Temperature extreme records: World Meteorological Organization metrological and meteorological evaluation of the 54.0°C observations in Mitribah, Kuwait and Turbat,

10 September 2024

International Journal of Climatology

World Meteorological Organization Evaluation and Calibration Testing of 2016/17 temperatures of 54.0 °C recorded in Mitribah, Kuwait and Turbat, Pakistan as Record Temperature Extremes

Journal:	*International Journal of Climatology*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Merlone, Andrea; Istituto Nazionale di Ricerca Metrologica INRiM, Thermodynamics; Istituto di Scienze dell'Atmosfera e del Clima Consiglio Nazionale delle Ricerche Sede di Roma, Al-Dashti, Hassan; Directorate General of Civil Aviation, Meteorology Faisal, Nadeem; Pakistan Meteorological Department, Meteorology Cerveny, Randall; Arizona State University, School of Geographical Sciences AlSarmi, Said; Oxford University, Climate Research Lab, Oxford University Centre for the Environment; Ministry of Transport and Communication, Meteorology Brunet, Manola; University Rovira i Virgili, Geography Bessemoulin, PIerre; Meteo France, Climatology Driouech, Fatima; Universite Mohammed VI Polytechnique, Meteorology Khalatyan, Yelena; Armstatehydromet, Climate Peterson, Thomas; unaffiliated, WMO Commission for Climatology Rahimzadeh, Fatima; Atmospheric Science and Meteorological Research Center, Meteorology Trewin, Blair; Australian Bureau of Meteorology, Climatology Wahab, Magdy; Faculty of Science, Cairo University, Astronomy, Space Sciences and Meteorology Department Coppa, Graziano; Istituto Nazionale di Ricerca Metrologica (INRiM), Metrology for the Quality of Life Smorgon, Denis; Istituto Nazionale di Ricerca Metrologica, Metrology Musacchio, Chiara; Istituto Nazionale di Ricerca Metrologica, Temperature group Thermal metrology for Environment – Quality of life division Krahenbuhl, Daniel; Arizona State University, School of Geographica Sciences and Urban Planning
Keywords:	Observational data analysis < 1. Tools and methods, Observations < 1. Tools and methods, Surface-based observations < 1. Tools and methods
Country Keywords:	Pakistan, Kuwait

World Meteorological Organization Evaluation and Calibration Testing of 2016/17 temperatures of 54.0 °C recorded in Mitribah, Kuwait and Turbat, Pakistan as Record Temperature Extremes

Andrea Merlone 1

Hassan Al-Dashti 2

Nadeem Faisal 3

Randall S. Cerveny 4*

Said AlSarmi 5

Pierre Bessemoulin 6

Manola Brunet 7, 10, 15

Fatima Driouech 8

Yelena Khalatyan 9

Thomas C. Peterson 10

Fatima Rahimzadeh 11

Blair Trewin 12

M.M. Abdel Wahab 13

Serpil Yagan 14

Graziano Coppa 1

Denis Smorgon 1

Chiara Musacchio 1

Daniel Krahenbuhl 4

1. Istituto Nazionale di Ricerca Metrologica (INRIM), Italy
2. Meteorology Department, Directorate General of Civil Aviation, Kuwait
3. Pakistan Meteorological Department, Karachi, Pakistan
4. Arizona State University, Tempe AZ USA
5. Ryiadh-Saudi Arabia
6. Metéo France
7. University Rovira i Virgili, Tarragona Spain
8. Mohammed VI Polytechnic University, Morocco
9. Climate Research Division, Armenian Hydromet Service, Armenia
10. WMO Commission for Climatology
11. Atmospheric Science and Meteorological Research Center, Tehran, Iran
12. Bureau of Meteorology, Australia
13. Cairo University, Egypt
14. Meteroloji Genel Müdürlüğü, Araştırma Dairesi Bşk., Klimatoloji Şb. Md., Keçiören/Anakara, Turkey
15. University of East Anglia, Norwich, UK

*corresponding author

Abstract:

A World Meteorological Organization (WMO) committee officially evaluated temperature record extremes of 54.0 °C at two locations, one in Mitribah, Kuwait on 21 July 2016 and a second in Turbat, Pakistan on 28 May 2017. The committee agreed that quantity and quality of documentation of both observations were excellent. Additional metrological testing of the equipment focused on three aspects: the calibration of both thermometers, an effort to estimate the factors influencing the measurements and a direct comparison of the two thermometers when exposed simultaneously to 54 °C. The metrological analysis's conclusion for the Mitribah value is a temperature estimated to be 53.87 °C with an expanded uncertainty of ± 0.08 °C. Correspondingly, for the Turbat value the temperature is estimated to be 53.72 °C with an expanded uncertainty of ± 0.40 °C. Following that analysis, the committee recommended acceptance of the calibrated observations to the first decimal digit such that the Mitribah observation is accepted as 53.9 °C ± 0.1 °C and the Turbat as 53.7 °C ± 0.4 °C. The Mitribah, Kuwait temperature is now accepted by the WMO as the highest temperature ever recorded for Asia (WMO RA II) and the two observations are the third (tied within uncertainty limits) and fourth highest WMO-recognized temperature extremes and, significantly, they are the highest, officially-recognized temperatures recorded in the last 76 years. This evaluation has involved the most extensive temperature extremes analysis ever been undertaken by an international evaluation committee of the WMO CCl Archive of Weather and Climate Extremes.

Key words: **Temperature Extreme, Middle East, Metrology, Calibration, Uncertainty**

1. Introduction

The World Meteorological Organization (WMO)'s Commission for Climatology (CCl) has created an online archive of officially recognized weather and climate extremes (e.g., highest recorded global temperature, strongest wind speed, most deadly tropical cyclone). That WMO Archive of Weather and Climate Extremes (https://wmo.asu.edu/) currently recognizes the 56.7 °C temperature recorded at a location in Death Valley, USA in 1913 as the hottest temperature for the globe, for the Western Hemisphere and for Northern Hemisphere. The organization also recognizes 55.0 °C temperature recorded in Kebili, Tunisia in 1931 as the hottest temperature recorded in Africa and 54.0°C ± 0.5°C temperature recorded in Tirat Tsvi (Tirat Zevi), present-day Israel in 1942. Two locations, one in Kuwait in 21 July 2016 and another in Pakistan in 28 May 2017, recorded temperatures of 54.0 °C. If verified, those temperatures would be recognized as the highest temperatures recorded for Asia (WMO RA region II) and the two observations would be the third highest WMO-recognized high temperature extremes and, significantly, the highest, officially-recognized temperatures recorded on the planet in the last 76 years.

Consequently, an international panel of atmospheric scientists was tasked with an analysis and verification of the 54.0 °C temperatures recorded in Kuwait and Pakistan in 2016. As this very detailed and comprehensive evaluation sets a new and very high standard for acceptance of new temperature extremes, the evaluation's specifics contain important information for the scientific community, and for the public and media at large. In particular, detailed discussion of the temperature sensor calibration testing is given for use in future extreme temperature assessment.

2. Metadata for Mitribah, Kuwait and Turbat, Pakistan

Starting in late 2016, a WMO committee was tasked with evaluation of two purported occurrences of a temperature of 54 °C, specifically on 21 July 2016 at Mitribah, Kuwiat and on 28 May 2017 at Turbat Pakistan. Background information of each station's observation was collected.

The Mitribah, Kuwait (WMO#40551) station is located at 29°49'N, 47°21'E (Figure 1). The station is automated and began operation in 2006. Temperature measurements are based on the uses of a HMP155 from Vaisala mounting a resistive platinum sensing element (Pt100). The sensor is covered by a naturally ventilated Vaisala DTR13 shield. Data are recorded at a one-minute interval by means of an Almos datalogger. Detailed equipment history logs were obtained along with photocopies of the actual observation log. Maps of the installation site as well as photographs of the station site were shared with the WMO evaluation committee.

The Turbat, Pakistan (WMO#41738) station is located at 25°59'N, 63°04' E (Figure 1). The station began operation in 1997 and with measurements manually recorded. Air temperature is measured with a mercury in glass thermometer, DDE 7461 manufactured by G.H Zeal, England, which is kept in a Stevenson screen located 2 m above ground. Detailed equipment history logs were given to the committee along with photocopies of the actual observation logs, photographs of the station sites, and maps of the installation site. Calibration history of the thermometer was given to the committee.

The national meteorological units of Kuwait and Pakistan shared with the WMO committee extensive weather data of their respective stations prior to and proceeding from the time of the record observations (Supplemental materials). This allowed evaluation committee members to ascertain the degree to which the extreme observation is in context with observations at the same location prior to and following the extreme occurrence. Contextual temporal consistency in observations, for example, was not evident in an earlier WMO committee's evaluation of a 1912 observation of 58 °C in El Azizia Libya and led, in part, to a refutation of that extreme (El Fadli et al. 2013).

Additionally, both Kuwait and Pakistan supplied the evaluation committee with detailed weather data for surrounding stations to Mitribah and Turbat respectively. This allowed the evaluation committee to assess the geographic consistency existing across the entire region at the time of the purported extreme. Geographical consistency in observations, for example, was also not evident in the El Azizia's 1922 observation.

Unanimously, the committee was quite impressed with the degree and quality of documentation associated with both observations. Data from both locations show good temporal consistency (e.g., the record observation was not a "spike" that failed to closely replicate data earlier or later at that location) and good geographic consistency (e.g., the surrounding stations, while obviously not exceeding the record observations, demonstrated a good degree of agreement in temperature change and magnitude with the extreme-recording station). The committee's analysis of the synoptic weather situation for both high temperature extremes indicated the presence of a large upper air ridge capping a strong surface high pressure system consistent with producing high surface air temperatures.

However, at this point in the discussion, the committee balked on recommending acceptance. They urged that both sensors be independently calibrated and possibly compared to ensure that the data were as accurately obtained as possible. This additional request has never before been made in a WMO evaluation committee of a record weather observation. It sets a new and higher standard than any previously accepted extreme.

The committee suggested that detailed investigation could achieve a quite reliable estimation of uncertainty associated to a recorded air temperature value, specifically instruments calibration uncertainty and measurement condition associated uncertainty. Moreover, the unique situation of having the same measured temperature values of 54.0 °C in both records makes the comparison of the sensors, when exposed to 54 °C, an unparalleled aspect to the study. Italian National Institute of Metrology (INRiM) staff was requested to join the evaluation committee. Recent work at INRiM included research activities and scientific production of direct interest for the present study: development of specific calibration systems and procedures (Musacchio et al. 2016), also for application in extreme environmental conditions (Musacchio et al. 2015), inclusion of changes in temperature standards for temperature series and extremes (Pavlasek et al. 2013), and specifically with the coordination of the large European project "MeteoMet – Metrology for Meteorology (Merlone et al. 2015, Merlone et al 2017).. Both the Mitribah and Turbat stations were asked to send the thermometers that recorded the top temperature values to INRiM, where staff was made available for this specific study that requested specific calibration and comparison procedures. The Kuwait sensor was sent to INRiM in May 2017, while the Pakistan sensor arrived in January 2018. Details of the metrology analysis are given in Section 3 for the Mitribah sensor and in Section 4 for the Turbat sensor. Section 5 details a direct comparison test of the two sensors. The analysis was made following the definitions and prescriptions of the Guide to the expression of uncertainty in measurement "the GUM" [JCGM 100:2008]

3. Metrology Analysis (Mitribah, Kuwait sensor)

3.1 Instrument Calibration

The Kuwait HMP155D was calibrated at INRiM on 12 to 16 June 2017 using the procedure for calibration of thermometers PT-T.3.3-01 Rev2, associated to the Institute's Calibration and Measurement Capability (CMC) as contained in the appendix C of the Mutual recognition Arrangement (MRA) of the CIPM, the International Committee for weights and measures (Comité International des Poids et Mesures). The CIPM MRA was signed also by WMO on 1 April 2010. The calibration results were reported in the calibration certificate n. 17-0496-01 issued on 2017-06-22. The calibration was made by comparison against a reference traceable to the ITS-90 fixed points in a comparator block inside the reference humidity generator. The sensor was positioned at 45° facing downward to avoid the embedded electronics warming the sensing element. The calibration was performed in air flowing at 0.04 m/s and with 50 %rh. The calibration uncertainty accounts for 0.06 °C and takes into account all contributions including the reference sensors, the calibration mean, the sensor's stability during calibration and its resolution.

The calibration demonstrated a deviation of the thermometer's reading in line with the instrument specification and declared uncertainty. The thermometer's sensing element is a Pt100 resistance thermometer, a platinum wire with nominal resistance of 100 Ω at 0 °C. The instrument output was recorded in resistance R and the temperature conversion T_{calc} was obtained applying the curve used by the Almos datalogger for the conversion, at the Mitribah station using equation (1).

$$T_{calc} = [(R\text{-}100.0)/0.39082]+(5.802/39082.0)*[(R\text{-}100.0)/0.39082]^2 \qquad (1)$$

The calibration points and results are reported in Table 1.

Consequently, the calibration correction at 54 °C was -0.12 °C with a calibration uncertainty of 0.06 °C (k=2), being k the numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty to cover 95 % of the distribution.

3.2 Evaluation of measurement uncertainty components

Full evaluation of measurement uncertainty for near-surface air temperature records using contact thermometer is not a trivial issue due to the numerous quantities of influence and non-perfect knowledge of the amplitude of their effects on different typologies of instruments and under different conditions. Moreover, evaluating an uncertainty budget “back in time” presents some difficulties. As reported above, the Kuwait Mitribah station measured and kept record of the main atmospheric variables: this is of fundamental help in this validation process. The main effects requiring qualitative and quantitative analysis are: wind speed at the time of the record, solar shield aging in terms of increased heat transferred to the sensor, datalogger contribution, sensors self-heating, sensor stability or drift in years.

3.3 Solar shield aging

Previous work [Lopardo et al. 2014] evaluated the effect of aging of solar shields in atmospheric thermometers. That analysis considered that exposure to meteorological conditions over time reduces the shield’s white-painting reflectivity, thus slowly introducing a temperature increase to the sensor reading. In this work three Vaisala screens were involved, a brand new one, a three-years old screen and a five-year old screen, equipped with the same sensors, calibrated in a climatic chamber were compared using identical thermometers. The results of this study are therefore interesting for this record validation, since the screen used at the Mitribah station is also in this case a Vaisala one and is three years old.

Experiments carried out on the five-year old and three-year old screens confirm the existence of shield ageing effect due to the degradation of the protective paint, in both cases. However, as expected, this effect is more evident when comparing shields with longer time of field use. The work also took into account the effect of wind on the magnitude of the aging contribution to temperature deviations. The differences between one-year and three-year old screens were distributed around 0 °C with a *thermal noise* of about 0.06 °C in presence of wind between 3.5 m/s and 5 m/s (as the case of Kuwait).

To evaluate the uncertainty associated to the aging, we therefore considered the max-min difference = 0.06 °C as a rectangular distribution (uniform-shape probability function) with associated standard uncertainty equals to 0.06 / $(12)^2$). This makes the contribution to the standard uncertainty due to the aging of about ±0.02 °C (thus ±0.04 °C with *k*=2). Consequently, the solar shield aging correction was 0 °C with a solar shield aging uncertainty: ±0.04 °C (*k*=2).

3.4 Datalogger contribution

Any electronics interface contributes to the uncertainty of an associated measurement. The magnitude of the uncertainty for the sensor with its electronic interface needs to be determined if a combined uncertainty for the measurement network is to be gauged. The data process and collection unit used at the Mitribah station is an ALMOS datalogger. Among the numerous activities on performance tests of Automatic Weather Station (AWS), WMO delivered a report made by Bruce Forgan [WMO 1999] on ALMOS AWS MSI2 - Sensor Interface Card Testing. This work evaluated the effect of the datalogger in the output of the different sensors potentially associated to the AWS, including the temperature value as translated from a resistance measurement. The board was also tested for effects of ambient temperature on the temperature measurements, in particular, on the electronics used for temperature measurement. The system was placed in a climatic chamber and exposed to temperatures up to 55 °C, thus in line with the present work. Decade box resistors were used to input the resistance in temperature channels. The 95% uncertainty in the average of the measurement of the entire temperature range was no greater than 0.0007 °C, which represents the repeatability of the system. This value is a negligible contribution in the present study. On the contrary, at a resistance close to the one read by the HMP155 at 54 °C, the Almos logger showed a deviation from the input resistance equivalent to -0.01 °C with an uncertainty of ±0.02 °C (*k*=2). In any case, the accuracy over the range of -10 °C to +55 °C met the specification as laid out in Equipment Specification A2672 of ±0.05 °C. For this investigation, the value measured in the report are considered to avoid over-estimation of uncertainty for the single 54 °C point. Consequently, the datalogger correction was -0.01 °C with associated uncertainty of ±0.02 °C (*k*=2) and with datalogger repeatability considered to be negligible.

3.5 Self-heating

The HMP 155 is a platinum resistance thermometer. The measurement of temperatures with this type of thermometer necessarily requires the pass of an electrical current through the thermometer's sensing element. The resistance of the thermometer is then calculated by observing the generated voltage and using the Ohm's law. This electrical current heats the thermometer element, by the Joule effect, causing a difference between the temperature of the sensor and the temperature to be measured. This effect is known as the self-heating. The sensor self-heating is usually determined in calibration laboratories under fixed conditions of temperature and wind speed but these conditions are highly variable when the thermometer performs air temperature measurements under real environmental conditions.

In the framework of European project MeteoMet [Merlone et al. 2015, Merlone et al. 2017] the self-heating of HMP 155 sensors was evaluated both in climatic chamber and in wind tunnel [Izquierdo et al. 2018]. Results of this study showed a significant contribution to temperature records due to self-heating. Moreover, the study showed differences of the same magnitude of the effect itself depending on whether the thermometer is calibrated in bath or in air. Considering that in the majority of the meteorological and climate applications the air temperature is the quantity to be measured, if the thermometers have been previously characterized in stirred liquid baths, the error due to self-heating can be under-estimated when on site temperatures are being performed. For this reason, the study mentioned and the present evaluation of the record both focused on tests in air. The temperature increase was evaluated to range up to more than 0.5 °C in case of currents of 3 mA in climatic chamber with 0.3 m/s of air flow. The uncertainty was evaluated to be 0.015 °C (k=2). At increasing airflow, the heat added by the passing current is removed by convection. The investigation continued in wind tunnel to evaluate the self-heating change when sensors are exposed to winds up to 5 m/s. As expected, higher wind speeds caused the heat brought by the current to the sensor to be removed by increasing convection. At wind speed of 5 m/s, close to the air velocity recorded at the time of the record, the self-heating was evaluated to be 0.027 °C.

For this record evaluation, the HMP 155 was calibrated, tested and measured in climatic chamber and any investigation in liquid was avoided. In the climatic chamber, airflow is approximately around 0.04 m/s. The calibration is so made with the sensor in heat equilibrium with the convection and self-heating at such wind speed. The calibration curve therefore already keeps into account the self-heating, which cannot be reduced to zero, as in standard thermometry, where measurements at multiple currents are made. For this reason, the value recorded in Mitribah at 2 pm of 2016-07-21 under a wind speed of 5.5 m/s does not require a correction with respect to the calibration value, due to self-heating which is already included in the calibration procedure that is made at lower air velocity. No correction is then applied for the self heating effect in field, with respect to the calibration correction, while a self heating uncertainty contribution of ±0.015 °C (k=2) is included.

3.6 Sensor drift

Platinum resistance thermometers are quite stable sensors. Their drift is normally very low and allows calibrations to be scheduled at more than one-year intervals. One-year recalibration is recommended for reference climate observations. Such sensors are moreover also quite stable at temperature changes in the range -40° C to + 60 °C and do not present significant hysteresis, as also studied in MeteoMet. Despite such advantage and due to the importance of the measurement here investigated, since the HMP 155 involved in this investigation arrived at INRiM for calibration about one year after having made the record measurement, it is important to make an evaluation of the drift of the sensor, if any, and the associated uncertainty. For this purpose, three different and independent analysis contributed to evaluate the drift and associated uncertainty: a) Vaisala study, b) Field exposure effect on drift made during MeteoMet (2013-2015) and c) Specific laboratory analysis carried on at INRiM for this study.

3.6.1 Vaisala study.

Due to its ongoing active collaboration within the MeteoMet consortium, Vaisala has direct links with the metrology community. For this specific record investigation, Vaisala provided relevant information about a drift analysis for 21 HMP155D probes that have been calibrated at 40 °C at least twice within twelve months during three years (2014-2017). Most of the analyzed sensors have operated in Vantaa, Finland where temperature typically varies between -15 °C and +25 °C; all calibrations reached 40 °C. Ideally, it would have been better to do drift analysis for sensors that have operated in similar conditions to the sensor in Kuwait, but there is not enough calibration data available in order to do statistically significant analysis for such sensors. However, the drift analysis should represent well also the sensor in Kuwait. Weather conditions in Kuwait, where temperature do not reach such low values, cause less stress to the platinum sensor than in Finland, so annual drift is hardly more than in Finland. Results of 40 repeated calibrations are reported in the following graph where difference of calibrations curves one with respect to the previous one are plotted. Using Figure 2, it is possible to evaluate that long-term drift is centered to 0 °C with an uncertainty (rectangular) mainly within ±0.02 °C, with exception of an outlier compensated immediately after.

In addition to this drift analysis, the sensor manufacturer was asked about long-term stability of the platinum sensor used in HMP155D. They test sensors according to DIN EN 60751:2009. In practice, this means 1000 hours at highest temperature. They also answered that they have never observed any drift at temperatures between 0 - 200 °C.

3.6.2 Field exposure effect on drift made during MeteoMet (2013-2015)

During the MeteoMet project, an effect of environmental conditions on characteristics of temperature sensors used by meteorological services was studied by three Polish institutions: Central Office of Measures, the Polish National Institute of Metrology, Institute of Low Temperature and Structure Research acting as Designated Institute in charge of maintaining the national temperature standard in Poland and University of Wrocław. The aim of this study, carried out over a two years period, was to investigate the factors affecting meteorological air temperature thermometers during normal operational work, which have an influence on sensor characteristic variability and uncertainty. Several series of calibration were performed just after an exposition of the thermometer to different atmospheric factors such as high humidity, high and low temperatures and rapid temperature changes. This study surpassed the values of exposure met by the sensor in Kuwait, but contains important information for this validation process. The repeated calibration after having exposed the sensor to conditions met in Kuwait, were made up to 50 °C thus giving significant contribution to this study. This allows to include a further aspect on the sensor stability from the time the record was measured and the time the instrument was calibrated.

The uncertainty on this evaluation ranged up to ±0.09 °C (k=2) with exposure of the thermometer down to temperatures well below those met in Kuwait and at constant 100 % relative humidity for months (Table 2). This uncertainty limit is in any case of the same order of magnitude as the one here evaluated. Such major value of course arises from the more extreme and accelerated changes.

Figure 3 shows the calibration curves obtained according to the treatment sequence of exposure as given in Table 2. It is of interest for this investigation that the second and third cycles were more similar to the conditions in Kuwait and consequently the differences between calibrations number 3 and number 4. Moreover, before and after both second and third cycles the thermometer was calibrated up to 50 °C thus forcing also under calibration a temperature range similar to the one here of interest.

Considering the plots of calibration 3 and 4 in Figure 3, temperatures around 50 °C demonstrate a repeatability of about 0.03 °C. This value differs from the one declared by Vaisala, since it is based on an used thermometer, already exposed to a number of conditions such as quick changes as well as forced and amplified thermal shocks. Additionally, the Vaisala calibration was limited to 40 °C, according to a general use in Finland, while the Polish test raised up to 50 °C.

The calibration uncertainties associated to the repeatability contribution evaluated both by Vaisala and in the MeteoMet study are respectively ±0.07 °C and ±0.09 °C: these values do not have to be added as source of uncertainty in the record investigation, since it has already been included as composition of the contributions of the calibration performed for the purpose at INRiM. This repeatability becomes then a source of uncertainty itself.

3.6.3 Specific laboratory analysis for this study

The Kuwait sensor was tested at INRiM for its repeatability when exposed to different temperatures. Four cycles of temperature changes were performed between 15 °C and 55 °C from September 2017 to March 2018. Due to this specific research, values were measured in terms of maximum change of readings at 54 °C, after thermal cycles, evaluated as differences with respect to reference thermometers in climatic chamber. Relative humidity was controlled during the cycles and ranged from 50 % to short exposure to 85 %.

The result of this repeatability test resulted for the Kuwait HMP 155 in line with the previous reported investigations, with a drift of no more than 0.04 °C all over the test duration and period. Considering this contribution having a rectangular distribution, centered symmetrically around 0 °C, the corresponding uncertainty is ±0.023 °C (k=2) with drift correction 0 °C.

3.7 Measurement value and uncertainty for the Kuwait record

The total uncertainty budget on the measured value of 54 °C for the Kuwait sensor is reported in Table 4, together with the associated corrections. All contributions are reported as expanded uncertainties in k=2 Therefore, the fundamental results of the metrology test for the Mitribah Kuwait thermometer reading of 54 °C is that the calibrated temperature is 53.87 °C ± 0.08 °C.

4. Calibration Analysis (Turbat Pakistan)

The sensor used at the Turbat station in Pakistan is a mercury in glass thermometer made in England by G.H.Zeal number DDE7461 range -10 °C to 65 °C. No calibration report is present nor is a recent calibration certificate available and therefore, as in the case of the Kuwait sensor, the validation process requires the calibration of the thermometer. Normally such sensors are calibrated in liquid bath at stable temperature and in adiabatic condition with the calibration mean, to have their readings associated to a reference thermometer also inserted in the same bath. Due to the specific purpose of this investigation, it was considered more significant to perform a calibration in air, to better represent the measurement conditions with the calibration process. No internationally recognized and accepted standard guidelines are at present available for the calibration of thermometers in air, and this is still an open issue under discussion both within the international metrology community as well as by WMO. Internal procedures are adopted, but a defined standard is not available. A project will start on this topic in 2018 by the European metrology organization, EURAMET.

A specific procedure was therefore adopted for the calibration of this thermometer in air. The thermometer was placed in a climatic chamber together with three INRiM Pt 100 thermometers coded GS01, GS02, NS02, NS05. The three Pt 100 are secondary thermometers directly traceable to the primary standards through a calibration by comparison against primary ones calibrated at ITS-90 fixed points. They were positioned close to the bulb of the Turbat thermometer, in a volume of a few cubic centimeters, to constantly evaluate the components due to the uniformity and stability, reported in the uncertainty table. The chamber used and the method is the one presented at the 2016 world meteorology exposition with the WMO TECO and called 'Meteocal.' It is based on a Kambic climatic chamber, a Fluke 1586A Super DAQ Precision temperature scanner equipped with m2588 STAQ Multiplexer unit. According to the temperature range occurring in Turbat, the calibration was made between 0 °C and 60 °C at the following points: 0 °C, 20 °C, 40 °C, 54 °C and 60 °C. The calibration showed a linear response of the thermometer and originated a calibration curve.

A general problem of the mercury-in-glass thermometer is that an air bubble can form in the liquid mercury column, thus introducing systematic errors in the readings and affecting the reproducibility. It is assumed that the thermometer was in correct operating condition at the time of the record; during tests and calibration at INRiM at the occurrence of the bubble formation, due mainly to handling and thermal cycles, the bubble was removed by slightly shaking the thermometer to allow mercury column to re-compact. The Pakistan thermometer was calibrated in vertical position and checked at 54 °C both in vertical and horizontal position showing no changes in the 54 °C indication.

The resulting calibration curve is close to linear with a constant term of -0.33 °C and is the following:

$$T_{cal} = T^2_{read}*a + T_{read} * b + c \qquad (2)$$

where a= -9.181 * 10^{-5} °C^{-1}, b= 1.006, c= -0.333 °C, T_{cal} is the temperature value corrected by the calibration curve, and T_{read} is the temperature marked on the sensor.

The residuals of this curve were accounted to determine the fitting uncertainty and originated a value of ±0.03 °C all over the range. The uncertainty components for such calibration are reported in the following table under the “calibration” group. The analog resolution of the thermometer plays the major role, being it 0.5 °C, but since this is a uniformly distributed uncertainty (rectangular distribution) its value is 0.15 °C which at the same time is very close to the human sensibility. A good operator, correctly positioned in front at the thermometer, can detect changes in the mercury column of about 0.15 °C. This was confirmed by ten people at INRiM, independently requested to read a temperature value. The resulting uncertainty was of the same order.

To achieve a better knowledge on the uncertainty associated with the measurement recorded on the 27th of May 2017 in Turbat, a couple of additional components are required. This being a mechanical system, there is no contribution from any datalogger. The ageing of the Stevenson screen is unknown and its effect in any case would introduce a zero to positive error in temperature reading. Therefore, the aim is intended at validating the maximum reliable air temperature at the time of the record, based on the available information. No self-heating is present in the case of mercury in glass thermometers.

The two main components on the measurement uncertainty to be added to the calibration uncertainty are therefore the thermometer repeatability or drift and the reading resolution as reported in the record table and accounted for 0.5 °C with rectangular distribution (i.e., the data indicate that measurements were taken to the nearest 0.5 °C). As reported for the Kuwait sensor, a series of thermal cycles were made also to the Pakistan sensor, by exposing it to temperatures between 0 °C and 60 °C and the reproducibility at 54 °C was evaluated. This included asking different people to read in different days the value of 54 °C restored after keeping the thermometer at room temperature (approximately 20 °C) and re-warming it. The test, corresponding to evaluating the reproducibility of the readings, did not showed a significant correction to be applied. After a statistical analysis on the standard deviation, the thermal cycles analysis originated a distribution of values of 0.052 °C around 54 °C, that corresponds to a rectangular distribution originating a value of about ±0.015 °C.

A further data analysis regarded the repetition of the points at 54 °C. A fitting curve was calculated, together with its residuals as check in the points around 54 °C. This curve confirmed the deviation of about 0.3 °C as calculated by the calibration curve over the 0 °C to 60 °C temperature interval. Values recorded by the Pakistan sensor and associated uncertainities are given in Table 4. Therefore, the fundamental results of the metrology test for the Turbat Pakistan thermometer reading of 54 °C is the calibrated temperature is 53.7 °C ± 0.42 °C.

5. Comparison of the Mitribah and Turbat Thermomters

This specific case allowed the unique opportunity directly to compare the two thermometers response when exposed to the same temperature at the same time. The fact that both measurements recorded equal values of 54 °C makes the comparison of the instrument reading at that temperature more significant. For this reason, despite the fact that calibration and characterization of the Kuwait HMP155 sensor was completed, its shipping back was delayed, to wait for the Pakistan thermometer to be received at INRiM. In early 2018, a test comparison was made by keeping both sensors at the same time in the climatic chamber (Figure 4). The sensing element of the HMP 155 and the bulb of the G.H.Zeal were kept in close vicinity to reduce thermal differences due to the inner gradient in the chamber; the HMP155 body and element were positioned in a way to avoid possible heat generated by the inner electronic to affect the G.H.Zeal readings (Figure 5).

Four calibrated INRiM PRTs were positioned in the surrounding volume to check for stability, gradient and accurate temperature value. The temperature was set to constantly 54 °C at 50 % of Relative Humidity and its reference value was evaluated as mean of the three PRTs closer to the two thermometers under test.

Four repeated comparisons were conducted in different days and bringing back the thermometers to room temperature between each measurement. Measurement values were recorded when the temperature of the chamber was stable within a few millikelvin and for at least 20 minutes, after hours of stabilization at 54 °C, to be sure any dynamic effect was concluded by both thermometers. Results are reported in Table 5.

Even though this procedure does not take into account an estimation of field conditions, it gives an extra information with reduced uncertainty on the response of both sensors at the same temperature of the record. As well known in metrology and measurement techniques, a relative process such this one strongly reduces the uncertainties due to common features. The main source of uncertainties, in this case are thus limited to:

a) temperature stability of the chamber, due to the very different dynamics of the two sensors this is a primary source of uncertainty. It is evaluated as range between the minimum and maximum values of the means of the closer three out of four reference thermometers, considering the more stable selected data. It accounts for 0.02 °C as a uniform (Max-Min) PDF;

b) the temperature gradient evaluated as standard deviation of the difference of the readings of three out of four of the reference thermometers in the volume around the two devices under test;

c) the uncertainty on the reference thermometers;

d) the uncertainty due to the resistance bridge used to read the reference thermometers

These components are common to the two sensors and have to be combined with the specific contributions:

e) the reading resolutions of the two sensors;

f) the repeatability, here evaluated as standard deviation on the differences among the reference value and the reading of respectively the Kuwait and Pakistan sensors over the four cycles.

Table 6 reports the uncertainty and the value read by both sensors.

The results give a robust confirmation on the independent analysis on the two sensors already showed. First of all, a direct evaluation at 54 °C confirm the value calculated applying the complete calibration curve. This means that within the evaluated uncertainty, at 54 °C the Kuwait sensor likely overestimated the temperature by about 0.16 °C and the Pakistan by about 0.28 °C. These results are reported in Table 7 and Figure 6.

The fundamental conclusions of the metrology analysis is, for the Mitribah Kuwait 54 °C thermometer reading, the calibrated temperature is estimated to be 53.87 °C with an expanded uncertainty of ±0.08 °C. Correspondingly, for the Turbat Pakistan 54 °C thermometer reading of 54 °C, the calibrated temperature is estimated to be 53.72 °C with an expanded uncertainty of ± 0.40 °C.

6. Evaluation, Record Determination and Implications

6.1 Evaluation

Because the establishment of these temperature record extremes has the potential for long-term significance as the highest temperatures recorded for Asia (WMO RA II), for the Eastern Hemisphere, and potentially the highest, officially-recognized temperatures recorded on the planet in the last 76 years, this evaluation demanded more extensive analysis and testing of these extremes than has ever been undertaken by an international evaluation committee of the WMO CCl Archive of Weather & Climate Extremes.

Following standard evaluation of data from both locations demonstrating good temporal consistency and good geographic consistency, along with analysis of the synoptic weather situation for both high temperature extremes (Supplemental material), an extensive metrological analysis of the two sensors was undertaken, including: a) the calibration of both thermometers, b) an effort of understanding and estimating the factors influencing the measurement and associated uncertainties and, c) a direct comparison of the two thermometer's readings when exposed simultaneously at 54 °C. The fundamental conclusions of the metrology analysis for the Mitribah Kuwait is a *calibrated* temperature estimated to be 53.87 °C with an expanded uncertainty of ± 0.08 °C (k=2). Correspondingly, for the Turbat Pakistan, the calibrated temperature is estimated to be 53.72 °C with an expanded uncertainty of ± 0.40 °C (k=2).

The committee noted that the observed temperatures of the two measurements compared extremely well with the metrological analysis with the Turbat Pakistan sensor's 54 °C value being well-within the limits of uncertainty and the Mitribah Kuwait sensor's 54 °C value being only 0.04 °C off the maximum uncertainty limits of the calibration tests. The critical question next addressed by the committee was what values to accept officially into the WMO Archive, the observed 54 °C value or the calibrated values?

Discussion by the committee on this point was influenced by the recent decision of another evaluation committee for the Antarctic region temperature (Laska et al. 2018). That evaluation committee recommended that a temperature observation atop the Davies Dome glacier be adjusted downward from its observed value of 17.9 °C to 17.0 °C ± 0.2 °C (62.6°F ± 0.4°F). In essence, this decision set the precedent of using the calibrated adjusted measurements as the official values in the WMO records. Therefore, it was the consensus (but not unanimous) decision to recommend acceptance of the calibrated observations.

However, a concern was raised by the committee with regard to the degree of accuracy that should be reported for the record. The calibrated values were measured (with uncertainties) to the second decimal place. Panel members noted that the accuracy of most modern meteorological measurements is to one decimal place (0.1 °C). Therefore, the consensus of the committee was that the Mitribah Kuwait observation be accepted as 53.9 °C and the Turbat Pakistan observation as 53.7 °C.

6.2 Record Determination

Following the precedent for acceptance of calibrated values by the Antarctic temperature evaluation committee, the official values are recorded as:

Mitribah, Kuwait [53.9 °C ± 0.1 °C (129.0°F ± 0.2°F), adjusted by calibration analysis from a reported observation of 54 °C]

Turbat Pakistan [53.7 °C ± 0.4 °C (128.7°F ± 0.7°F), adjusted by calibration analysis from a reported observation of 54 °C]

As such, they are the third and fourth highest WMO-recognized extremes.[1] Significantly, the Mitribah and Turbat values are the highest, officially-recognized temperatures recorded on the planet in the last 76 years. Therefore, the WMO CCl Archive of Global Weather and Climate Extremes has accepted the temperature values of 53.9 °C at Mitribah Kuwait in 21 July 2016 and the temperature value of 53.7 °C in Turbat Pakistan in 28 May 2017 as verified and the Mitribah Kuwait temperature is now accepted as the highest temperatures recorded for Asia (WMO RA II).

6.3 Implications

This calibration analysis marks a major advancement in WMO CCl extreme evaluations. Although international comparisons of meteorological instruments have been conducted in the past under WMO guidance, the use of detailed calibration of the instruments producing the global extremes data marks a new level of assurance that our record of global, hemispheric and regional extremes are, and will remain, as accurate as possible. The calibration (and indeed the affiliation of this paper's authorship)

[1]It should be noted that (a) other values of equal or near equal temperatures have not been formally evaluated by the WMO (e.g., 129°F (53.9 °C) value recorded on 30 June 2013 in Death Valley USA) and (b) both the Mitribah and Turbat value are within limits of error with 1942 value 54 °C ± 0.5 °C at Tirat Tsvi (present-day Tirat Zevi Israel) for WMO RA VI (Europe) although the Israeli location is geographically within Asia.

demonstrates that a high level of trust and collaboration exists between WMO, national Meteorological Services, and research institutions around the world.

Second, this calibration analysis sets a precedent that may increase the level of difficulty for the World Meteorological Organization's CCl to approve a new temperature record if the instrument calibrations and evaluations applied in this analysis are not allowed. At this time, new records are examined with "all available data" but evaluations do not mandate the use of calibration as a prerequisite to acceptance. Committee members noted that Regional Instrument Centres can and do assist Regional Members in calibrating their national meteorological standards and related environmental monitoring instruments for variables such as temperature, humidity and pressure (CIMO 2014). Indeed, the WMO Commission for Instruments and Methods of Observation (CIMO) explicitly states (page 64), "All temperature-measuring instruments should be issued with a certificate confirming compliance with the appropriate uncertainty or performance specification, or a calibration certificate that gives the corrections that must be applied to meet the required uncertainty. This initial testing and calibration should be performed by an accredited calibration laboratory. Temperature-measuring instruments should also be checked subsequently at regular intervals, the exact apparatus used for this calibration being dependent on the instrument or sensor to be calibrated." However, as committee members also noted, accurate assessment of important and critical global extremes often must require a complete re-analysis of the instruments field performances, station site characteristics, systems ageing and identify quantities of influence which are normally not included in the calibration. Uncertainty should then be estimated, with the calibration uncertainty as one of the components and other identified factors to complete the total expanded uncertainty.

Third, it should be noted that this level of calibration potentially accords an unfair advantage to earlier observations (pre-2000 extremes) where calibration analysis and evaluation simply is not possible compared to a new temperature-record challenger which undergoes calibration tests. For example, records from the early 1900s conceivably might have been rejected based on calibration with current equipment but such reanalysis is not possible given that those equipment/sensors are not available for modern analysis. However, as committee members note, WMO guidelines to assess the quality/properties of long time series have been published (WMO, 2009). At this time, the WMO Archive of Weather and Climate Extremes does not remove historical records without sufficient and significant cause, e.g., new (ideally physical) evidence of error in the historical record.

Lastly, it should be noted that there are limits to the degree that calibration can reveal the true air temperature. For example, laboratory calibration cannot assess air movement through the shelter on that day, or other in situ processes that unknown or impossible to replicate in the laboratory. The potential relative importance of all those other factors may influence the actual observation in the real world and may be a part of any WMO evaluation and recommendation.

7. Acknowledgments

Manola Brunet thanks support from the Copernicus for Climate Change Services (C3S)/ European Centre for Medium Range Weather Forecast (ECMWF) C3S 311A_LOT 1_DATA RESCUE contract and the European Research Area for Climate Services (ERA4CS) Consortium (FP7-ERA-NET Cofund, contract number 690462).

References

Cerveny, R.S., J. Lawrimore, R. Edwards, C. Landsea 2007: Extreme Weather Records: Compilation, Adjudication and Publication, *Bulletin of the American Meteorological Society* 88 (6): 853-860

Cerveny, R.S., and co-authors, 2017 : WMO Assessment of Weather and Climate Mortality Extremes: Lightning, Tropical Cyclones, Tornadoes, and Hail, *J. Wea. Clim. Soc.* https://doi.org/10.1175/WCAS-D-16-0120.1

CIMO, 2014: CIMO Guide, Part IV Quality assurance and management of observing systems, Chapter 4. Testing, Calibration and Intercomparison. World Meteorological Organization, Geneva Switzerland. https://library.wmo.int/index.php?lvl=notice_display&id=12407

El Fadli, K., and co-authors, 2013: World Meteorological Organization Assessment of the Purported World Record 58ºC Temperature Extreme at El Azizia, Libya (13 September 1922), *Bulletin of the American Meteorological Society*. doi: http://dx.doi.org/10.1175/BAMS-D-12-00093.1

Forgan, B., 1999: *Instrument test report number 657*, Almos AWS MSI2- Sensor Interface Card Testing, World Meteorological Organization, Geneva Switzerland, 8 pages.

JCGM, 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement.

Izquierdo, C.G. and co-authors , Evaluation of the self-heating effect in a group of different meteorological thermometers, *Meteorological Applications* 2018

Lang, T.J., and co-authors, 2016 : WMO World Record Lightning Extremes: Longest Reported Flash Distance and Longest Reported Flash Duration, *Bulletin of the American Meteorological Society*, http://dx.doi.org/10.1175/BAMS-D-16-0061.1

Lopardo, F., and co-authors 2013: Comparative analysis of the influence of solar radiation screen ageing on temperature measurements by means of weather stations, *International Journal of Climatology*, 34(4),:1297–1310, DOI: 10.1002/joc.3765

Merlone A., and co-authors, 2015: The MeteoMet project – metrology for meteorology: challenges and results, *Meteorological Applications*, 22: 820-829, DOI: 10.1002/met.1528

Merlone, A., and co-authors. 2018: The MeteoMet2 project – Highlights and results, *Meas. Sci. Technol*. 29 025802 Measurement Science and Technology, 29(2), https://doi.org/10.1088/1361-6501/aa99fc

Musacchio C. and co-authors, 2015: Arctic Metrology: calibration of radiosondes ground check sensors in ny-Ålesund, *Meteorological Applications*, 22: 854-860, DOI: 10.1002/met.1506

Musacchio, C., A. Merlone, A. Viola, V. Vitale, M. Maturilli, 2016: Towards a calibration laboratory in Ny-Ålesund, Rend. Fis. Acc. *Lincei*, 27:243,249, DOI 10.1007/s12210-016-0531-9

Pavlasek, P., A. and co-authors, 2016: Effect of changes in temperature scales on historical temperature data, *International Journal of Climatology*, 36(2): 1005-1010, DOI: 10.1002/joc.4404

Thorne, P.W., and co-authors, 1028: Towards a global land surface climate fiducial reference measurements network, *International Journal of Climatology*, Open access https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.5458

WMO, 2009: "Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation",WCDMP-No. 72, World Meteorological Organization, Geneva Switzerland, (https://www.ecad.eu/documents/WCDMP_72_TD_1500_en_1.pdf)

Figure Captiosn

Figure 1. Geographic location of the two extreme temperature stations (Mitribah, Kuwait and Turbat Pakistan).

Figure 2. Drift between calibrations. Calibration is done in a liquid bath. Calibration uncertainty is ±0.07 °C (*k*=2).

Figure 3. Calibration curves evaluated after exposures to environmental conditions given in Table 2.

Figure 4. The Kuwait HMP155 and the Pakistan G.H.Zeal thermometers in the climatic chamber for comparison.

Figure 5. Close up view of the sensing element, mercury bulb and INRiM four Pt 100 reference thermometer. The four Pt 100 are not in contact with any of the two thermometer to avoid self-heating to be transferred to the devices under test.

Figure 6. Summary results. Values calculated applying calibration curves (A) and from direct comparison at 54 °C (B). Uncertainties of values A include calibration uncertainty and estimation of measurement uncertainty.

Table 1. HMP 155 calibration results where T_{ref} is the temperature indication of the reference thermometer (°C), R is the resistance (Ω), T_{calc} is the temperature conversion through application of the curve used by the Almos datalogger, Δt is the temperature difference (°C) between the readings of the HMP155 translated into temperature values using the equation adopted by the station datalogger and the reference traceable temperature.

t_{ref}	R	T_{calc}	Δt
20.27	107.9246	20.34	+0.07
39.85	115.5145	39.93	+0.08
54.26	121.0840	54.38	+0.12
59.93	123.2712	60.07	+0.14

Table 2. MeteoMet study on HMP15 stability after environmental exposure, using period, factor, place and conditions parameters.

Calibration		Exposure			
No	Date	Period	Factor	Place	Condition
1	06. 2013				
		06.2013 - 03.2014	time	Lab	ambient
		03.2014 - 07.2014	high humidity	cave	t = 7.5 °C± 0.5 °C humidity ≈ 100%
		07.2014 - 11.2014	time	Lab	Ambient
2	11.2014				
		12.2014 - 05.2015	mid temperature	Stevenson Screen	t_{min} = -9 °C t_{max} = 26°C
3	05.2015				
		05.2015 - 09.2015	high temperature	Stevenson Screen	t_{min} = 11°C t_{max} = 36°C
4	09.2015				
		09.2015 - 11.2015	time	Lab	ambient
		11.2015 - 12.2015	low temperature	refrigerator	t = -30 °C
5	12.2015				

Table 3. Kuwait sensor - Determining the corrected record temperature and the associated uncertainty.

Quantity/Contribution	Estimated Value or correction °C	Uncertainty °C	Divisor	Distribution	Uncertainty. (*k*=1) (°C)
Measured value	54.00				
Calibration					
Calibration (Procedure INRiM PT-T.3.3-01 Rev. 2)	-0.12	0.06	2.00	normal	0.03
Measurement					
Solar shield ageing	0	0.06	3.46	rectangular	0.02
HMP Logging ALMOS	-0.01	0.02	1.00	normal	0.01
Self heating	0	0.015	1.00	normal	0.015
Sensor drift and repeatability (1 year or short term after exposure to whole temperature range)	0	0.02 (Vaisala) 0.04 (MeteoMet) 0.04 (INRiM)	3.46	rectangular	0.01
HMP resolution	0	negligible			0.00
Repeatability during test	0	negligible			0.00
Corrected value	**53.87**				
				Combined uncertainty (*k*=1)	0.04
				Expanded uncertainty (*k*=2)	**0.08**

Table 4. Pakistan sensor - Determining the corrected record temperature and the associated uncertainty

Quantity/Contribution	Estimated Value or correction °C	Uncertainty °C	Divisor	Distribution	Uncertainty. (*k*=1) (°C)
Measured value	54.00				°C
Calibration					
Chamber temperature stability	0	0.02	3.46	rectangular	0.006
Chamber temperature uniformity	0	0.052	3.46	rectangular	0.015
Reference Thermometers calibration	0	0.017	1	normal	0.017
Read-out for reference PRTs	0	0.01	1	normal	0.01
Pakistan Thermometer resolution (includes repeatability and readings from different operators)	0	0.5	3.46	rectangular	0.14
Fitting	0	0.03	1	normal	0.03
Corrected value	**53.72**				
				Combined uncertainty (*k*=1)	0.15
				Calibration uncertainty (*k*=2)	**0.30**
Measurement					
Pakistan Thermometer resolution	0	0.5	3.46	uniform	0.14
Reproducibility @ 54 °C	0	0.052	3.46	uniform	0.015
Corrected value	**53.72**				
				Combined uncertainty (*k*=1)	0.20
				Measurement uncertainty (*k*=2)	**0.40**

Table 5. Results of the comparison at 54 °C between Kuwait and Pakistan sensor.

Reference Temperature (°C)	Kuwait Sensor Temperature (°C)	Kuwait – Reference Temperature (°C)	Pakistan Sensor Temperature (°C)	Pakistan – Reference Temperature (°C)
54.009	54.168	0.158	54.268	0.259
54.012	54.184	0.171	54.264	0.251
53.832	53.976	0.144	54.125	0.293
53.704	53.887	0.183	54.006	0.303
Mean Difference		0.164		0.276
Repeatability (st. dev.)		0.017		0.025

Table 6. Uncertainties on direct comparison of Kuwait and Pakistan thermometers at 54 °C.

Uncertainty contribution	**Value (°C)**	**Distribution**	**Uncertainty (*k*=1) (°C)**
Reference system			
Chamber temperature uniformity	0.052	rectangular	0.015
Chamber temperature stability	0.02	rectangular	0.006
Reference Thermometers calibration	0.017	normal	0.017
INRiM Data-logger and resolution	0.01	normal	0.010
		Combined uncertainty (*k*=1)	**0.025**
Kuwait instrument			
INRiM Data Acquisition for HMP155	0.01	normal	0.010
Repeatability	0.017	normal	0.017
		Combined uncertainty (*k*=1)	0.032
		Expanded uncertainty (*k*=2)	**0.064**
Pakistan instrument			
G.H.Zeal thermometer resolution	0.5	rectangular	0.14
Repeatability	0.025	normal	0.025
		Combined uncertainty (*k*=1)	0.144
		Expanded uncertainty (*k*=2)	**0.29**

Table 7. Summary results of the metrological analysis………

	Corrected Value (°C)	Uncertainty (°C)
Kuwait calibration (A)	53.87	±0.080
Kuwait comparison (B)	53.84	±0.064
Pakistan calibration (A)	53.72	±0.40
Pakistan comparison (B)	53.72	±0.29

Supporting Materials.

To establish the synoptic weather conditions for the Mitribah Kuwait 2016 and Turbat Pakistan 2017 events, the ERA Interim Reanalysis (Dee et al. 2011) was queried for specific meteorological data for those days.

Synoptic weather conditions for both 21 July 2016 at Mitribah, Kuwait and on 28 May 2017 at Turbat Pakistan show evidence of extensive high pressure over the Middle East at the two locations and times respectively. Figure A1 indicates that a large 500 hPa ridge was centered of the northern Arabian (Persian) Gulf in July of 2016 with some of the highest heights located over Iran and Kuwait (~5830 geopotential meter 500 hPa heights). This was coupled with surface dew points in the 0 °C range and surface pressures on the order of 1004 hPa over Kuwait. Nearby surrounding stations recorded markedly high temperaures as well (51.6 °C Sabriya, 51.1 °C Jal Alyah and 50.8 °C at Al Abraque) with comparable trends to Mitribah.

Correspondingly, on 28 May 2017 at Turbat Pakistan, a high pressure ridge is also firmly entrenched over the Middle East, centered over Iraq (Figure A2). 500 hPa heights over Turbat Pakistan were on the order of 5820 geopotential meters). With the Mitrihah Kuwait observation in 2016, the 500 hPa ridge over Turbat was coupled with dry air (wet bulb temperatures 20 °C) and sealevel pressures at Turbat on the order of 1000 hPa. Nearby surrounding stations recorded markedly high temperatures as well (46 °C Panjgur, Gwadar 45 °C) with comparable trends to Turbat, noting that Panjgur is at over 900 metres elevation, and Gwadar is on the coast, so both would normally be expected to be substantially cooler than Turbat.

Reference:

Dee, D. P., and co-author, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society* 137(656): 553-597 https://doi.org/10.1002/qj.828

Supporting Material Figure Captions

Figure A1. 500 hPa heights (meters) for 00 UTC, 06 UTC, 12 UTC and 18 UTC for 20-23 July 2016. Dashed red lines indicate 500 hPa air temperature (°C). Black dot indicates location of Mitribah Kuwait. Extreme occurred 12 UTC 21 July 2016.

Figure A2. 500 hPa heights (meters) for 00 UTC, 06 UTC, 12 UTC and 18 UTC for 27-29 July 2017. Dashed red lines indicate 500 hPa air temperature (°C). Black dot indicates location of Turbat Pakistan. Extreme occurred 12 UTC 28 May 2017.

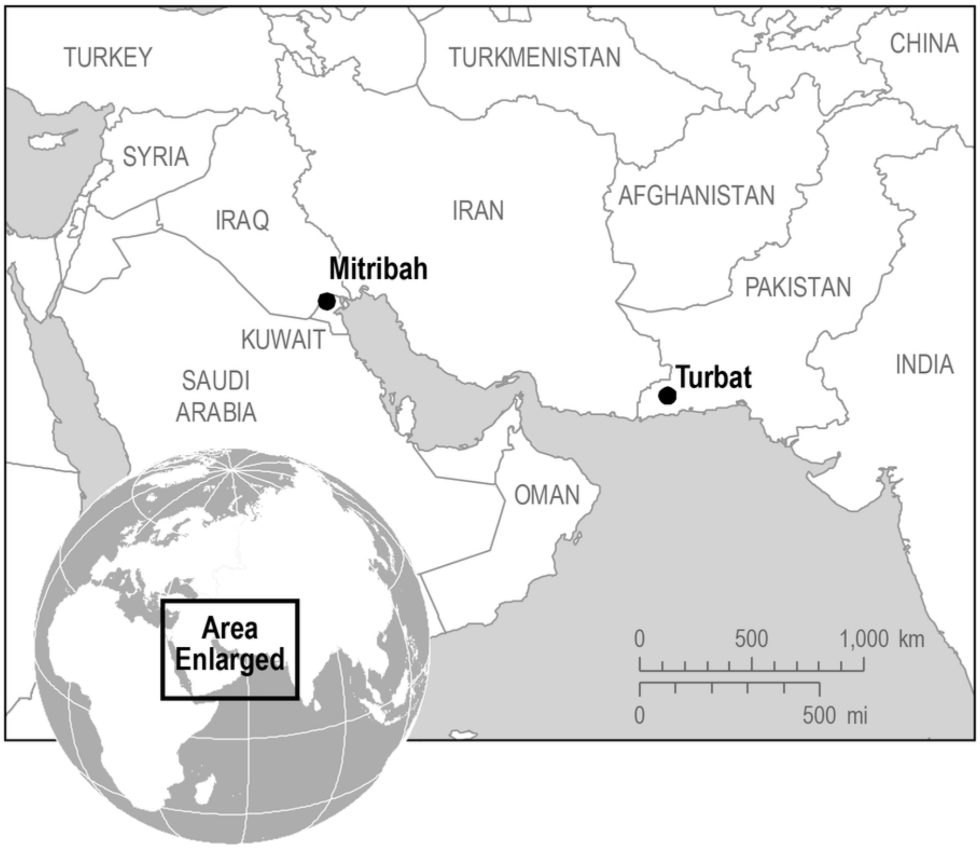

Figure 1. Geographic location of the two extreme temperature stations (Mitribah, Kuwait ad Turbat Pakistan).

78x67mm (300 x 300 DPI)

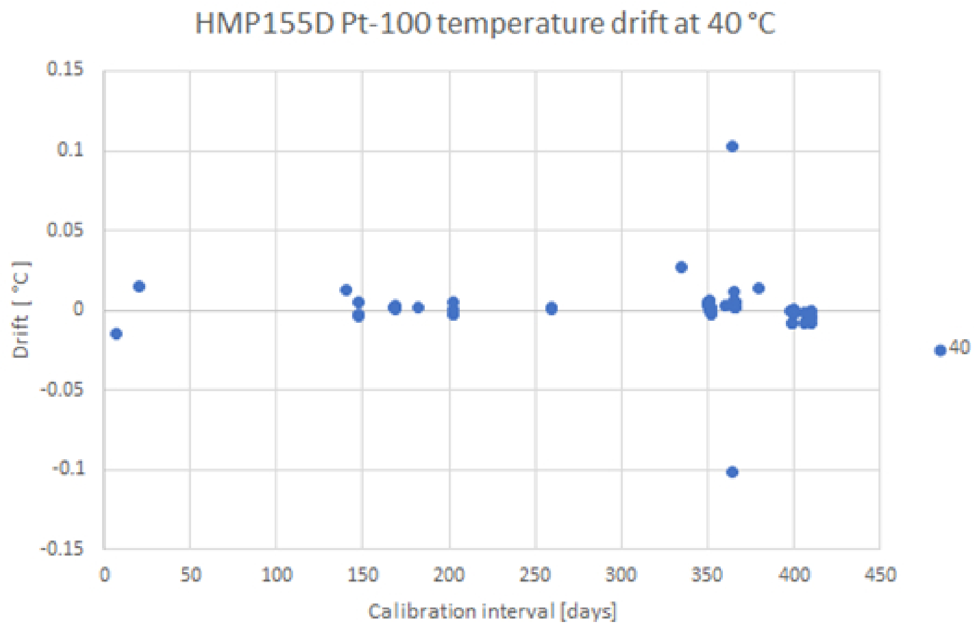

Figure 2. Drift between calibrations. Calibration is done in a liquid bath. Calibration uncertainty is ±0.07 °C (k=2).

45x30mm (300 x 300 DPI)

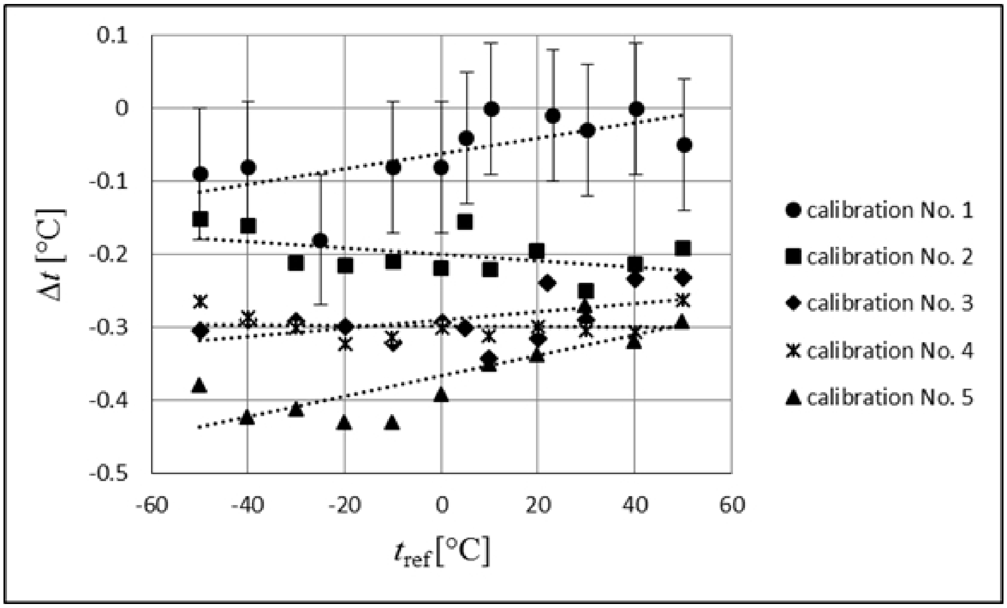

Figure 3. Calibration curves evaluated after exposures to environmental conditions given in Table 2.

49x29mm (300 x 300 DPI)

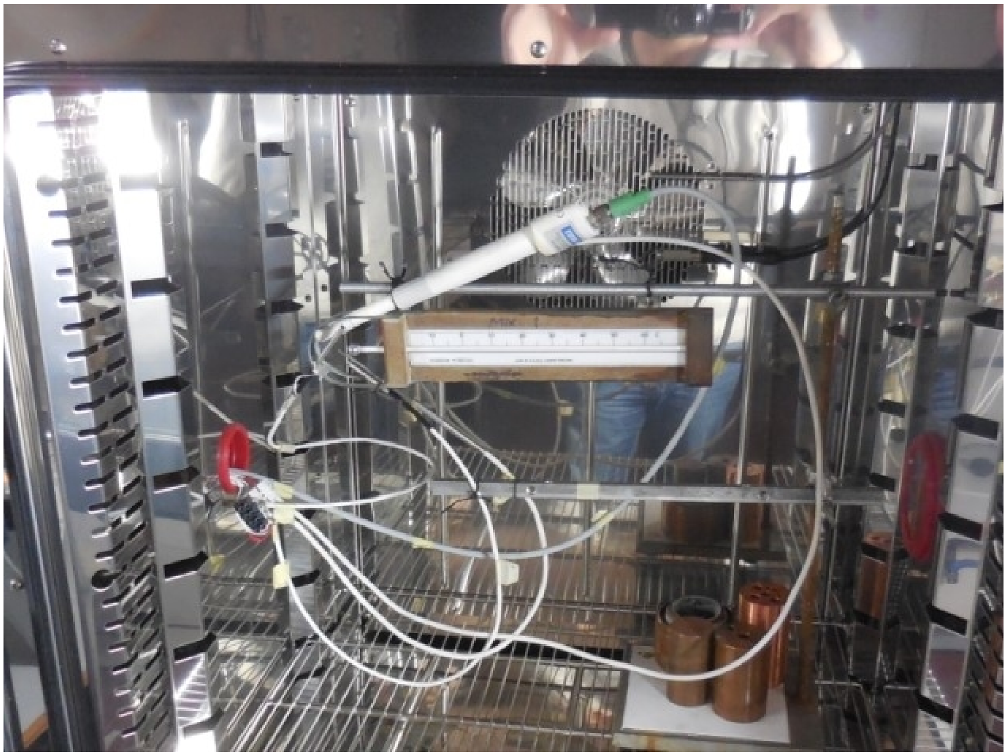

Figure 4. The Kuwait HMP155 and the Pakistan G.H.Zeal thermometers in the climatic chamber for comparison.

53x39mm (300 x 300 DPI)

Figure 5. Close up view of the sensing element, mercury bulb and INRiM four Pt 100 reference thermometer. The four Pt 100 are not in contact with any of the two thermometer to avoid self-heating to be transferred to the devices under test.

102x76mm (220 x 220 DPI)

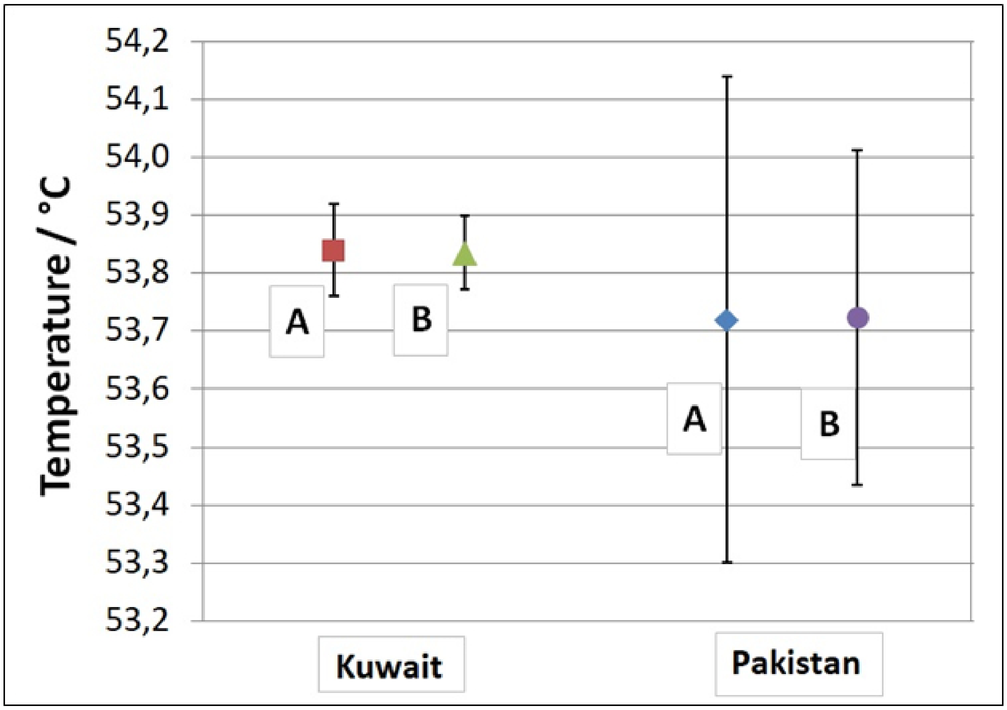

Figure 6. Summary results. Values calculated applying calibration curves (A) and from direct comparison at 54 °C (B). Uncertainties of values A include calibration uncertainty and estimation of measurement uncertainty.

53x37mm (300 x 300 DPI)

Graphical Table of Contents

Title: World Meteorological Organization Evaluation and Calibration Testing of 2016/17 temperatures of 54.0 °C recorded in Mitribah, Kuwait and Turbat, Pakistan as Record Temperature Extremes

Authors: Andrea Merlone, Hassan Al-Dashti, Nadeem Faisal, Randall S. Cerveny*, Said AlSarmi, Pierre Bessemoulin, Manola Brunet, Fatima Driouech, Yelena Khalatyan, Thomas C. Peterson, Fatima Rahimzadeh, Blair Trewin, M.M. Abdel Wahab, Serpil Yagan, Graziano Coppa, Denis Smorgon, Chiara Musacchio, Daniel Krahenbuhl
(* corresponding author)

Caption (80 words or 3 sentences): A World Meteorological Organization (WMO) committee officially evaluated temperature record extremes of 54.0 °C at two locations, one in Mitribah, Kuwait on 21 July 2016 and a second in Turbat, Pakistan on 28 May 2017. Metrological testing concluded the Mitribah value is a temperature estimated to be 53.87 °C with an expanded uncertainty of ± 0.08 °C. Correspondingly, for the Turbat value the temperature is estimated to be 53.72 °C with an expanded uncertainty of ± 0.40 °C.

Figure: