary Dry Bulb Thermometer

Instrument

The Ordinary Dry Bulb thermometer is a manually read mercury in glass thermometer. It is mounted vertically within the screen. The ordinary dry bulb thermometer is also known as both an ordinary thermometer and, as it uses mercury as its indication fluid, a MIG (mercury in glass).

The scale is divided in 1.0 °C intervals and subdivided into 0.5 °C intervals. The user further subdivides these divisions, by eye, to read the thermometer to a resolution of 0.1C.

The temperature is read from the top of the curved mercury meniscus. Refer to Figure 2.

Figure 2. Ordinary dry bulb thermometer with a temperature measurement of 23.7 °C.

Measurement

The Ordinary Dry Bulb thermometer is used to provide a manual reading of the air temperature within the screen.

The method for reading the thermometer is described in the Surface Observations Handbook Volume 1, Part 3, Section 3.2 Ordinary Dry Bulb Thermometer [BoM 2001]. The measurement parameter is described in WMO Guide 8 Part 1 Chapter 2 [WMO 2018].

The screen temperature is recorded, for the synoptic records, every three hours. This is performed continuously over the 24-hour period from 9:00 am to 9:00 am on the following day.

Measurement Specifications

The purchase and measurement specifications for ordinary dry bulb thermometers is covered by a Bureau specification A410(issue 8) [BoM 2011] “Meteorology thermometers Ordinary - Maximum and Minimum”, and the withdrawn Australian Standard AS2819-1985 [AS 1985]. The measurement tolerances from the Australian Standard for the thermometer are provided in Appendix A.

Ordinary Dry Bulb thermometers were purchased from manufacturers who were accredited under the international laboratory quality assurance standard ISO 17025-2005 [ISO 2005] to provide traceably calibrated thermometers. The Bureau does not perform additional verification on purchased ordinary dry bulb thermometers.

In 2008, the Bureau continued purchasing ordinary dry bulb thermometers for the field, but the Inspectors thermometers were replaced by Electronic field references. The mercury-in-glass thermometers are still used by co-operative observers but are being phased out at other stations. A few staffed sites continue to use thermometers as a cross check of the air temperature PRT probe.

In the field, the ordinary dry bulb thermometers are checked (at least annually, twice per year at climate sites) by an inspector against their electronic reference to ensure it is within ±0.40 °C of their inspection reference.

The relevant measurement specifications for ordinary dry bulb thermometers are:

Range	-20 °C to +60 °C
Divisions	Major division every 1 °C and a minor division every 0.5 °C.
Index Fluid	Triple distilled mercury
Tolerance	The thermometer shall be tested at intervals of 5 °C. The error at any such point on the scale, and the algebraic difference in the errors at the beginning and the end of any 10 °C interval shall not exceed the values shown Appendix A.

Appendix C

Air Temperature PRT Probe

Instrument

The air temperature probe's sensor determines temperature via the change in resistance of a pure thin platinum wire. The Bureau uses platinum sensors which conform to the international standard BS EN 60751-2008 [IEC 2008], band 5 (1/10 DIN) which have a $R0 = 100\ \Omega$ and $\alpha = 0.003850\ \Omega/(\Omega.\ °C)$.

The equation which translates the four-wire resistance measurement of the platinum sensor into temperature is also specified in the standard [IEC 2008] and is also known as the Callendar Van Dusen equation.

To provide mechanical, chemical and electrical protection, the sensor is mounted within a stainless-steel sheath. The sheath design has changed over time; most significantly, the sheath's diameter has reduced, to reduce its time constant. Refer to Figure 3.

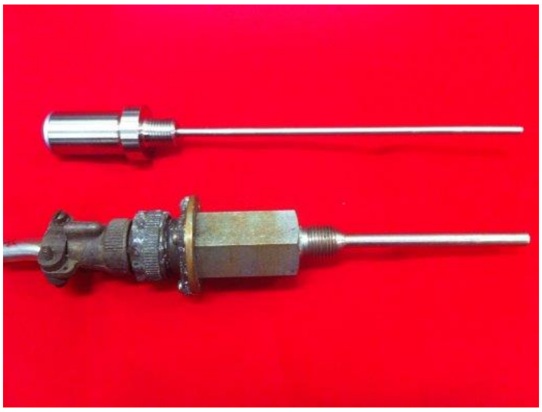


Figure 3. Air Temperature (PRT) probes (top) slim current (bottom) thick previous.

The electronic thermometer is mounted within the Stevenson screen vertically in the same way as the mercury ordinary dry bulb thermometer refer Appendix E Figure 4 (item 10).

Measurement

The PRT probe is an electronic thermometer used to provide an automated reading of the air temperature within the screen. Refer to Appendix E Figure 4.

The temperature from the PRT probe is recorded for synoptic records every three hours by an Automatic Weather Stations (AWS), additionally the AWS continuously provides several other temperature measurements products, from one-second instantaneous values to averages, maximums and minimums for both one-minute and ten-minute periods. This analysis covers the reported one-minute temperatures.

The method for reading and reporting the instantaneous temperature from the electronic thermometer is described in a Bureau Specification [BoM 2003]. The measurement parameter is described in WMO Guide 8 Part 1 Chapter 2 [WMO 2018].

The PRT probe is also known under a variety of different names such as an industrial platinum resistance thermometer (IPRT or PRT) or resistance temperature device (RTD).

Measurement Specifications

The purchase and measurement specifications for thermometers is covered by a Bureau specification A3001 [BoM 2011] and it refers to the withdrawn Australian Standard AS2819 [AS 1985]. The measurement tolerances from the Australian Standard are provided in Appendix A.

The relevant measurement specifications for electronic thermometers are covered in the Bureau specification A3001 [BoM 2011] "Equipment Specification - Electronic Air Temperature Probe":

The measurement specifications for this thermometer are:

Range	-10 °C to +55 °C pre 2019 and -30 °C to +60 °C post 2019[24]
Sensor	Meet the minimum requirements of IEC 60751 [IEC 2008] (1/10 DIN) (Band 5).
R0	100 Ω
Alpha	0.00385 Ω/°C/Ω
Tolerance	The thermometer shall be tested at intervals of 5 °C. The errors at any point shall not exceed the values shown Appendix A.

[24] This is the purchase specification range for devices purchased prior to 2019. The current purchase specification has been extended to -30 °C to 60 °C. In practice there is no physical difference in the sensors as they use the same IEC 60751 (1/10 DIN) (Band 5) element. The specified base accuracy for the sensing element is no worse than 0.05 °C at -30 °C, 0.03 °C at -10 °C and 0 °C is 0.07 °C at 50 °C and 0.08 °C at 60 °C [IEC 2008]. This defines the lower limit of the uncertainty budget for the temperature instrument. The instrument package which includes the steel sheath, packing material and sensing element has an uncertainty greater than that specified in IEC 60751. See Appendix D.

Appendix D Measurement Uncertainties

Ordinary Dry Bulb Thermometer – (See Method Section, Sources of Uncertainty for explanation of Uncertainty Contributors)

Isolated/Unsupported Single Measurement - Ordinary Dry Bulb Thermometer								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
Field Instrument Performance								
Calibration uncertainty of the determined corrections	0.10	°C	2.000	1	1	0.050	30	A
Conformance of the divisions to temperature (Residual and Scatter)	0.15	°C	1.732	1	1	0.087	10	B
MIG secular change between inspection (Drift/year)	0.01	°C	1.732	1	1	0.006	60	B
Measurement Interface - Human Performance								
Resolution & human ability to subdivide scale.	0.20	°C	1.732	1	1	0.115	30	B
Parallax	0.10	°C	1.732	1	1	0.058	30	B
Human reproducibility (between operators)	0.14	°C	2.000	1	1	0.070	30	B
Human Repeatability (single operator)	0.07	°C	2.000	1	1	0.035	30	B
Message Reporting Output								
Resolution	0.05	°C	1.732	1	1	0.029	30	B
Inspection Reference								
Inspection instruments' residual corrections	0.070	°C	2.000	1	1	0.035	60	B
The inspection instruments resolution	0.005	°C	1.732	1	1	0.003	60	B
Calibration uncertainty on inspection reference	0.037	°C	2.000	1	1	0.019	60	B
Inspection Method (Environmental Scatter)								
Environment and coupling of the reference to the sensor	0.333	°C	2.000	0.66	1	0.110	3	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	1	0.047	30	B
Discolouration impacts	0.020	°C	2.000	1	1	0.010	30	B
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	1	0.000	30	B
Combined Standard Uncertainty		°C				0.223	39.9	
Combined Expanded Uncertainty (95%)	0.45	°C	2.023					

Typical Measurement - Ordinary Dry Bulb Thermometer								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
MIG Performance								
Calibration uncertainty of the determined corrections	0.100	°C	2.000	1	1	0.050	30	A
Conformance of the divisions to temperature	0.150	°C	1.732	1	10	0.027	10	B
MIG secular change between inspection (Drift/year)	0.010	°C	1.732	1	10	0.002	60	B
Human Performance								
Resolution & human ability to subdivide scale.	0.200	°C	1.732	1	10	0.037	30	B
Parallax	0.100	°C	1.732	1	10	0.018	30	B
Human reproducibility (between operators)	0.140	°C	2.000	1	10	0.022	30	B
Human Repeatability (single operator)	0.070	°C	2.000	1	10	0.011	30	B
Reporting Output								
Resolution	0.050	°C	1.732	1	10	0.009	30	B
Inspection Reference								
Inspection instruments' residual correction limits	0.070	°C	2.000	1	1	0.035	60	B
The inspection instruments resolution; and	0.005	°C	1.732	1	10	0.001	60	B
Calibration uncertainty on inspection instrument	0.037		2.000	1	10	0.006	60	B
Inspection Method (Environmental Scatter)								
Scatter introduced by unstable Field Environment	0.333	°C	2.000	0.66	10	0.035	3	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	5	0.021	30	A
Discolouration impacts	0.020	°C	2.000	1	10	0.003	30	A
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	10	0.000	30	B
Combined Standard Uncertainty		°C				0.092	85.5	
Combined Expanded Uncertainty (95%)	0.18	°C	1.988					

Long Term Measurement Statistic - Ordinary Dry Bulb Thermometer								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
MIG Performance								
Calibration uncertainty of the determined corrections	0.100	°C	2.000	1	1	0.050	30.0	A
Conformance of the divisions to temperature	0.150	°C	1.732	1	100	0.009	10.0	B
MIG secular change between inspection (Drift/year)	0.010	°C	1.732	1	100	0.001	60.0	B
Human Performance								
Resolution & human ability to subdivide scale.	0.200	°C	1.732	1	100	0.012	30.0	B
Parallax	0.100	°C	1.732	1	100	0.006	30.0	B
Human reproducibility (between operators)	0.140	°C	2.000	1	100	0.007	30.0	B
Human Repeatability (single operator)	0.070	°C	2.000	1	100	0.004	30.0	B
Reporting Output								
Resolution	0.000	°C	1.732	1	100	0.000	30.0	B
Inspection Reference								
Inspection instruments' residual correction limits	0.070	°C	2.000	1	1	0.035	30.0	B
The inspection instruments resolution; and	0.005	°C	1.732	1	100	0.000	30.0	B
Calibration uncertainty on inspection instrument	0.037	°C	2.000	1	100	0.002	30.0	B
Inspection Method (Environmental Scatter)								
Scatter introduced by unstable Field Environment	0.333	°C	2.000	0.66	100	0.011	3.0	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	50	0.007	30	A
Discolouration impacts	0.020	°C	2.000	1	100	0.001	30	A
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	100	0.000	30	B
Combined Standard Uncertainty		°C				0.065	66.7	
Combined Expanded Uncertainty (95%)	0.13	°C	1.997					

Air Temperature PRT Probe

Isolated/Unsupported Single Measurement –PRT Probe								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
Field Instrument Performance								
Laboratory calibration uncertainty	0.043	°C	2.000	1	1	0.022	30	A
Sensor residual correction and scatter	0.099	°C	2.000	1	1	0.049	10	B
Sensor drift between inspections	0.001	°C	2.000	1	1	0.0003	60	B
Measurement Interface - AWS Electronics								
Calibration uncertainty on resistance source	0.023	°C	2.000	1	1	0.012	30	B
AWS residual corrections	0.027	°C	2.000	1	1	0.013	30	B
Reproducibility (Temp and non-Linear included)	0.112	°C	2.000	1	1	0.056	30	B
Repeatability	0.033	°C	2.000	1	1	0.016	30	B
Algorithm conversion to temperature	0.005	°C	2.000	1	1	0.003	30	B
Message Reporting Output								
Resolution	0.050	°C	1.732	1	1	0.0289	60	B
Inspection Reference								
Inspection instruments' residual corrections	0.063	°C	2.000	1	1	0.032	60	B
The inspection instruments resolution	0.005	°C	1.732	1	1	0.003	60	B
Calibration uncertainty on inspection reference	0.037	°C	2.000	1	1	0.019	60	B
Inspection Method (Environmental Scatter)								
Environment and coupling of the reference to the sensor	0.506	°C	2.000	0.66	1	0.167	3	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	1	0.047	30	B
Discolouration impacts	0.020	°C	2.000	1	1	0.010	30	B
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	1	0.000	30	B
Combined Standard Uncertainty		°C				0.197	5.8	
Combined Expanded Uncertainty (95%)	0.51	°C	2.571					

Typical AWS Measurement - PRT Probe								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
Field Instrument Performance								
Laboratory calibration uncertainty	0.043	°C	2.000	1	1	0.022	30	A
Sensor residual correction and scatter	0.099	°C	2.000	1	10	0.016	10	B
Sensor drift between inspections	0.001	°C	2.000	1	10	0.0001	60	B
Measurement Interface - AWS Electronics								
Calibration uncertainty on resistance source	0.027	°C	2.000	1	1	0.013	30	B
AWS residual corrections	0.112	°C	2.000	1	10	0.018	30	B
Reproducibility (Temp and non-Linear included)	0.033	°C	2.000	1	10	0.005	30	B
Repeatability	0.005	°C	2.000	1	10	0.001	30	B
Algorithm conversion to temperature	0.005	°C	2.000	1	10	0.001	30	B
Message Reporting Output								
Resolution	0.050	°C	1.732	1	10	0.009	60	B
Inspection Reference								
Inspection instruments' residual corrections	0.063	°C	2.000	1	1	0.032	60	B
The inspection instruments resolution	0.005	°C	1.732	1	10	0.001	60	B
Calibration uncertainty on inspection reference	0.037	°C	2.000	1	10	0.006	60	B
Inspection Method (Environmental Scatter)								
Environment and coupling of the reference to the sensor	0.506	°C	2.000	0.66	10	0.053	3	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	5	0.021	30	B
Discolouration impacts	0.020	°C	2.000	1	10	0.003	30	B
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	10	0.000	30	B
Combined Standard Uncertainty		°C				0.075	11.8	
Combined Expanded Uncertainty (95%)	0.16	°C	2.201					

Long Term Measurement Statistic - PRT Probe								
Uncertainty Contributors	**Expanded Uncertainty**	**Unit**	**Coverage Factor**	**Sensitivity Coefficient**	**Independent/ Correlated Factor**	**Standard Uncertainty**	**Degrees of freedom**	**Type**
Field Instrument Performance								
Laboratory calibration uncertainty	0.043	°C	2.000	1	1	0.022	30.0	A
Sensor residual correction and scatter	0.099	°C	2.000	1	100	0.005	10.0	B
Sensor drift between inspections	0.0006	°C	2.000	1	100	0.0000	60.0	B
Measurement Interface - AWS Electronics								
Calibration uncertainty on resistance source	0.027	°C	2.000	1	1	0.013	30.0	B
AWS residual corrections	0.112	°C	2.000	1	100	0.006	30.0	B
Reproducibility (Temp and non-Linear included)	0.033	°C	2.000	1	100	0.002	30.0	B
Repeatability	0.005	°C	2.000	1	100	0.0003	30.0	B
Algorithm conversion to temperature	0.005	°C	2.000	1	100	0.0003	30.0	B
Message Reporting Output								
Resolution	0.050	°C	1.732	1	100	0.003	60.0	B
Inspection Reference								
Inspection instruments' residual corrections	0.063	°C	2.000	1	1	0.032	60.0	B
Inspection instruments resolution;	0.005	°C	1.732	1	100	0.000	60.0	B
Calibration uncertainty on inspection instrument	0.037	°C	2.000	1	100	0.002	60.0	B
Inspection Method (Environmental Scatter)								
Environment and coupling of the reference to the sensor	0.506	°C	2.000	0.66	100	0.017	8.0	B
Screen Type, Maintenance and Siting (WMO Class 1 or 2)								
Uncertainty due to screen size	0.094	°C	2.000	1	50	0.007	30	B
Discolouration impacts	0.020	°C	2.000	1	100	0.001	30	B
Siting (WMO Class 1 or 2)	0.000	°C	1.732	1	100	0.000	30	B
Combined Standard Uncertainty		°C				0.045	119.1	
Combined Expanded Uncertainty (95%)	0.09	°C	1.980					

Appendix E

Stevenson Screen Layout

The layout of the temperature measurement instruments mounted inside the instrument screen is provided in Figure 4 below.

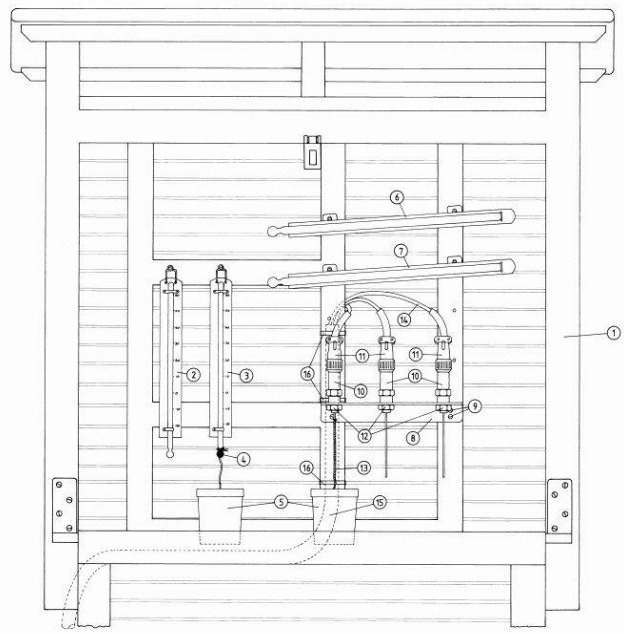


Figure 4. Screen Instruments Layout. Items are identified as follows:
(1) Stevenson Screen, (2) Ordinary Dry Bulb and (3) Wet Bulb Thermometers, (6) Maximum and (7) Minimum Thermometers, and (10) Air Temperature (PRT) Probes