Challenges in the Transition from Conventional to Automatic Meteorological Observing Networks for Long-term Climate Records

2017 edition

WEATHER CLIMATE WATER

WORLD
METEOROLOGICAL
ORGANIZATION

WMO-No. 1202

Challenges in the Transition from Conventional to Automatic Meteorological Observing Networks for Long-term Climate Records

2017 edition

WORLD
METEOROLOGICAL
ORGANIZATION

WMO-No. 1202

EDITORIAL NOTE

METEOTERM, the WMO terminology database, may be consulted at http://public.wmo.int/en/resources/meteoterm.

Readers who copy hyperlinks by selecting them in the text should be aware that additional spaces may appear immediately following http://, https://, ftp://, mailto:, and after slashes (/), dashes (-), periods (.) and unbroken sequences of characters (letters and numbers). These spaces should be removed from the pasted URL. The correct URL is displayed when hovering over the link or when clicking on the link and then copying it from the browser.

WMO-No. 1202

Chairperson, Publications Board
World Meteorological Organization (WMO)
7 bis, avenue de la Paix
P.O. Box 2300
CH-1211 Geneva 2, Switzerland

Tel.: +41 (0) 22 730 84 03
Fax: +41 (0) 22 730 81 17
Email: publications@wmo.int

ISBN 978-92-63-11202-6

PUBLICATION REVISION TRACK RECORD

Date	*Part/ chapter/ section*	*Purpose of amendment*	*Proposed by*	*Approved by*

CONTENTS

ACKNOWLEDGEMENT

Mr Blair Trewin, Bureau of Meteorology, Australia, is thanked for his outstanding contribution to this publication.

1. INTRODUCTION

Note: This guidance note discusses, and provides examples and guidance regarding, the homogeneity of long-term climate records in the transition from conventional to automatic meteorological observing networks.

Automatic weather stations (AWSs) are playing an increasing role in meteorological observing networks, in both developed and developing countries. They offer numerous advantages in weather applications, by allowing observations at a high time resolution to be received in real time, and at a relatively low recurrent cost after the initial capital cost of installation. Many locations that previously had observations only a few times per day now have a continuous feed of data every minute. These AWSs also allow observations to be made in areas without permanent human populations, and in locations not readily accessible by human observers (such as in the centres of airports). On the other hand, they typically require more-frequent and more-specialized maintenance than manual systems, which can place a significant burden on network management in some places, especially countries with limited resources. They also introduce challenges for climate observation programmes, because of issues associated with the transition from conventional to automatic measurements, and because of some characteristics of automatic measurements, such as different rates of instrument malfunctions.

Some countries have either fully transitioned their synoptic networks to AWSs, or have declared their intention to do so (for example, Germany plans to move to a fully automatic network by 2020; Claussnitzer et al., 2015), while some developing countries have used the introduction of AWSs to support a major expansion of previously sparse networks.

The term “automatic weather station” can encompass a wide variety of station types. At the basic end of the scale, stations measuring a relatively narrow range of variables with limited but useful accuracy can be purchased for a few hundred dollars through normal commercial channels, and are widely used by private individuals and small organizations in many countries. At the advanced end of the scale, professional stations that meet WMO performance standards can cost tens of thousands of dollars; in addition to basic variables such as temperature, humidity, wind speed and direction, air pressure and precipitation, they can also include sensors for parameters such as visibility, cloud amount and type, and present weather. In some countries, only data from National Meteorological and Hydrological Service (NMHS)-owned AWSs are used by the NMHSs, while in other countries, the NMHSs also incorporate data from AWSs owned by other organizations or individuals into their products and analyses. For the purposes of this guidance note, the AWSs will not be considered to include autographic instruments, such as barographs or anemographs, that record continuously but require manual intervention to read information from a chart or similar; however, they do include devices that produce a digital output which requires manual intervention to communicate it to the broader communications network.

Guidelines for managing changes in climate observation programmes, with a set of recommended practices by which such changes can be managed, have already been published (WMO, 2007). This guidance note will focus on those AWSs likely to form part of long-term climate records. These will normally be owned by NMHSs or associated agencies, or sometimes by third parties (such as aviation, agriculture or road transport agencies) under standards endorsed by the NMHSs. Experience shows that private AWSs can be useful for some climate purposes (such as providing information on individual local-scale extreme events) but rarely have the length of record, accuracy or exposure standards, or long-term stability of site and instrumentation, to have much value for monitoring on timescales of a decade or longer.

Automation of an observing network has many advantages, but it also introduces challenges for long-term climate monitoring. Any change in an observing system potentially introduces an inhomogeneity into the climate record, which needs to be assessed and adjusted for, if necessary. Assessment of an inhomogeneity at a site is especially challenging if similar changes occur at a large number of sites (which might otherwise be reference stations) over a short period of time, as discussed in Chapter 4. In many cases, the introduction of AWSs will also be accompanied by site relocations. In addition, the introduction may lead to the loss of observations of variables that are difficult to measure automatically (or where automatic measurements are not directly comparable with manual observations), such as cloud amount, snow depth or pan evaporation. Once AWSs become widespread in a network, it may also be difficult to maintain a consistent instrument type because of competitive tendering or procurement policies in some countries.

This guidance note should be read in conjunction with other WMO publications, including the Commission for Instruments and Methods of Observation (CIMO) *Guide to Meteorological Instruments and Methods of Observation* (WMO, 2014; henceforth referred to as CIMO 2014). In particular, CIMO 2014 provides detailed guidance (Part II, Chapter 1) on the use of AWSs. This note will focus on aspects of AWSs that are relevant to long-term climate measurements; readers are referred to CIMO 2014 for detailed information on AWSs in general. The *Guide to Climatological Practices* (WMO, 2011) discusses aspects of AWSs within Chapter 2 (Climate observations, stations and networks), and refers to the overarching climate monitoring principles established by the WMO Global Climate Observing System. The *Guidelines on Climate Metadata and Homogenization* (WMO, 2003) discuss and provide guidance regarding homogeneity test methods and homogenization (an update is expected at the time of writing this note).

2. GENERAL AND ELEMENT-SPECIFIC ASPECTS OF THE TRANSITION FROM CONVENTIONAL TO AUTOMATIC MEASUREMENTS

2.1 General aspects

2.1.1 *Data completeness*

When operating optimally, AWSs are capable of providing data flows with a continuity and time resolution well beyond the capacity of even the most diligent of observers. They can provide data at resolutions of 1 min or less, and operate just as effectively at night and over weekends as they do during the standard working week.

Both conventional and automatic observations can suffer from breaks in data, but the way in which breaks happen can take different forms. In the case of conventional observations, the most-common reason for missing observations is the absence of the observer, either scheduled (for example, no observations over weekends) or unscheduled (for example, through illness). There can also be extended periods of missing data, or even cessation of observations altogether, if an observer leaves their position (for example, through resignation, illness or death) and a replacement cannot be recruited quickly – this is a particular risk in remote areas with small local populations. Instrument faults or failures occasionally cause missing data for conventional observations, but communications failures rarely do, as the data can normally be retained for later transmission.

Automatic systems do not need the presence of an observer. However, electronic systems tend to be more susceptible to failure than conventional systems (especially those parts of conventional systems that do not have any moving parts such as thermometers and raingauges), and can be affected by issues such as lightning strikes (to which anemometers are particularly susceptible because they are on tall masts) or power supply failure (whether that is mains, battery or solar power). Automatic observations can also be affected by failures of the communication system. These can occur at the station, at the point of ingest into a database or in the telecommunications system in between. The extent to which the affected data are recoverable depends on factors such as the methods used for data transmission and whether data are logged on site (and if so, for how long). Instrument or communications systems failures often require specialist technical expertise to remedy them, whereas a broken thermometer, for example, can be replaced with a spare one kept on site or sent to the station within a few days. This can result in extended outages if that expertise is not available locally, especially in remote areas. Experience with Australian data since the introduction of AWSs indicates that the overall percentage of missing data is similar at conventional and automatic stations, but that the mean duration of outages at automatic stations is longer. Data from Australia and Spain indicate that intermittent short outages, repeated over a longer period, are another failure mode that can occur at AWSs. These can affect daily observations and hence the totals and averages for longer periods.

2.1.2 *Differences in error modes between conventional and automatic measurements*

Conventional observations are susceptible to a wide range of human errors, which are eliminated in automatic systems. Examples of these include:

(a) Instrument misreading errors (for example, reading the wrong end of a thermometer index or misreading the value by 5 ° or 10 °);

(b) Transcription and data entry errors (for example, miskeying of data when information from paper forms is entered into a database);

(c) Calculation errors (for example, conversion of station-level to mean sea-level pressure, where it is done manually or through lookup tables, which is common with historical data).

For some meteorological elements, conventional observations will also inevitably depend on observer judgement, which may vary depending on the skill, experience and diligence of the observer. An obvious example is that of cloud amount, type and height, where considerable observer judgement is required, and a change in observer is a common source of data inhomogeneities (Jovanovic et al., 2010). Other examples include visibility estimates, and wind observations derived from Beaufort scale estimates.

While the sources of error mentioned above are largely or totally eliminated for AWSs, AWSs can generate erroneous data for other reasons, including electronic or mechanical failures. A particularly common error mode for AWSs involves very rapid changes in an observed variable ("spikes") to an often totally unrealistic value (Figure 1). These often arise from power surges or electrical interference at some point in the system. In sophisticated AWS networks, this is becoming less of a problem as algorithms are applied to filter out spurious data (see below), but it remains a significant problem for early-generation or lower-sophistication AWSs, as are sometimes found in developing countries, or in the earlier years of AWS use in more developed countries. Errors can also occur in AWS internal systems, such as the algorithms used to convert signals from the instruments to the values of meteorological variables. These can be particularly challenging to detect and correct if the AWS software is a "black box" provided by the manufacturer and not directly accessible to the NMHS.

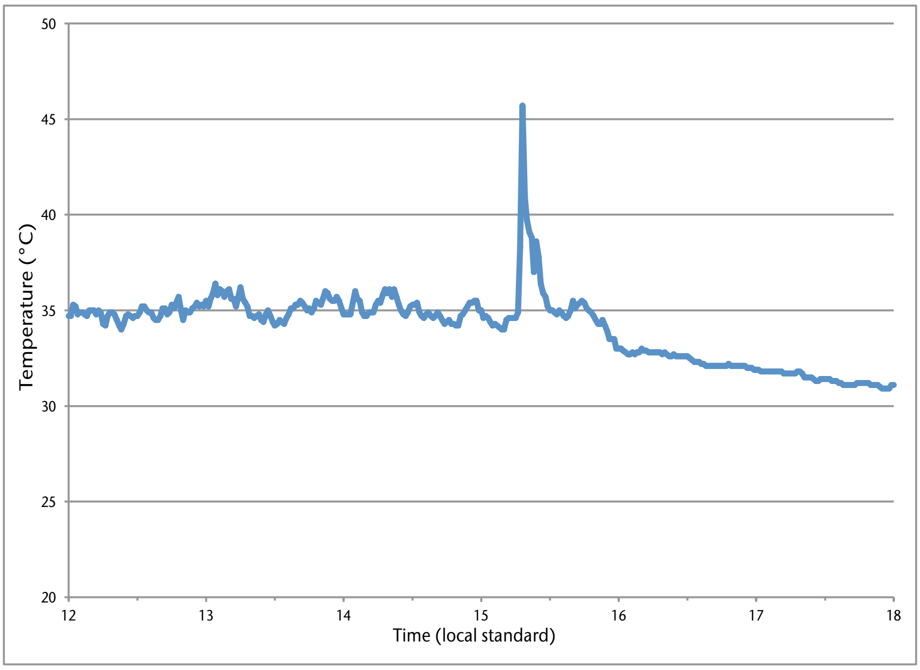

Figure 1. Example of a data spike: dry bulb temperatures at Bulman, Australia, 5 May 2016

2.1.3 *Consistency of practices between conventional and automatic systems*

In general, it is desirable to define climatological variables measured with automatic stations in a way that matches, as closely as possible, the equivalent variables measured with conventional stations. This is particularly necessary for variables that are defined over a day. The way in which daily variables are calculated from AWS data streams (which are often at a resolution of 1 min or shorter) is therefore important.

The higher time resolution of automatic stations, and their functioning at all times of the day, can introduce measurement possibilities that do not exist for conventional stations. Two examples are:

(a) For an automatic station, it is easy to define a climatological day as ending at midnight, whereas for a conventional station, it is often difficult to find observers who are available at midnight, and other observation times are often chosen (for example, 0900);

(b) A daily mean value at an automatic station (for example, mean sea-level pressure) can be calculated continuously across all observations, whereas at conventional stations, only a small number of observations (for example, four 6-hourly observations) can be used.

In these cases, while making use of the additional possibilities of an AWS may appear to be an "improvement", doing so can introduce potential inhomogeneities into the climate record. There may also be possible inconsistencies across an observation network if, for example, different stations are reporting at different times. Matching, as much as possible, the definitions used for variables at AWSs to the definitions used for conventional stations is preferable.

2.1.4 *Maintenance, calibration and tolerance checks*

It is recommended by WMO that all synoptic land stations and principal climatological stations should be inspected no less than once every two years (CIMO 2014, Part I, section 1.3.5). In addition, WMO guidelines state that manufacturers' recommended checks on automatic instruments (which will vary from instrument to instrument) should be carried out. Automatic sensors, particularly those that are electrically based, can drift in the field, and regular tolerance checks[1] are an important part of ensuring the long-term stability and homogeneity of observations. There are also benefits to conducting more-rigorous calibrations of sensors from time to time. Should these not be carried out, the risk arises that instrument drift can produce spurious trends in data, which may be difficult to detect if they develop gradually over time. A lack of tolerance checks or calibrations also reduces the traceability of data.

In addition to regular inspections and maintenance in the field, AWSs will inevitably have some unexpected outages. This risk can be mitigated through measures such as having some redundant sensors (in case of failure of a single sensor), or having a substantial on-station data logging capacity to allow data recovery in the event of a communications failure (this requires an adequate power supply). Many outages are resolved without requiring direct intervention by the observation network provider, and some can be resolved by local personnel; however, some will require intervention by specialist technical staff.

A common scenario is that an AWS network is installed, but the funding source for the installation makes limited or no provision for ongoing maintenance (scheduled or unscheduled). Other challenges can include a shortage of suitably qualified technical staff, and sites that are difficult for technical staff to reach quickly (for example, they are a long distance from the NMHS headquarters, or at difficult-to-reach locations, such as up high mountains or on offshore islands).

Therefore, in countries where maintenance support is limited or non-existent, some AWSs are prone to extended outages or other maintenance problems (Page et al., 2004). This problem

---

[1] Tolerance checks in the field are often referred to as "calibrations", but strictly speaking, the term "calibration" is restricted to a comparison with a formal reference standard.

is particularly acute in some developing countries, where funds (sometimes provided by aid donors or similar) are limited in scope and technical expertise is also limited. In the worst cases, AWS networks have become almost inoperative within a few years of installation.

2.1.5 *Spike filtering, time sampling and other algorithms*

As noted above, a common data quality issue for AWSs is the occurrence of spikes. Algorithms are increasingly deployed in data processing to filter such spikes from data streams. Nevertheless, data quality control methods still need to be used in climate analyses to detect and flag erroneous values, especially for historical data. This is because it cannot be assumed that the quality standards that apply to present data necessarily also apply to data from the early years of AWS use. Data spikes are a particular issue for analyses of climate extremes because, by their nature, they often fall outside the normal range of observations and will hence show as spurious extremes unless filtered.

A wide range of software and hardware is utilized by AWSs to convert signals from instruments (for example, the electrical resistance of a temperature probe or electrical signals generated each time an anemometer rotates) into the values of meteorological variables. This is a process that will also involve the selection of time windows over which to sample the variables. In some instances, software changes may produce inhomogeneity in one or more climate variables. It is therefore important for software version changes and changes in any internal hardware (including electronic interface modules) to be included in station metadata. An additional challenge is provided by the fact that AWS software is often proprietary, and the effect of version changes implemented by manufacturers may not be transparent to the network operator or to data users.

2.2 Specific meteorological elements

2.2.1 *Temperature*

Temperature is the variable that has undergone the greatest level of homogeneity scrutiny. Historically, substantial changes in observing technology, notably changes in the shields or shelters used to protect the instruments from direct or indirect solar radiation, have resulted in sometimes substantial inhomogeneities in temperature measurements. As an example, the introduction of standard instrument shelters (Stevenson screens or similar) in numerous countries in the late nineteenth and early twentieth centuries resulted in widespread shifts of the order of 0.2 °C in mean annual temperatures (Parker, 1994) and much larger changes at some individual stations (for example, Brunet et al., 2011; Ashcroft et al., 2012).

Specific issues relevant to automatic measurements of temperature include:

(a) Changes in instruments (from liquid-in-glass thermometers to electronic probes);

(b) Changes in instrument shelters;

(c) Changes in data processing algorithms (for example, time frequency of observations and definition of the daily mean);

(d) Changes in observation times or other observing practices;

(e) Changes in sites associated with the introduction of AWSs.

The actual change in instrumentation is sometimes thought of as the major change that occurs with the introduction of AWSs. However, in well-run observation networks, the measurement system itself is well calibrated (in the sense of spot readings being consistent with a laboratory standard) in accordance with standard metrological procedures (Bertiglia et al., 2015), and temperature inhomogeneities associated with the introduction of AWSs arise principally from other sources, as discussed below.

(a) Changes in instruments

The introduction of AWSs almost invariably involves a change of instrument, typically from a manually read liquid-in-glass thermometer (mercury or alcohol) to a platinum resistance thermometer or similar.

Best practice is for instruments to be calibrated and tested (WMO, 2014). In countries where this occurs, it is unusual for there to be any significant difference between instantaneous observations from a manual and an automatic instrument under controlled and stable laboratory conditions. However, even if instruments produce identical instantaneous readings, they may have different response times (the length of time they take to respond to an instantaneous change in the air temperature). Typically, unfiltered observations from an automatic probe will have a faster response time than a liquid-in-glass thermometer.

In some cases, once in the field, automatic probes can drift from a standard, or may do so during transport. Such drifts will introduce inhomogeneities into a climate record. Their prevention requires regular inspections of sites (at least every 6 months for AWSs; WMO, 2011) and tolerance checks and calibrations of instruments against a standard. Some countries lack the resources to carry out frequent inspections, making instrument drift a heightened risk.

(b) Changes in instrument shelters

Some countries (such as Australia and Canada) retained the same screen design when they introduced AWSs, and others (such as the United Kingdom of Great Britain and Northern Ireland) retained the same screen shape and dimensions but used plastic materials instead of wood. In the British case, Perry et al. (2007) found that temperature differences between plastic and wooden screens, on average, were less than 0.1 °C (and that plastic/wood differences were often less than the differences between co-located wooden screens). Retention of the same screen design is the preferred option where possible, as doing so removes a potential source of inhomogeneities.

However, many countries have introduced new screen designs – often changing to small plastic screens that are cheaper and easier to maintain. Some of these new screen designs produce outcomes that are similar (typically within 0.1 °C) to those obtained from conventional wood Stevenson screens (Brandsma and van der Meulen, 2008). But in other cases, such as the transition from the Cotton Region Shelters (CRSs) to various other screens in the United States of America, impacts on certain temperature variables are of the order of several tenths of a degree. It has been found in some cases that some new screen designs can degrade significantly with age, with effects on mean maximum temperature of up to 0.5 °C over five years (Lopardo et al., 2014). Such degradation, where it occurs, is difficult to detect statistically or through field inspection.

In some cases (for example, the Climate Reference Network in the United States; Diamond et al., 2013), ventilated shelters have also been introduced, but naturally aspirated sensors remain in use in most operational networks.

Further details of relevant comparisons are discussed in section 3.2.

(c) Changes in data processing algorithms

As noted in (a) above, the automatic probes used in AWSs normally have different response time characteristics to those of liquid-in-glass thermometers. Most commonly, the probes will have faster response times, which means that they are more capable of sampling very short-term fluctuations in temperature, resulting in higher maximum and lower minimum temperatures, than liquid-in-glass thermometers. This gives a positive bias in the diurnal temperature range.

Increased maximum temperatures and decreased minimum temperatures can cancel out to some extent in mean temperatures, in those countries where daily maximum and minimum temperatures are used as the basis for calculating mean temperatures. However, there are some locations (Figure 2), particularly in arid regions, where temperature fluctuations on very short timescales (less than 1 min) are stronger during the day than at night (due to the different structure of the boundary layer, which is often stable during the night and well mixed during the day). Hence, in such cases, more-rapid sampling will result in a positive shift in maximum temperatures that is larger than the negative shift in minimum temperatures, thus producing a shift in the mean temperatures.

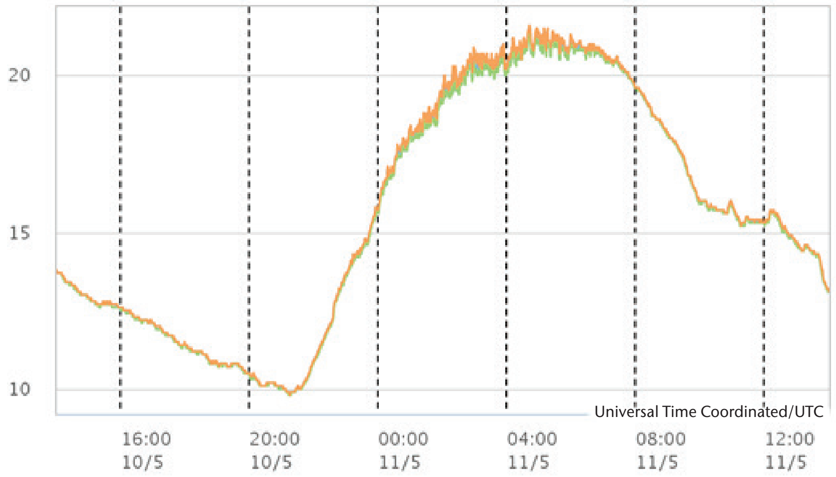

Figure 2. Temperatures (°C) at Birdsville, Australia, from 1400 UTC 10 May 2016 to 1400 UTC 11 May 2016 (0000 to 2400 on 11 May 2016 local time), showing discernible minute-to-minute fluctuations from around 0000 to 0700 UTC (1000 and 1700 local time) 11 May 2016 between the highest (orange line) and lowest (green) temperatures of each minute

Source: Australian Bureau of Meteorology

It is possible to use filters in data processing algorithms to smooth the outputs from automatic probes to response times that are characteristic of liquid-in-glass instruments. However, such processes cannot be assumed to give a precise match, and may not be in place in all circumstances. They are also unlikely to account for changes in the thermal properties of a screen in situations where automatic stations have a different screen design to that of conventional stations.

(d) Changes in observation times or other observing practices

Changes of observation time are a well-established source of inhomogeneities in temperature records (Menne et al., 2009; Vincent et al., 2009).[2] While the automation of networks does not necessarily lead to changes of observation times, it does potentially facilitate such changes. For example, the fact that a human observer is no longer required makes it much more practical to measure maximum and minimum temperatures for a calendar day, from midnight to midnight. As noted earlier, while it is preferable to avoid changes in observation time for daily variables when moving from conventional to automatic stations, many such changes have occurred historically.

The introduction of AWSs can result in changes in data precision. On the one hand, human observers, even if instructed to read to the nearest 0.1 °, have a tendency to use round numbers, resulting in values ending in .0 and, to a lesser extent, .5 being over-represented (Trewin, 2002). Automatic stations should be free of such tendencies, but in some cases, they have reported values rounded to whole degrees (for example, because of limitations in data transmission codes, particularly in early-generation AWSs) in countries where the normal standard of precision is

---

[2] While the case discussed by Menne et al. (2009) only partly relates to the introduction of AWSs, it illustrates that a change in observation time over a large part of a network can significantly bias values – in this case, a systematic shift from afternoon to morning observations at many stations resulted in a negative bias in mean temperatures.

0.1 °. Providing that any rounding tendency has no upward or downward bias, rounding should have a negligible impact on mean values but may have a noticeable impact on the frequency of exceedances of thresholds (for example, days where the temperature is 30 °C or higher) (Zhang et al., 2009; Trewin, 2012), and may also influence observed weather variability.

(e) Changes in sites associated with the introduction of automatic weather stations

In many cases, the introduction of automatic observations is accompanied by a site change. There are several reasons for this, but a common scenario is for a long-standing manual site located in an urban area (for historical reasons, or because of the availability of observers) to be moved to a new location in a less built-up area (often an airport or similar location). The new location would better meet standard observation specifications as defined in CIMO 2014, and may be suitable for observing variables such as wind that cannot be satisfactorily observed in most urban environments. Another common scenario is for a site to be moved to a location, such as the centre of an airport, that is not readily accessible by human observers but which is more representative for key user groups of the information obtained.

While every site change is unique in its exact impact upon temperature and other variables, it is well established that urban areas are typically warmer (especially at night) than non-urban areas. Hence, in the absence of other influences (such as local topography or exposure to a nearby coast-line), a site relocation from an in-town to an out-of-town location, something often associated with AWS installation, will frequently lead to an artificial decrease in minimum temperatures. Depending on the exact nature of the old and new sites, inhomogeneities of 1 °C or more in minimum temperatures, resulting from a move out of town, are not unusual. There may also be less-consistent impacts on maximum and daily mean temperatures.

2.2.2 ***Precipitation***

The most-common transition that occurs in precipitation measurements during the transition from conventional stations to AWSs is a change from manually read, accumulating gauges (where rainfall accumulates in a container that is read and then emptied at a set time) to automatic gauges. The most-common type of automatic precipitation gauge is a tipping bucket raingauge (where water enters the gauge through an intake and accumulates in a small bucket, which tips when full, generating a signal for transmission), while another common type of automatic gauge is a weighing gauge. Gauges that measure the water level in a container also exist, as do non-catching gauges using tools such as impact measurement, microwave radar and lasers. Non-catching gauges are more often used for present weather measurements than for long-term records of precipitation accumulation (Vuerich et al., 2009).

As with temperature, the impact of transition to automatic measurements can include the effects of the change in instrument and changes in the exposure or local site environment of the instrument. As with any change of instrument type, there can be systematic differences between the conventional and automatic instruments, the nature of which will depend on the specific instruments involved.

The major impact of siting occurs in locations where the wind is strong enough to cause undercatch of precipitation, because wind is a well-established source of this (Sieck et al., 2007). In these cases, a significant change in the local wind environment (which can occur as a result of even a modest site change) can significantly influence the extent to which undercatch occurs, and thus cause an inhomogeneity in the measured precipitation. This issue is especially pronounced in highly exposed locations, such as mountain tops or coast-lines (for example, at lighthouses). It is also especially pronounced at locations where a substantial proportion of the precipitation falls in frozen form. Siting is generally less important in light-wind conditions.

2.2.2.1 **Impact of outages on precipitation data**

While the operators of meteorological networks endeavour to minimize data loss, some level of outage is inevitable in any observing network. Outages are a particular issue for additive elements, of which precipitation is the most prominent. This is because the loss of any part of a single day of observations results in the loss of the monthly (annual) total for the month (year), unless the data for the missing period are estimated, for example, by use of surrounding stations. (In contrast, the loss of one day of temperature observations for a month normally only introduces a marginal uncertainty into the monthly mean value.) This problem is most acute for tipping bucket raingauges where a continuous count is required, and is less likely to affect weighing gauges depending on the way that data ingests are configured.

A period of missing precipitation data from an automatic station may result in the affected daily value being set to missing, or it may be recorded as a (possibly spurious) value of 0, depending on the way a database ingest is structured. In the latter case, this will create a negative bias in the recorded precipitation, the extent of which depends on the frequency of outages and the circumstances in which they occur. (It could be surmised that outages might be more likely to occur during storms, which are also likely to be times of high precipitation, although this proposition is not known to have been objectively tested.)

With conventional observations, if a day’s observations are missed, a multiday total can still (normally) be obtained, which may lead to the loss of daily observations but not normally those of the month and year. However, conventional observations can occasionally be affected by gauge overflows in very wet conditions, leading to extreme high rainfalls being underestimated or lost altogether – a failure mode that is much less likely for automatic instruments.

2.2.2.2 **Recording of small precipitation amounts**

Potential biases exist in the recording of small precipitation amounts using conventional and automatic instruments. This creates a potential for biases in the frequency of such small amounts when transitioning from conventional to automatic stations.

Experience with manual sites indicates that it is relatively common for small amounts (less than 2 mm, and especially less than 1 mm) not to be reported, especially at sites that do not have professional observers. A study of Australian manual rainfall data found that more than 50% of all daily rainfall amounts less than 1 mm across the network were not reported (Trewin, 2001). While the impact of this on monthly and annual totals is small (as the small amount will usually remain in the gauge and be added to the next, more-substantial rainfall), such under-reporting of small amounts will affect the number of observed rain days, and indices based on these.[3]

Conversely, automatic raingauges cannot reliably distinguish between rain and dew or frost, which can lead to an exaggerated number of days with small precipitation amounts unless there is manual intervention to remove them. Tipping bucket raingauges are also susceptible to spurious readings from tips registered as a result of a disturbance to the gauge, foreign objects entering the gauge[4] or similar.

Manual gauges are known to suffer from under-recording due to wetting losses for small precipitation amounts (this can also occur for some automatic systems). This issue is most significant for locations that receive a high number of very light precipitation falls, especially when these are in frozen form. In Canada, it has been estimated that the total wetting loss in some locations can be of the order of 15–20% (Goodison et al., 1998). Such a precipitation pattern is typical of very cold climates, where trace amounts can also contribute significantly to total precipitation.

---

[3] This phenomenon partly explains why the Expert Team on Climate Change Detection and Indices uses a rain-day threshold of 1 mm in its precipitation indices.

[4] A case of a spurious 0.2 mm reading at a site in northern Australia was observed in 2009, on a day with no cloud within several hundred kilometres. This was caused by flying debris from a passing lawnmower entering the gauge intake.

2.2.2.3 Frozen precipitation

The measurement of frozen precipitation has long presented a particular challenge. Undercatch of snow is a problem, even in light winds, as is distinguishing between falling snow and blowing snow that has been lifted from the ground. It is commonplace for different methods of measuring frozen precipitation to produce results that differ by a factor of 2 or more, with differences even larger at high wind speeds (Goodison et al., 1998; Wolff et al., 2014).

A major intercomparison study was carried out by WMO between 1986 and 1993, largely based on conventional observations (Goodison et al., 1998). A second intercomparison study, the WMO Solid Precipitation Intercomparison Experiment, made observations between 2012 and 2015.

Many of the issues with the measurement of frozen precipitation are common to conventional and automatic instruments. In particular, the impact of wind on frozen precipitation measurements (both in causing undercatch of falling snow and in causing blowing snow) will be a function of the position (height and exposure) and size of the gauge intake, and the extent to which it is shielded from wind. How the frozen precipitation is measured once it enters the gauge should not be an influence.

Tipping bucket raingauges are typically designed to measure liquids. A common instrument configuration is to heat the gauge intake to melt any frozen precipitation that enters it, which allows the gauge to measure liquid equivalents (with a small time lag). This can be an effective way of measuring frozen precipitation, although it depends on the proper functioning of the heating system. It will hence break down should that heating system fail,[5] and even when operating well, it may have systematic biases (for example, through evaporation at the gauge intake). Where such a heating system is not provided (most likely in places where frozen precipitation is relatively rare), snow may accumulate in the gauge and not be measured until it melts. This may result in precipitation being measured at a time well after that precipitation fell, even if the total amount was approximately correct and was not affected by, for example, the gauge intake becoming completely filled with snow. Evaporation losses from heated gauges can be significant, and Goodison et al. (1998) recommended that heated gauges not be used to measure frozen precipitation in locations where temperatures were below 0 °C for prolonged periods.

Weighing gauges, which can readily measure either liquid or frozen precipitation, generally do not rely on heating. Intercomparison studies have suggested that they are better suited to climates that receive substantial amounts of frozen precipitation, although they can also experience data quality issues (Goodison et al., 1998).

A CIMO survey (WMO, 2010) reported the results from nine countries. It found that at that time, 74% of operating automatic sites used heated tipping bucket raingauges to measure frozen precipitation, and 23% used weighing gauges.

2.2.3 *Atmospheric moisture (humidity, dewpoint and vapour pressure)*

Conventional observations of atmospheric moisture typically involve the use of dry and wet bulb thermometers. Readings from these are used to derive the current vapour pressure and saturation vapour pressure using the psychrometric method; these can then be used to calculate the dewpoint temperature and relative humidity.

Automatic stations typically employ one of two types of technology: a wet bulb probe, which is operated in a similar environment to a conventional wet bulb thermometer, or a relative humidity probe, which measures the change in capacitance of a thin film, a quantity dependent on the relative humidity.

---

[5] Operational experience is that gauge heating systems often suffer from reliability problems and have a high power demand.

Both types of instrument have potential systematic differences to measurements derived from conventional instruments (Lucas, 2010). In the case of methods involving the wet bulb temperature, the equation used to derive the vapour pressure from dry and wet bulb temperatures involves a quantity A, the psychrometric constant. The “real” value of A is a function of factors such as the ventilation of instruments, the shape of the wet bulb instrument, and the wick length and cleanliness. In practice, the value of A used operationally will typically be a constant across a national network, and hence a change in instrument type that affects the underlying “true” value of A without a change in the value used operationally will produce an inhomogeneity in vapour pressure and associated elements.

Humidity probes are most often deployed at locations with no regular observer presence, or where temperatures fall below freezing on a regular basis. They are designed to operate most effectively within a specified relative humidity range, and may be less reliable outside that range. This can lead to potential biases at extreme high or low humidities. Observed changes in the occurrence of extreme low humidity can have implications for applications such as fire weather, for which extreme low atmospheric moisture is a substantial risk factor. Observations of extreme high humidity are important for the identification of fog and mist, which are especially important for aviation and shipping.

Wet bulb probes require the associated reservoir of water to be periodically topped up (in the same way as wet bulb thermometers). If this reservoir dries up, the wet bulb probe will start to behave as a dry bulb, leading to spurious high dewpoint temperatures and humidity readings near 100%. This mode of failure can occur for either conventional or automatic wet bulb temperature instruments. However, at sites where the instruments are regularly checked (as would happen if an observer is making an observation), it is expected that it would be more likely that a depleted reservoir would be noticed before it became dry enough to affect observations.

The relationship between saturation vapour pressure and temperature (and hence between observed vapour pressure and dewpoint temperature) is highly non-linear. At dewpoint temperatures below 0 °C, a relatively small absolute change in vapour pressure (or in wet bulb temperature) can have a substantial impact on dewpoint temperature. For example, reducing the vapour pressure by 1 hPa will reduce the dewpoint temperature by 0.7 °C at a dewpoint temperature of 20 °C, by 2.4 °C at 0 °C and by 5.3 °C at –10 °C. The effect of this is that any biases between different instrument types will most commonly have their largest impacts on observations under conditions of extreme low humidity. An example of this is at Adelaide, Australia, where the mean number of days per year on which the dewpoint at 0900 local time was –5 °C or lower was 0.2 per year from 1978 to 1995, when a conventional wet bulb thermometer was used, but 1.5 per year from 1997 to 2015, when an automatic wet bulb probe was used. Meanwhile, at the nearby site of Parafield, where a humidity probe was in use from 1990 onward, the mean frequency was 0.5 per year.[6]

2.2.4 *Other elements*

Some elements that have traditionally been observed at many conventional stations: either cannot be measured at all with automatic instruments; are elements for which automatic observations are currently in their infancy; can be measured in some way with automatic instruments but not in a way that is compatible with conventional observations (even with homogeneity adjustments); or for which the use of automatic instruments is still largely experimental. These include: cloud amount, height and type; pan evaporation; present weather; visibility and snow depth.

Conversion from conventional to automatic measurements normally results in the cessation of observations for these elements. In countries where a substantial proportion of the network has been, or is in the process of being, automatic, this can lead to the observing network for those elements being reduced to a level below that required for viable national analyses. (For example, updating of the Australian total cloud amount dataset was suspended at the end of 2015 due to

---

[6] See section 3.4 for a detailed discussion of the conventional to automatic transition for Australian dewpoint measurements.

insufficient contributing stations; Jovanovic et al., 2010.) In some cases, there may be possible alternatives to conventional data. For example, there is potential to develop long-term cloud amount analyses by merging modern satellite data with older station-based observations, while solar radiation data, where available, may be used as a measure of sunshine duration.

The introduction of automatic observations will usually involve a change in instrument type for wind observations (or a change from winds estimated by an observer, using the Beaufort scale or similar, to instrumental measurements). Any change in instrument type introduces a potential inhomogeneity into the record, the nature of which will depend on the specific instruments involved and the specific algorithms used to convert an instrument output (for example, the rotation rate of anemometer cups) into the measurement of interest.

Another significant issue for wind observations is that the introduction of automatic observations also involves a site relocation. In many cases, this will be a relocation from an urban location to an airport or similar facility. Wind observations are highly sensitive to the presence of obstructions to airflow in the vicinity of the observing site. (CIMO 2014 recommends that wind instruments be placed in a location where the distance between the anemometer and any obstruction is at least 10 times the height of the obstruction.) Hence, relocating to a more open site – even without a change in instrument – would be expected to result in an increase in observed wind speed.

3. EXAMPLES OF DOCUMENTED INHOMOGENEITIES ARISING DUE TO THE TRANSITION FROM CONVENTIONAL TO AUTOMATIC OBSERVATIONS

3.1 Introduction

Many countries now have experience in making the transition from conventional to automatic observations. The results of these are widely documented, although not always in an accessible form (for example, many results from NMHS field trials are documented only in internal reports, which are often difficult for external users to access, although such material is becoming increasingly accessible online).

The purpose of this chapter is to provide some examples to indicate to users the approximate nature and type of differences between conventional and automatic observations for a range of variables, not to carry out an exhaustive review of the literature on inhomogenities.

The Parallel Observations Science Team (POST) of the International Surface Temperatures Initiative (Venema et al., 2016) is undertaking (at the time of writing of this note) an assessment of case studies, across a broad range of countries, of parallel measurements between conventional and automatic observations for temperature and precipitation. Preliminary findings from the work of POST are included below.

3.2 Temperature

Preliminary results from POST (Aguilar et al., 2015) from 10 countries have indicated that, in the majority of studies, the average bias in mean temperatures due to the transition from conventional to automatic observations was negative (that is, automatic observations were cooler). The mean biases ranged from +0.19 °C in Peru to –0.36 °C in Argentina, although substantially larger values occurred at some individual stations. The magnitudes of biases were considerably larger for the diurnal temperature range, as some countries had contrasting results for maximum and minimum temperatures. For example, in Spain and Sweden, positive biases were found for maximum temperature and negative biases for minimum temperature, while in the United States case (discussed further below), the average bias found was –0.50 °C for maximum temperature but it was near to zero for minimum temperature. It should be noted that the reported results do not systematically separate the instrument change from associated site relocations or changes in instrument shelters; in some countries, these have a systematic impact of their own (for example, in Australia, some automatic weather installations were also associated with site moves from locations in towns to locations outside towns, which would be expected to have a cool bias).

One closely studied transition from conventional to automatic instruments occurred in the United States, where, between 1984 and 1988, about 60% of the cooperative station temperature network transitioned from conventional instruments (in a box-shaped wooden screen) to automatic instruments (in a multiplate screen).[7] Initial analyses based on comparisons of operational stations (Quayle et al., 1991) found an average impact of this transition of −0.4 °C for the maximum temperature and +0.3 °C for the minimum temperature. Dedicated field comparisons at an experimental site (Wendland and Armstrong, 1993; Doesken, 2005) supported this conclusion for maximum temperature, and also found that the differences were largest under conditions of light winds and strong solar radiation. Little evidence of any significant difference for minimum temperature was found, suggesting that the minimum temperature differences observed by Quayle et al. (1991) were possibly an artefact of site changes associated with the transition (for example, establishing automatic sites closer to buildings than the conventional observations they replaced, to reduce the amount of cabling required). Hubbard and Lin (2006) reinforced this by finding that, while the overall impacts across the United States network given by Quayle et al. (1991) were sound, the impacts at individual locations were highly site specific.

Studies involving other parts of the United States network (the automated surface observing system (ASOS) used at major airports, and the climate reference network (CRN), both of which used different screen designs to MMTS) also found differences of the order of a few tenths of a degree between MMTS on the one hand, and ASOS and CRN on the other (Hubbard et al., 2004; Sun et al., 2005). However, Guttman and Baker (1996) found that differences between ASOS and other systems were site specific, even when both systems were installed within the boundaries of the same airport. Leeper et al. (2015) found that the aspirated screen used in CRN had lower maximum temperatures (mean of −0.48 °C) and higher minimum temperatures (+0.36 °C) than MMTS. When assessing the transition from conventional to automatic temperature measurements, these results reinforced the challenges of separating out impacts due to the sensors themselves, those relating to screen type and those relating to other associated changes such as site moves.

Overlapping observations from conventional and new Campbell Scientific automatic stations with standard configurations were examined at 22 locations across Canada for the purpose of preserving the continuity of long-term climate records. First, the observing windows for the daily maximum and minimum temperatures were matched at some of the station pairs. Second, biases were computed and used to adjust the time series of the automatic stations. Two- and five-year overlapping periods were used to analyse the biases. The mean absolute difference between the station pairs, which ranged up to 0.7 °C and 1.4 °C for the maximum and minimum temperatures, respectively, was found (Milewska and Vincent, 2016).

3.3 **Precipitation**

Preliminary results from POST (Stepanek et al., 2015) from nine countries indicated that in the majority of studies, automatic observations recorded lower precipitation than conventional observations. The average magnitude of the dry bias was small (less than 5% in most of the countries studied), but with widely varying responses among different stations, with substantial numbers of stations showing biases greater than +20% or less than −20%. The preliminary findings indicated that biases were greater for frozen than for liquid precipitation, and greater for early-generation AWSs than for more-recently developed instruments. As for temperature, the reported results also incorporated the effect of associated site relocations, and much of the

---

[7] These are widely referred to as the Cotton Region Shelter (CRS) system and the maximum minimum temperature system (MMTS), respectively. The MMTS is not fully automatic in the sense that it transmits digital data to a terminal (typically inside the observer’s premises), but manual intervention is required to transmit the data to the broader communications network.

spread in the POST findings is likely to arise from those site relocations rather than the instrument change (especially for frozen precipitation, given the influence of wind speed on undercatch, and the sensitivity of wind speed to local site environment).

While conventional measurements were not part of the comparison, there have been WMO intercomparison projects involving a wide variety of precipitation gauges, both in the laboratory (Lanza et al., 2006; Sevruk et al., 2009) and in the field (Vuerich et al., 2009). These found that uncorrected tipping bucket raingauges typically had large negative biases, relative to a reference, at very high rainfall intensities (typically 15–20% at rainfall rates of 300 mm/h), although biases were much smaller at lower intensities (typically less than 5% at rates of 50 mm/h or lower). These results indicated that uncorrected tipping bucket raingauges are likely to be seriously affected by long-term biases in climates where a substantial proportion of annual precipitation falls in high-intensity events, as is typical in many parts of the tropics and subtropics. The study by Vuerich et al. (2009) found generally good results for tipping bucket gauges with automatic correction factors applied depending on rainfall intensity. Solid precipitation was not part of this intercomparison.

3.4 Other elements

Most published studies of the transition from conventional to automatic measurements have focused on temperature and precipitation, with fewer studies of other variables.

Lucas (2010) and Gorman (2003) reported on comparisons of conventional and automatic dewpoint/humidity measurements in Australia (Lucas through a regression analysis across the conventional and automatic components of the network, and Gorman through a field trial). Lucas (2010) found that automatic wet bulb probes in the Australian network had a mean dewpoint bias of –0.5 °C and humidity probes had a mean dewpoint bias of –0.3 °C (although it was somewhat site dependent in the latter case) relative to conventional wet bulb instruments. Gorman (2003) found large differences between dewpoints derived from wet bulb and humidity probes at low dewpoints (typically 3 °C or more at dewpoints below –5 °C) and attributed these to an unsuitable value of the psychometric coefficient A (see section 2.2.3). These results are specific to the instrumentation and observation practices used in Australia and may not necessarily apply to other instruments and networks.

4. GUIDANCE FOR MANAGING THE TRANSITION FROM CONVENTIONAL TO AUTOMATIC MEASUREMENTS

4.1 Parallel measurements between conventional and automatic systems

In the climate context, the most significant challenge presented by the transition from conventional to automatic measurements is the homogeneity of the data, across a wide range of variables, and with respect to mean values and to extremes. As discussed earlier, and for specific elements in section 2.2, the introduction of AWSs has the potential to introduce inhomogeneities into any climatological time series, as do the site relocations that often accompany the transition.

Recommended practice for the transition from conventional stations to AWSs is to carry out a period of parallel observations between the two systems. Guidance provided by WMO for the optimal period of parallel observations is not fully consistent. A minimum of 12 months for wind speed and direction, 24 months for temperature, humidity, sunshine and evaporation, and 60 months for precipitation is suggested by CIMO 2014 (which also notes that “A useful compromise would be an overlap period of 24 months”). The Guide to Climatological Practices (WMO, 2011) recommends an overlapping period of at least one year, and preferably two or more years. In practice, the necessary period is likely to be site dependent and is not necessarily predictable ahead of time. Twenty-four months may be sufficient to determine whether a significant inhomogeneity exists or not, but it may not be sufficient to fully quantify an inhomogeneity (including the impact on extremes of the distribution, as well as the mean). The risk of a significant inhomogeneity is greater the larger the system change; if the introduction of an AWS is in the same instrument enclosure as that of the conventional observations, the risk of a large inhomogeneity is low for many variables. However, if there is also a substantial site relocation involved, then the risk of a large inhomogeneity increases. It is therefore preferable to have the option of extending a parallel observations period should analysis of the initial period of parallel observations reveal substantial differences.

Much of the discussion in this chapter relates to the use of parallel observations in general, and is applicable to the use of parallel observations in scenarios other than a transition from conventional to automatic observations (such as a site relocation with no change in instrument type).

Detailed guidance on detecting and adjusting for inhomogeneities in climate data is provided in the Guidelines on Climate Metadata and Homogenization (WMO, 2003) and is outside the scope of this note.

4.2 Testing prior to deployment of operational automatic weather stations

The CIMO 2014 guide recommends that new instruments should be tested prior to their operational deployment at AWSs. This incorporates environmental tests and calibration in a laboratory setting, and functional testing in the field. Such testing will determine how closely sensors match a standard (if one exists), as well as the performance of instruments in the field with respect to a reference.

Field testing is typically carried out at one or more dedicated experimental site(s). Where field testing is being carried out to support the introduction of AWSs, the existing instruments in use for conventional observations should be included in the field testing so that the instruments can be compared in a consistent local environment. A field test programme should encompass at least one full annual cycle, and should seek to isolate individual components of a change (for example, by comparing a conventional and automatic sensor within the same screen).

Conducting environmental and field testing does not eliminate the need for parallel observations at individual locations where AWSs are deployed in place of conventional stations, as inhomogeneities can be site specific depending on local conditions. However, if both environmental and field testing demonstrate that the difference between the two in an experimental setting is negligible, this reduces the risk that there will be significant differences at individual sites (providing no major site change is involved). If, as is often the case, AWSs have been deployed across a substantial proportion of the network within a short space of time, the absence of systematic differences between automatic and conventional systems also reduces the risk of a bias across a network (see section 4.4).

4.3 Representativeness of a parallel observations period

A period of parallel observations between two observations systems will be a reliable indication of the differences between those systems if, and only if, the period of parallel observations is representative of those periods before and after (that is, the old station during the period of parallel observations is consistent with the old station before the period of parallel observations started, and the new station during the parallel period is representative of the new station after the period finishes). It is therefore important to test the individual site time series for inhomogeneities during, or shortly before/after, the period of parallel observations. It is also important to ensure that the parallel observations are carried out properly; there have been cases where the “conventional” observations during a comparison were actually read from the automatic instruments, making the comparison of limited use.

Inhomogeneities during a period of parallel observations can occur for many reasons. A common scenario is that the decision is made to move a station because the exposure of the observing site has become poor, or is in imminent danger of doing so, due to urban development. If that development occurs during the period of parallel observations, it is likely that it will render the old site unrepresentative of its former self, and hence the parallel observations period will be unrepresentative.

Even in the absence of any long-term inhomogeneities, the relationships among meteorological variables at parallel observing sites may be subject to interannual variability. For example, soil moisture and the greenness of nearby vegetation are more likely to influence temperatures at a rural site than they are at a site in an urban environment, and hence the rural/urban temperature difference in an unusually wet (or dry) year may not be representative of normal conditions. A parallel observations period of several years minimizes the risk that that the period samples an unusual climatic phase, especially in climates that are strongly influenced by the El Niño Southern Oscillation or are otherwise subject to high interannual variability.

If the period of parallel observations is long enough, inhomogeneities during a period of parallel observations can be dealt with by not using the unrepresentative proportion of the parallel observations period. An example is shown in Figure 3. In this case, there were parallel observations between the new and old sites from 1992 to 2002. However, there were inhomogeneities, of about –0.9 °C at the new site in 1995 (due to a change in the data smoothing algorithm with a new version of the AWS), and of about +0.5 °C at the old site in 1999 (due to the site becoming increasingly enclosed by new buildings). Hence, only the period from 1995 to 1998 was used to assess the long-term difference between the two sites. Depending on the length of period of parallel observations and the timing of any inhomogeneities, such an approach may not be possible. In this case, better results may be obtained by ignoring the period of parallel observations and instead assessing potential inhomogeneities by other means, such as independent reference stations in the region (see below).

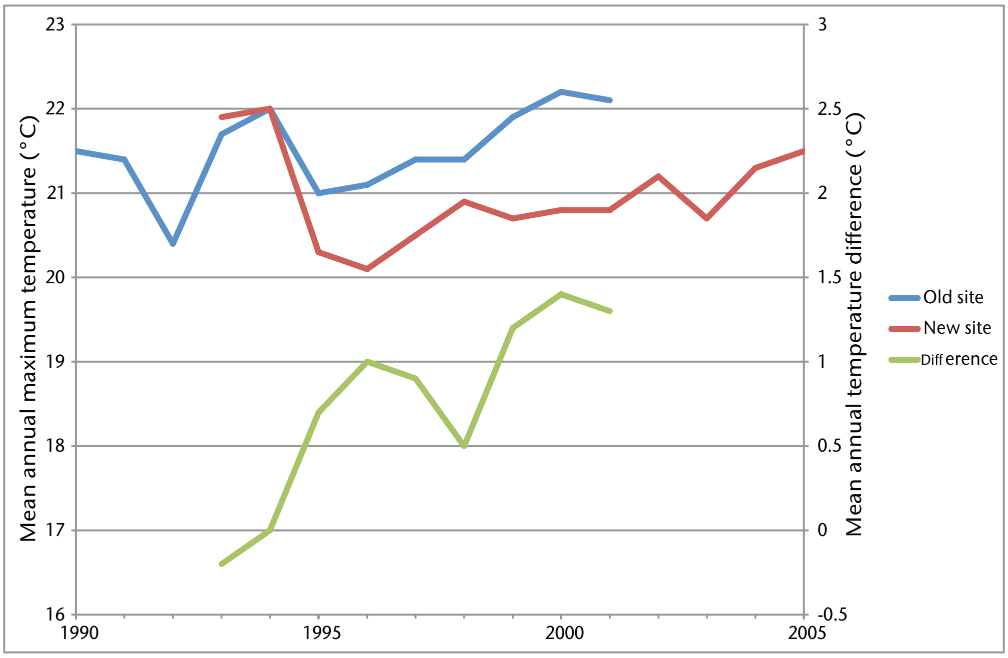

Figure 3. Mean annual maximum temperatures at Port Lincoln, Australia, around a period of parallel observations between a conventional site in the town centre ("old site") and an automatic site at the airport, 14 km to the north ("new site")

4.4 What happens if there is no useful parallel observations period?

While planning for a transition from conventional to automatic observations should include provision for field testing, and parallel observations during the deployment phase, there are many situations where no useful period of parallel observations exists. Some of the reasons for this may include:

(a) The transition took place some time ago and there was no policy to carry out parallel observations at that time;

(b) Parallel observations were carried out but covered only a limited range of variables;

(c) The parallel observations period was not useful because of inhomogeneities in the record of either the conventional or automatic station during or around that period;

(d) A planned parallel observations period was terminated early because of the unavailability of the old site (for example, if a site move occurred because of development pressures at an old site, a landowner may be unwilling to continue to lease the land required for the observations).

In any of these cases, assessing the impact of the transition from conventional to automatic measurements becomes a problem of assessing the size of an inhomogeneity, with metadata support (assuming that the date of installation of the automatic station is known). This process typically makes use of other stations (reference stations) nearby.

Handling situations where changes occur across a large part of a network at a similar time

The homogenization of climate data is particularly challenging when changes occur across a large part of a national network at a similar time. This is because the most-accurate homogenization methods rely on the use of data from reference stations (stations that are well correlated with the candidate station, and whose data in the few years before and after the inhomogeneity can be used to detect the inhomogeneity and assess its impact). However, if much of a network is affected, many potential reference stations will also be affected and hence will not provide a true indication of the inhomogeneity's impact. A further issue is that an inhomogeneity that affects a large part of a national network may be large enough to be significant on a national scale even though it is marginal at an individual station. For example, a 0.2 °C inhomogeneity in temperature is likely to be undetectable at an individual station, but it can be significant on a national scale (in the context of century-scale trends of the order of 0.1 °C per decade).

This problem is most acute when changes are simultaneous, as may occur, for example, with a national change in observation time. The change of observation time in Canada in 1961 from a day ending at 0000 UTC to one ending at 0600 UTC resulted in an inhomogeneity of −0.6 °C to −0.8 °C in minimum temperature at many stations in eastern Canada (Vincent et al., 2002). Such a simultaneous change is rarer for the transition from conventional to automatic networks as it would not be expected that all the AWSs in a network would be made operational on the same day.[8] However, it is a common scenario for a large number of stations to make the transition over a period of a few years, which is a sufficiently short time window to produce a substantial effect on homogenization practices.

If no suitable parallel observations exist, potential strategies for dealing with this problem include:

(a) Use as reference stations only those stations that are unaffected by the change – in this context, stations that retained conventional observations throughout, or possibly automatic stations that were installed some years before broader network-wide changes.

(b) If no such reference stations exist (for example, because the entire network was affected), use an independent dataset that has some relationship with the variable of interest. In the case of land surface temperatures, this might include observations in neighbouring countries, upper-air temperatures (for example, 850 hPa temperatures) or sea surface temperatures in nearby oceans, while for wind measurements (in middle and higher

---

[8] However, there are cases where AWSs become the principal instrument simultaneously; for example, at those Australian principal observing sites that had both conventional and automatic observations, the automatic observations became the primary readings for most variables on 1 November 1996.

latitudes), geostrophic winds derived from pressure data may be a useful proxy. Such datasets are unlikely to be sufficiently well correlated to produce useful results for a specific station, but may be useful for assessments of impacts at national or regional scales.

(c) Establish a field trial retrospectively (or make use of the results of an existing field trial) in which currently operational automatic stations are compared with a replica of the conventional instruments that were formerly in use, to identify and quantify any differences in performance that can then be used/extrapolated to the same AWS types used at other locations. This approach has been followed in a number of studies seeking to assess the characteristics of instrument shelters used for temperature measurements in the nineteenth century and earlier (Brunet et al., 2006, 2011; Böhm et al., 2010).

Where such network-wide changes have occurred, it is considered best practice to assess, and adjust for, the overall effect of these changes before attempting to carry out homogenization at the individual site level (Milewska and Vincent, 2016; Vincent et al., 2017).

4.5 Data management during transition from conventional to automatic observations

The transition from conventional to automatic observations presents a number of data management challenges.

It is recommended that, when a station is automatic, it is assigned a new identifier, to make the change as transparent as possible to data users. (If parallel observations are carried out as recommended, a new identifier will be needed anyway, as the conventional and automatic stations will be operating simultaneously for a period of time, unless a climate database is structured so that it can accept data from multiple sensors at the same location.)

It is recognized that a change of identifier will create issues for certain applications. For example, the lack of a sufficient averaging period may prevent the calculation of climate normal for an automatic station with a new identifier, while to obtain long-term records, it will be necessary to merge data from two or more separately identified stations in the same general location. Best practice in these cases is to develop a long-term homogenized dataset for the location in question. However, if differences between the conventional and automatic stations have been found to be small, it may be possible to merge the series without adjustment for certain applications (for example, public information), although fully homogenized data should always be used for applications relating to long-term climate change.

5. REFERENCES

Aguilar, E., P. Stepanek, V. Venema, R. Auchmann, F.D. dos Santos Silva, E. Engström, A. Gilabert, Z. Kretova, J.A. Lopez-Díaz, Y.L. Rico, C.O. Rojas, M. Prohom, D. Rasilla, M. Salvador, G. Vetacnik, Y. Yosefi and M. de los Milagros Skansi, 2015: *Biases Found in Temperature Records by the Transition from Conventional to Automatic Measurements in European and American Parallel datasets*. EUMETNET Data Management Workshop held at St. Gallen, Switzerland, 28–30 October 2015.

Ashcroft, L., D. Karoly and J. Gergis, 2012: Temperature variations of south-eastern Australia, 1860–2011. *Australian Meteorological Oceanographic Journal*, 62:227–245.

Bertiglia, F., G. Lopardo, A. Merlone, G. Roggero, D. Cat Berro, L. Mercalli, A. Gilabert and M. Brunet, 2015: Traceability of ground-based air-temperature measurements: a case study on the meteorological observatory of Moncalieri (Italy). *International Journal of Thermophysics*, 36:589–601, doi:10.1007/s10765-014-1806-y.

Böhm, R., P.D. Jones, J. Hiebl, D. Frank, M. Brunetti and M. Maugeri, 2010: The early instrumental warm-bias: a solution for long central European temperature series 1760-2007. *Climatic Change*, 101:41–67.

Brandsma, T. and J.P. van der Meulen, 2008: Thermometer screen intercomparison in De Bilt (the Netherlands) – Part II: description and modelling of mean temperature differences and extremes. *International Journal of Climatology*, 28:389–400.

Brunet, M., J. Asin, J. Sigró, M. Bañón, F. García, E. Aguilar, J.E. Palenzuela, T.C. Peterson and P. Jones: 2011. The minimisation of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis. *International Journal of Climatology*, 31:1879–1895, doi:10.1002/joc.2192.

Brunet, M., O. Saladié, P. Jones, J. Sigró, E. Aguilar, A. Moberg, D. Lister, A. Walther, D. Lopeza and C. Almarzae, 2006: The development of a new dataset of Spanish daily adjusted temperature series (SDATS) (1850–2003). *International Journal of Climatology*, 26:1777–1802.

Claussnitzer, A., F. Maier and R. Spengler, 2015: *QualiMET2.0, the New Quality Control System of Deutscher Wetterdienst*. EUMETNET Data Management Workshop held at St. Gallen, Switzerland, 28–30 October 2015.

Diamond, H.J., T.R. Karl, M.A. Palecki, C.B. Baker, J.E. Bell, R.D. Leeper, D.R. Easterling, J.H. Lawrimore, T.P. Meyers, M.R. Helfert, G. Goodge and P.W. Thorne, 2013: U.S. Climate Reference Network after one decade of operations: status and assessment. *Bulletin of the American Meteorological Society*, 94:424–448, doi:10.1175/BAMS-D-12-00170.1.

Doesken, N.J., 2005: *The National Weather Service MMTS (Maximum-Minimum Temperature System) – 20 Years After*. Thirteenth Symposium on Meteorological Observations and Instrumentation held at San Diego, 9–13 January 2005.

Goodison, B.E., P.Y.T. Louie and D. Yang, 1998: *WMO Solid Precipitation Measurement Intercomparison* (WMO/TD-No. 872). Geneva, World Meteorological Organization.

Gorman, J., 2003: *AWS Determination of Dew Point in the Field*. Technical Note 2003-0001. Melbourne, Bureau of Meteorology.

Guttman, N.B. and C.B. Baker, 1996: Exploratory analysis of the difference between temperature observations recorded by ASOS and conventional methods. *Bulletin of the American Meteorological Society*, 77:2865–2873.

Hubbard, K.G. and X. Lin, 2006: Re-examination of instrument change effects in the U.S. Historical Climatology Network. *Geophysical Research Letters*, 33:L15710, doi:10.1029/2006GL027069.

Hubbard, K.G., X. Lin, C.B. Baker and B. Sun, 2004: Air temperature comparison between the MMTS and the USCRN temperature systems. *Journal of Atmospheric and Oceanic Technology*, 21:1590–1597.

Jovanovic, B., D. Collins, K. Braganza, D. Jakob and D.A. Jones, 2010: A high-quality monthly total cloud amount dataset for Australia. *Climatic Change*, 108:485–517.

Lanza, L., M. Leroy, C. Alexandropoulos, L. Stagi and W. Wauben, 2006: *WMO Laboratory Intercomparison of Rainfall Intensity Gauges* (WMO/TD-No. 1304). Geneva, World Meteorological Organization.

Leeper, R.D., J. Rennie and M.A. Palecki, 2015: Observational perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: temperature and precipitation comparison. *Journal of Atmospheric and Oceanic Technology*, 32:703–721, doi:10.1175/JTECH-D-14-00172.1.

Lopardo, G., F. Bertiglia, S. Curci, G. Roggero and A. Merlone, 2014: Comparative analysis of the influence of solar radiation screen ageing on temperature measurements by means of weather stations. *International Journal of Climatology*, 34:1297–1310, doi:10.1002/joc.3765.

Lucas, C., 2010: *A High-quality Historical Humidity Database for Australia*. CAWCR Technical Report 24. Melbourne, Centre for Australian Weather and Climate Research.

Menne, M.J., C.N. Williams and R.S. Vose, 2009: The U.S. Historical Climatology Network monthly temperature data, version 2. *Bulletin of the American Meteorological Society*, 90:993–1007.

Milewska, E.J. and L.A. Vincent, 2016: Preserving continuity of long-term daily maximum and minimum temperature observations with automation of reference climate stations using overlapping data and meteorological conditions. *Atmosphere-Ocean*, 54(1):32–47, doi:10.1080/07055900.2015.1135784.

Page, C.M., N. Nicholls, N. Plummer, B. Trewin, M. Manton, L. Alexander, L.E. Chambers, Y. Choi, D.A. Collins, A. Gosai, P. Della-Marta, M.R. Haylock, K. Inape, V. Laurent, L. Maitrepierre, E.E.P. Makmur, H. Nakamigawa, N. Ouprasitwong, S. Mcgree, J. Pahalad, M.J. Salinger, L. Tibig, T.D. Tran, K. Vediapan and P. Zhai, 2004: Data rescue in the southeast Asia and south Pacific region: challenges and opportunities. *Bulletin of the American Meteorological Society*, 85:1483–1489.

Parker, D.E., 1994: Effects of changing exposure of thermometers at land stations. *International Journal of Climatology*, 14:1–31, doi:10.1002/joc.3370140102.

Perry, M.C., M.J. Prior and D.E. Parker, 2007: An assessment of the suitability of a plastic thermometer screen for climatic data collection. *International Journal of Climatology*, 27:267–276.

Quayle, R.G., D.R. Easterling, T.R. Karl and P.Y. Hughes, 1991: Effects of recent thermometer changes in the cooperative station network. Bulletin of the American Meteorological Society, 72:1718–1723.

Sevruk, B., M. Ondras and B. Chvila, 2009: The WMO precipitation measurement intercomparisons. *Atmospheric Research*, 92:376–380.

Sieck, L.C., S.J. Burges and M. Steiner, 2007. Challenges in obtaining reliable measurements of point rainfall. *Water Resources Research*, 43:W01420.

Stepanek, P., E. Aguilar, V. Venema, R. Auchmann, F.D. dos Santos Silva, E. Engström, A. Gilabert, Z. Kretova, J.A. Lopez-Díaz, Y.L. Rico, C.O. Rojas, M. Prohom, D. Rasilla, M. Salvador, G. Vetacnik, Y. Yosefi and M. de los Milagros Skansi, 2015: *Biases in Precipitation Records Found in Parallel Measurements*. EUMETNET Data Management Workshop held at St. Gallen, Switzerland, 28–30 October 2015.

Sun, B., C.B. Baker, T.R. Karl and M.D. Gifford, 2005: A comparative study of ASOS and USCRN temperature measurements. *Journal of Atmospheric and Oceanic Technology*, 22:679–686.

Trewin, B.C., 2001: *The Under-reporting of Small Daily Rainfall Amounts*. Australia-New Zealand Climate Forum held at Darwin, Australia, 18–21 September 2001.

Trewin, B.C., 2002: Extreme temperature events in Australia. PhD Thesis, *School of Earth Sciences, University of Melbourne, Australia*.

Trewin, B.C., 2012: *Techniques Involved in Developing the Australian Climate Observations Network – Surface Air Temperature (ACORN-SAT) Dataset*. CAWCR Technical Report 49. Melbourne, Centre for Australian Weather and Climate Research.

Venema, V., R. Auchmann, E. Aguilar, I. Auer, C. Azorin-Molina, T. Brandsma, M. Brunetti, M. Dienst, P. Domonkos, A. Gilabert, J. Lindén, E. Milewska, O. Nordli, M. Prohom, J. Rennie, P. Stepanek, B. Trewin, L. Vincent, K. Willett and M. Wolff, 2016. *A Global Database with Parallel Measurements to Study Non-climatic Changes*. European Geosciences Union General Assembly held at Vienna, Austria, 17–22 April 2016.

Vincent, L.A., E.J. Milewska, R. Hopkinson and L. Malone, 2009: Bias in minimum temperature introduced by a redefinition of the climatological day at the Canadian synoptic stations. *Journal of Applied Meteorology and Climatology*, 48:2160–2168, doi:10.1175/2009JAMC2191.1.

Vincent, L.A., E.J. Milewska, X.L. Wang and M.M. Hartwell, 2017: Uncertainty in homogenized daily temperatures and derived indices of extremes illustrated using parallel observations in Canada. *International Journal of Climatology*, doi:10.1002/joc.5203.

Vincent, L.A., X. Zhang, B.R. Bonsal and W.D. Hogg, 2002: Homogenization of daily temperatures over Canada. *Journal of Climate*, 15:1322–1334.

Vuerich, E., C. Monesi, L.G. Lanza, L. Stagi and E. Lanzinger, 2009: *WMO Field Intercomparison of Rainfall Intensity Gauges* (WMO/TD-No. 1504). Geneva, World Meteorological Organization.

Wendland, W.M. and W. Armstrong, 1993: Comparison of maximum-minimum resistance and liquid-in-glass thermometer records. *Journal of Atmospheric and Oceanic Technology*, 10:233–237.

Wolff, M.A., K. Iskasen, K. Ødemark, R. Brækken and A. Petersen-Øverleir, 2014: *How Much Snow is not being Measured*? WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation held at St. Petersburg, Russian Federation, 7–9 July 2014.

World Meteorological Organization, 2003: *Guidelines on Climate Metadata and Homogenization* (WMO/TD-No. 1186). Geneva.

———, 2007: *Guidelines for Managing Changes in Climate Observation Programmes* (WMO/TD-No. 1378). Geneva.

———, 2010. *CIMO Survey on National Summaries of Methods and Instruments for Solid Precipitation Measurement at Automatic Weather Stations* (WMO/TD-No. 1544). Geneva.

———, 2011: *Guide to Climatological Practices* (WMO-No. 100). Geneva.

———, 2014. *Guide to Meteorological Instruments and Methods of Observation* (WMO-No. 8). Geneva.

Zhang, X., F.W. Zwiers and G. Hegerl, 2009: The influences of data precision on the calculation of temperature percentile indices. *International Journal of Climatology*, 29:321–327.

For more information, please contact:

World Meteorological Organization

7 bis, avenue de la Paix – P.O. Box 2300 – CH 1211 Geneva 2 – Switzerland

Communication and Public Affairs Office

Tel.: +41 (0) 22 730 83 14/15 – Fax: +41 (0) 22 730 80 27

E-mail: cpa@wmo.int

public.wmo.int

JN 171141