Metrologia 43 (2006) S46–S50 doi:10.1088/0026-1394/43/2/S10

Low-background temperature calibration of infrared blackbodies

Adriaan C Carter[1], Raju U Datla[1], Timothy M Jung[2], Allan W Smith[2] and James A Fedchak[2]

[1] National Institute of Standards and Technology, Gaithersburg, MD, USA
[2] Jung Research and Development Corporation, Washington, DC, USA

Received 18 October 2005
Published 23 March 2006
Online at stacks.iop.org/Met/43/S46

Abstract
The Low Background Infrared (LBIR) facility at the National Institute of Standards and Technology (NIST) has performed ten radiance temperature calibrations of low-background blackbodies since 2001, when both the calibration facility and method of calibrating blackbodies were significantly improved. Data from nine of these blackbody calibrations are presented, showing a surprisingly large spread in blackbody performance. While some blackbodies performed relatively well, in no case did the measured radiance temperature agree with the temperature sensors in the blackbody core to within 0.3 K over the entire operating temperature range of the blackbody. Of the nine blackbodies reported, five showed temperature errors greater than 1 K at some point in their operating temperature range. The various sources of uncertainty, such as optical geometry and detector standard uncertainty, are presented with examples to support the stated calibration accuracy. Generic blackbody cavity design features, such as cavity thermal mass, cavity volume and defining aperture placement are discussed and correlated with blackbody performance. Data are also presented on the performance of the absolute cryogenic radiometers (ACRs) that are used as detector standards in the calibration of blackbodies. Recent intercomparisons of all the LBIR ACRs with a trap detector calibrated against the NIST primary optical power measurement standard show that ACRs used to calibrate blackbodies are suitable detector standards and contribute less than 0.02% uncertainty ($k = 1$) to radiance temperature measurements of the blackbody cavities.

1. Introduction

The stability of blackbody sources and their ability to be modelled accurately make them well suited for radiometric calibration activity. However, the design and fabrication of the blackbody in industry or the laboratory may, and often does, result in blackbody performance that deviates from ideal because of performance requirements other than calibration accuracy. From the standpoint of this paper, an ideal blackbody is one that can be used as a radiometrically accurate source.

The Low Background Infrared (LBIR) facility at the National Institute of Standards and Technology (NIST) was established in 1989 to fulfil the national need for calibrations in the infrared wavelength region in a low, 15 K to 80 K, background environment [1]. The calibration capability is based on the ability to measure infrared power absolutely using absolute cryogenic radiometers (ACRs) [2, 3]. LBIR maintains two ACR I model radiometers which have a time constant of 22 s, a responsivity of 29.7 K mW^{-1} and the capability to measure powers from 5 nW to 250 µW as well as two more sensitive ACR II model radiometers which have a time response of 17 s, a responsivity of 210 K mW^{-1} and the capability to measure powers from 40 pW to 100 µW [4]. Power measurements of 500 pW with an uncertainty 4 pW (type A, $k = 1$) are now routinely achieved with the ACR II radiometers. The four ACRs are periodically compared against the NIST primary national optical power measurement standard [5, 6]. The results of the most recent intercomparisons are presented below.

In 2001 upgrades to the LBIR facility were completed which allow for higher accuracy in the calibration of blackbodies [7]. Blackbody calibrations before then had significantly higher uncertainties and are excluded from this survey. Since 2001 ten blackbodies have been calibrated with the LBIR ACRs. Nine of these calibrations are reported in this paper and represent blackbodies incorporating six different blackbody designs. Sources of systematic uncertainty are discussed and show that the quality of a particular blackbody calibration depends on the combined performance of the customer's blackbody and the ACR together, not just the optimal performance capabilities of the LBIR calibration facility. Unless otherwise stated, all uncertainties presented in this paper are type A with $k = 1$.

2. Low-background calibration facilities

The LBIR facility performs calibration activity in a high-vacuum, low-temperature background environment in one of two cryogenic-vacuum chambers. Both chambers are approximately 2 m in length and have cryo-shrouds that enclose a cylindrical volume that is about 1.5 m in length and about 0.5 m in diameter. At test conditions the chambers are typically operated around 1×10^{-7} Pa and the cryo-shrouds are cooled to between 15 K to 80 K with closed-cycle He refrigerators. Inside the chambers the ACRs are mounted on liquid He cryostats which are cooled to 2 K by pumping the liquid He down to 2670 Pa.

3. Absolute cryogenic radiometer performance

Intercomparison between the four LBIR-maintained ACRs and the NIST primary optical power measurement standard was conducted with a QED-150 trap detector in a manner similar to previous intercomparisons [8]. The QED-150 is a three Si photodiode reflectance trap detector having five internal reflections so that the total reflection at 632.8 nm should be less than 0.4%. To determine an absolute responsivity, the optical power from an intensity-stabilized He–Ne laser was measured using first the NIST primary optical power measurement standard, the high accuracy cryogenic radiometer (HACR) and then again later by the newer Primary Optical Watt Radiometer (POWR). The external responsivity of the trap was measured to be 0.502 448 A W^{-1} by the HACR at the beginning of the intercomparison effort and 0.502 457 A W^{-1} by the POWR at the end. The uncertainty associated with the calibrated responsivity is 0.02%.

The light source for the intercomparison was a polarized, stabilized He–Ne laser with a maximum power output of 1.5 mW. The laser light was transmitted into the cryogenic-vacuum chamber through a Brewster angle window and a 2 cm diameter hole in the cryo-shroud. The ACR under test was placed 0.8 m from the 2 cm hole in the cryo-shroud so that it received only 20 μW to 30 μW of background radiation through the hole in the shroud. The calibrated QED-150 trap detector was used to measure the laser power just in front of the Brewster angle window. The trap detector was then removed from the beam path to allow the power to be measured by the ACR inside the chamber. The effective transmission of the He–Ne laser from the trap detector location in front of the Brewster window to the location of the ACR 1.5 m away was measured in a separate experiment. This effective transmission is then used to correct the ACR power measurements, allowing a direct comparison with the QED-150 trap detector power measurements. Table 1 shows the intercomparison results as the ratio of radiometric power measured in the calibrated QED-150 trap to that measured in the ACR. The total uncertainty of the individual intercomparisons is dominated by the Brewster window transmission measurement uncertainty. The data suggests that on average the LBIR ACRs measure 0.08% more power than the QED trap detector, with an uncertainty of 0.04%.

Table 1. Ratio of radiometric power measured in the calibrated QED-150 trap to that measured in the ACR.

ACR Name	$R_{\text{trap/ACR}}$	Combined uncertainty
ACR Ia	0.999 339	0.024%
ACR Ib	0.999 130	0.030%
ACR IIa	0.999 595	0.053%
ACR IIb	0.998 765	0.029%

The thermal non-equivalence of the ACRs was evaluated as well. Each of the four ACRs, except the ACR IIa, was equipped with two temperature sensors and two heaters for the electrical substitution power measurement. The placement of the individual heaters and sensors in the ACRs was chosen such that in theory one pair would provide the smallest thermal non-equivalence and the other the largest non-equivalence in the optical power measurement. By measuring the effect of each combination on the power measurement the maximum possible error introduced by thermal non-equivalence for a given cavity can be determined. For the ACR I model the maximum measured non-equivalence was 0.013%. For the ACR II model the maximum non-equivalence observed was 0.64%. The ACR II model was designed to be more sensitive than the ACR I while still maintaining a reasonable response time. The increased sensitivity came from an increase in responsivity from a weaker (more resistive) thermal link. To counter an increase in time constant, the thermal mass was lowered by thinning the walls of the ACR light trapping cavity. The measured non-equivalence of 0.64% is considerably more than the 0.2% that had been estimated when the ACR II model radiometers were designed. A much more complete discussion of the intercomparison of the LBIR ACRs with the primary optical power measurement standards is presented elsewhere [6]. In all the blackbody calibrations performed by LBIR the results of the intercomparisons and thermal non-equivalence measurements of the ACR used in that calibration are incorporated into the calibration.

4. Blackbody calibrations and overall performance

Using the LBIR-maintained ACRs the performance of the nine blackbodies calibrated since 2001 can be accurately assessed and compared. The blackbodies calibrated at the LBIR facility were cryogenic-vacuum blackbodies that operated with cavity temperatures of 180 K to 800 K. All but one of the blackbodies tested included aperture wheels to vary the size of the defining aperture in front of the blackbody cavity. All the blackbodies had space between the defining aperture and the blackbody cavity for a shutter, one or two filter wheels or a chopper wheel. All of the calibrations were performed entirely within cryogenic-vacuum chambers at 1×10^{-7} Pa in a 20 K to 30 K background environment.

The radiance temperatures of the blackbodies were calibrated using length measurements of the calibration optical geometry, power measurements made by the ACRs and the Stefan–Boltzmann law. The optical geometries for the calibrations were intentionally made very simple so that they could be well-characterized. The ACRs are typically fitted with a 1 cm or 2 cm precision defining aperture whose radius was measured in ambient conditions to better that 0.005%. The uncertainty in the defining aperture radius is expanded to 0.03% (type B) due to range in reported thermal contraction of the material from which the aperture is made when the aperture is used at 2.1 K, the operating temperature of the ACRs. The results presented in this paper include only those for which the blackbody defining apertures were 1 mm to 3 mm in diameter and the blackbody cavity was large in surface area in comparison to the area of the defining aperture. For seven of the blackbodies the radii of these apertures were measured at NIST [9] and have a similar uncertainty of 0.03% (type B). The other blackbodies had apertures that carried a higher uncertainty in radius of approximately 0.1% (type B) at their operating temperature as determined from the supplier of the apertures. The defining apertures of the ACRs and blackbodies were positioned coaxial to each other with the aperture planes normal to the axis. With the use of an *in situ* distance measuring device the distance between two reference surfaces close to the defining apertures was measured under test conditions. Ambient dimensional relationships between these reference surfaces and their respective defining apertures, known coefficients of thermal contraction and the *in situ* distance measurement were used to compute the distance between the two defining apertures during calibration. The uncertainty in computed distance between the two defining apertures is typically around 0.04% (type B). Other circular, blackened baffles were placed coaxial with the optical axis to reduce stray radiation and to provide a stable scene for the ACRs. These baffle apertures have radii chosen such that their contribution to uncertainty due to diffraction is not dominant.

The power measured by the ACRs deviates from that expected from strictly geometric optical considerations due to diffraction [10]. Typically there are diffraction losses of the order of 1% of the measured power that result from the relatively small blackbody defining aperture and similarly sized diffraction gains that result from the baffle apertures that are directly illuminated with radiation from the blackbody cavity. These diffraction effects are modelled with a 5% uncertainty (type B) in the corrected value and then corrections are made to the power measurements so that the strictly geometric model can then be used to compute the radiance from the defining aperture of the blackbody. The Stefan–Boltzmann law is then used to compute the radiance temperature for the blackbody cavity.

There were a total of ten blackbodies calibrated that represent six different blackbody designs. The six different designs are labelled design A through F. The contact thermometry for all of the blackbodies was performed with platinum resistance thermometers (PRTs). Blackbody designs A through C all used calibrated PRTs. The resistance measurement accuracy of most of the temperature controllers used to control the blackbody cavity temperature was checked against a NIST-calibrated standard resistor and was typically found to be a few milliohms (~5 mK). Calibrated PRTs have a typical uncertainty of 0.025 K (type B) and uncalibrated PRTs can be expected to have an uncertainty of 0.5 K (type B). The confidential nature of most blackbody calibrations prevents a detailed display of all of their performances. Only nine of the blackbodies are discussed in this paper.

Table 2. Blackbody control temperature setpoint and radiometric temperature measurement results.

Temperature setpoint	Radiometric temperature	Temperature error
300.000 K	305.735 K	5.735 K

Table 3. Contributors to type B uncertainty.

Optical geometry	ACR non-equivalence	Diffraction	Total
0.071 K	0.054 K	0.046 K	0.100 K

Table 4. Total uncertainty computation. The expanded uncertainty is at the 95% confidence level [11].

Type B	Type A	Combined u_c	Expanded U
0.100 K	0.207 K	0.230 K	0.68 K

As an informative demonstration some calibration details are provided for the blackbody designated design D. The values shown for this particular blackbody (tables 2 through 4) were determined at a blackbody cavity temperature of 300 K. Its values are then compared to the range of values spanned by all the others at a blackbody cavity temperature of 300 K. The temperature error of 5.735 K exhibited by the design D blackbody is much larger than the 0.5 K expected uncertainty in accuracy for the uncalibrated PRTs used on this blackbody cavity. It is unlikely that this error was caused by a faulty PRT because the second PRT on the cavity deviated by only 0.5 K to 0.4 K from the first PRT over the 180 K to 620 K over which this blackbody was calibrated. Table 3 shows the individual components of uncertainty that arise from systematic effects that contribute to the total type B uncertainty for the calibration. The optical geometry was very well known for this particular blackbody because of its particular design and therefore its contribution to uncertainty is at the low end of the 0.071 K to 0.230 K range demonstrated by the entire set of blackbodies. The uncertainty due to the ACR non-equivalence was measured for the ACR used in the calibration and is typical of the other ACRs. The contribution due to the diffraction correction uncertainty of the LBIR baffles installed between the customer's blackbody and the ACR is always designed to be minimal. This makes the diffraction correction uncertainty mostly dependent on the blackbody assembly design. For the entire set of blackbodies the diffraction correction uncertainty (type B) ranged from 0.046 K to 0.210 K. The uncertainty is conservatively taken as 5% of the correction; thus designing the blackbody to have low diffraction effects minimizes the diffraction correction uncertainty. The total type B uncertainty for the entire set of blackbodies ranged from about 0.10 K to 0.30 K. The design D blackbody showed an unusually

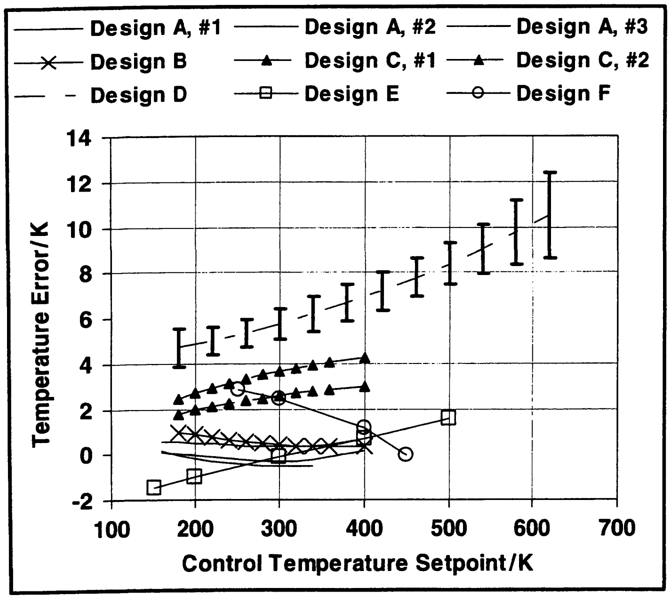

Figure 1. Blackbody cavity radiance temperature calibration data. The temperature error is equal to the measured radiance minus the contact temperature. The error bars shown for design D are the expanded uncertainties at the estimated 95% confidence level and are larger than that for most of the others. The error bars for the others were removed for clarity.

large type A uncertainty of 0.207 K (table 4). All the other blackbodies exhibited type A uncertainties of 0.070 K or less, with most of them around 0.020 K. The combined uncertainty shown in table 3 is the square root of the sum of the squares of the type A and type B uncertainties. The expanded uncertainty shown in table 4 is at the 95% confidence level and was determined and reported using standard guidelines [11]. The expanded uncertainty was determined by multiplying the combined uncertainty by a coverage factor derived from the fitting that is used to create a continuous calibration curve that spans the range of temperatures calibrated [1]. The coverage factor, k, and the degrees of freedom, ν, are different for each blackbody based on the number of data points taken for each blackbody and the order of the polynomial used to generate the best fit calibration curve. The expanded uncertainty for the approximate confidence interval of 95% for this set of blackbodies ranged from 0.24 K to 0.86 K.

The various sources of uncertainty were covered in detail to show that experimental error is not capable of explaining the temperature errors observed. Figure 1 shows the absolute calibration blackbody cavity temperature error for the nine blackbodies reported. The temperature error is equal to the measured radiance temperature minus the contact temperature. The radiance temperature of six out of nine of the blackbodies deviated by more than the uncertainty for the 95% confidence level at some point in their calibrated temperature range. The three design A blackbodies showed the best calibration results and had characteristics typically associated with accurate blackbodies. The design A blackbodies had large cavities with respect to the exit aperture size, were thermally massive and slow to change in temperature. All the other blackbodies tended to be lighter and smaller in design, favourable for trying to reduce power consumption and settling time but usually resulting in lower temperature accuracy. In addition, the type of PRTs used in the design A blackbodies were selected based on the recommendation of an expert in PRT technology. This expert's recommendation was based on the blackbody's planned operational environment. This careful selection of PRT type was not done for designs B and C and the selection process is not known for designs D, E and F. Figure 1 also shows the effects of other issues. For example, the two design C cavities deviate by about 1 K from each other over their operating range even though they were calibrated one after the other in identical test configurations and both used calibrated temperature sensors that should have agreed with each other to within 0.025 K. For this particular blackbody cavity it is known that the platinum elements in the PRTs are packed solidly in powder and are susceptible to strain because the PRTs are potted into the blackbody cavity block with a ceramic epoxy. Consultation with experts in the field of PRT technology indicates that strain on the PRT sensors and their leads can result in temperature measurement errors of the order of several kelvin. It is speculated that strain is the most likely cause of the temperature difference exhibited between these two blackbodies.

Figure 1 shows that there are no systematic sources of error that scale with temperature or power. The Stefan–Boltzmann law dictates that if there is a source of systematic power measurement error that scales as a constant percentage of power then there will be a systematic temperature error that scales as one-fourth of that percentage in temperature. Sources of such power measurement errors include defining aperture area errors, diffraction correction errors, errors in the distance between the defining apertures and stray reflection errors. If such sources of error existed and contributed significantly to the measured power, the data as plotted in figure 1 would appear as a straight line that intersected the origin. Most of the plotted lines are not linear and none of them can be reasonably extrapolated to intersect the origin. Simple temperature errors that might come from an uncalibrated PRT, for instance, also have this signature. Due to the nearly linear relationship between resistance and temperature of a PRT, a simple sensor temperature error also appears as a straight line that intersects the origin. Again, none of the blackbodies shows the signature of a simple temperature sensor error.

Evidence of poor blackbody performance does not need to be determined by radiometric calibration. In a particular blackbody, in order to increase the testing rates in an industrial process, the walls of the blackbody cavity were thinned to reduce the blackbody temperature settling time. However, the walls conducted so little heat that the cavity temperature as measured by the contact thermometry dropped by more than 0.3 K when the blackbody shutter was opened. In this same blackbody, two PRTs at dissimilar locations showed a temperature difference that varied from 0 K at a blackbody cavity temperature of 180 K to 6 K at a blackbody cavity temperature of 400 K. Further evidence of errors induced by mounting-related strain comes from blackbodies that contained two calibrated PRTs in close proximity to each other. In this case two co-located PRTs should agree to within 0.035 K. All the five blackbodies with this design feature showed a temperature reading difference of at least 0.4 K at some point in their operating temperature range. Note that this temperature difference is only an indication that there is a temperature sensor mounting issue and does not indicate the magnitude of the temperature error.

5. Conclusion

The ACRs used by the LBIR facility as absolute infrared detector standards were intercompared with the primary optical power measurement standard at NIST, Gaithersburg. The results showed that the LBIR ACRs performed well as absolute detector standards and contributed minimally to the total uncertainty of the blackbody calibrations.

Results from the calibrations of nine blackbodies that represented six different blackbody designs show an informative pattern of blackbody behaviour. First, there was not a single blackbody where the radiance temperature agreed to within 0.3 K of the contact thermometry over its entire operating temperature range. To within the uncertainty at the estimated 95% confidence level, agreement is achieved for most of the calibrated temperature range for only three of the nine blackbodies. A detailed discussion of the various sources of known errors was provided to support the calibration accuracy statements. Various symptoms of common sources of systematic errors were described and shown not to be significant contributors. The blackbody source design that performed the best had a blackbody cavity with a large thermal mass and long time response, the largest ratio of cavity surface area to exit aperture area, and PRTs that were chosen by an expert in the field of PRT technology. At this time it is believed that strain on the PRT sensors due to mounting methods is the primary cause of differences between the radiometric and contact thermometry. To investigate this, the LBIR facility is developing the capability to recalibrate PRTs after they are permanently mounted in the blackbody cavities to see if the agreement between contact and radiometric thermometry can be improved.

Acknowledgments

The authors are grateful to Joseph P Rice for assistance in calibration of the QED-150 optical trap detector against the primary national optical power measurement standard.

References

[1] Datla R U, Croarkin M C and Parr A C 1994 *J. Res. Natl Inst. Stand. Technol.* **99** 77–87
[2] Foukal P V, Hoyt C, Kochling H and Miller P 1990 *Appl. Opt.* **29** 988–93
[3] Datla R U, Stock K, Parr A C, Hoyt C C, Miller P J and Foukal P V 1992 *Appl. Opt.* **31** 7219–25
[4] Carter A C, Lorentz S R, Jung T M and Datla R U 2005 *Appl. Opt.* **44** 871–5
[5] Gentile T R, Houston J M and Cromer C L 1996 *Appl. Opt.* **35** 4392–403
[6] Fedchak J A, Carter A C and Datla R U 2006 *J. Res. Natl Inst. Stand. Technol.* submitted
[7] Carter A C, Jung T M, Smith A W, Lorentz S R and Datla R U 2003 *Metrologia* **40** S1–4
[8] Lorentz S R and Datla R U 1993 *Metrologia* **30** 341–4
[9] Fowler J and Litorja M 2003 *Metrologia* **40** S9–12
[10] Shirley E L and Terraciano M L 2001 *Appl. Opt.* **40** 4463–72
[11] 1997 *US Guide to the Expression of Uncertainty in Measurement* ANSI/NCSL Z540-2-1997, prepared by the Accredited Standards Committee on General Requirements for Calibration and Measuring and Test Equipment, writing group Z540-2