metrology

MDPI

Article

Metrological Validation of Low-Cost DS18B20 Digital Temperature Sensors Using the TH-001 Procedure: Calibration Models, Uncertainty, and Reproducibility

Juan Antonio Rodríguez-Rama [1,2,3,*], Leticia Presa Madrigal [2,3], Alfredo Marín Lázaro [3,4], Javier Maroto Lorenzo [3,5], Ana García Laso [2,3], Jorge L. Costafreda Mustelier [2,3] and Domingo A. Martín-Sánchez [2,3]

1 Grupo de Investigación en GeoEnergía, Departamento de Recursos Geológicos Para La Transición Ecológica, Instituto Geológico y Minero de España, Consejo Superior de Investigaciones Científicas (CN IGME-CSIC), 28006 Madrid, Spain
2 Departamento de Ingeniería Geológica y Minera, Escuela Técnica Superior de Ingenieros de Minas y Energía, Universidad Politécnica de Madrid (ETSIME-UPM), 28003 Madrid, Spain
3 Unidad de Emprendimiento Social, Ética y Valores en la Ingeniería (UESEVI-ETSIME-UPM), Comunidad EELISA Ethics, Social Commitment & Entrepreneurship (EELISA-ESCE), Universidad Politécnica de Madrid (UPM), 28003 Madrid, Spain
4 Laboratorio de Metrología Térmica, Departamento de Ingeniería Energética, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid (ETSII-UPM), 28003 Madrid, Spain
5 Unidad de Tecnologías de la Información y la Comunicación (UTIC), Escuela Técnica Superior de Ingenieros de Minas y Energía, Universidad Politécnica de Madrid (ETSIME-UPM), 28003 Madrid, Spain
* Correspondence: juan.rodriguez@igme.es

Abstract

This study presents the metrological validation of encapsulated DS18B20 digital temperature sensors. Eight units were tested, and seven were analysed (sensor 8 was excluded owing to a systematic failure). The evaluation was performed using a standard comparison calibration, where T_{ref} was defined as the mean of two calibrated Pt-100 probes in a Julabo DYNEO DD 601F thermostatic bath, following the TH-001 procedure of the Spanish Centre of Metrology (CEM). Four validation tests were performed: Test 1 (E1, 20 to 75 °C), Test 2 (E2, 20 to 72 °C), and with an extended range, Test 3 (E3, −12 to 86 °C) and Test 4 (E4, −12 to 86 °C; repetition to assess reproducibility relative to E3), with 10 steady-state readings per setpoint. Erroneous readings were defined and removed (probe 3, Test 4), and set points without valid readings from probe 4 above 68 °C were excluded. Without data processing, the errors were consistent with the manufacturer's stated ±0.5 °C, despite an inter-probe bias. Several correction models were evaluated (offset, affine linear, polynomial, and segmented); the probe-specific affine linear model provided the best overall compromise, reducing MAE (Mean Absolute Error) to 0.046 to 0.130 °C and RMSE (Root Mean Square Error) to 0.057 to 0.169 °C. The process uncertainty is dominated by the traceability of the Pt-100 probes and the effective nonuniformity of the isothermal volume, which limits the achievable accuracy. The results support the use of individually calibrated DS18B20 sensors for continuous monitoring, provided that the effective operating range is maintained.

Keywords: low-cost sensors; DS18B20; temperature monitoring; uncertainty; IoT

Check for updates

Academic Editors: Han Haitjema, Markus Bär and Samanta Piano

Received: 5 January 2026
Revised: 1 March 2026
Accepted: 17 March 2026
Published: 23 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The use of low-cost sensors (LCS) is transforming numerous scientific and technological fields, particularly environmental monitoring, preventive conservation of cultural

heritage, smart resource management, and the development of data networks across academic and urban infrastructures [1–4]. Their affordability, together with their versatility for implementation in scalable, real-world solutions, has made these devices key enablers for democratising access to scientific instrumentation and increasing the spatial density of conventional measurement systems [5–7].

Nevertheless, their widespread adoption raises important metrological challenges that must be addressed. Key limitations include high unit-to-unit and batch-to-batch variability, the presence of systematic errors in the absence of individual calibration, and the lack of internationally recognised traceability procedures [8,9]. These factors hinder their use in projects that require accurate and reproducible measurements, particularly in those in which data reliability is critical for scientific analysis, decision-making, or process certification [10,11].

Recent literature reflects this concern and proposes different statistical correction strategies and calibration models to improve the accuracy of the LCS. Abdinnoor et al. [12] show that calibrating DS18B20 sensors under controlled conditions can reduce the mean error from 3% to 0.85% and decrease the standard error by approximately 48% through linear correction models and appropriate handling of experimental uncertainty. Lin et al. [13] emphasise that linear regression is the preferred approach for field calibration, provided that environmental factors and sources of uncertainty are addressed explicitly. However, most of these studies focus on statistical approaches, and there remains a scarcity of work incorporating traceable metrological validation protocols and direct calibrations against certified standards [14–16].

In this context, the Tellus UPM Innovation Ecosystem, embedded within the Unit for Social Entrepreneurship, Ethics, and Values in Engineering (UESEVI) at the School of Mines and Energy Engineering of the Technical University of Madrid (ETSIME-UPM), operates as a multidisciplinary laboratory aimed at developing and validating technological solutions for sustainability, education, and social innovation [17,18]. Through the Tellus IoT Lab, this ecosystem provides the infrastructure and experimental framework for the design, characterisation, and verification of microelectronic prototypes and low-cost sensors with a strong emphasis on knowledge transfer. The European Engineering Learning Innovation and Science Alliance (EELISA), particularly its Ethics, Social Commitment & Entrepreneurship (ESCE) community [19], has enabled the inter-institutional collaboration required to structure joint work between teams with complementary technological and metrological expertise.

In this context, a collaboration with the Thermal Metrology Laboratory of the Department of Energy Engineering at the School of Industrial Engineering of the Technical University of Madrid (ETSII-UPM) enabled the implementation of the TH-001 procedure of the Spanish Centre of Metrology (CEM) [20] for the comparison calibration of direct-reading thermometers. The application of calibration models, strictly aligned with internationally recognized normative frameworks such as the International System of Units (SI) [21], the International Vocabulary of Metrology (VIM) [22], and the Guide to the Expression of Uncertainty in Measurement (GUM) alongside EA-4/02 M:2013 [23], establishes a traceable methodology for uncertainty estimation and provides reproducible criteria for the metrological validation against certified Pt-100 standards. This approach adds value relative to the previous literature by integrating metrological rigor into the use of LCS and, moreover, by enabling the identification and selection of the most stable and calibratable units within each batch.

This study reports the results of four calibration and metrological validation tests of DS18B20 sensors [24] conducted under the TH-001 protocol and within a university Living Lab environment [25,26]. The operational range of interest was defined as −10 °C

to +85 °C, in accordance with the most stringent specifications stated in the datasheet for these sensors.

The objective of this study was to demonstrate that, through the application of traceable calibration protocols, statistical quality control, and independent validation under the TH-001 methodology, DS18B20 sensors can achieve levels of accuracy and stability compatible with scientific, engineering, and advanced monitoring applications in IoT networks, heritage conservation, and sustainability.

The main contributions of this study are as follows: (i) the definition and implementation of a replicable DS18B20 calibration protocol grounded in international metrological standards; (ii) a systematic evaluation of different correction models, identifying the affine linear model as the optimal solution in terms of accuracy and robustness; (iii) a detailed sensor-by-sensor uncertainty estimation, including both Type A and Type B components; and (iv) demonstration of the practical transferability of the procedure to IoT networks deployed in a university Living Lab.

2. Materials and Methods

2.1. Instrumentation and Reference Standards

The sensors under calibration comprised eight encapsulated DS18B20 units, individually wired and configured to a 12-bit resolution (0.0625 °C per increment), and operated within the manufacturer's specified highest-accuracy interval (±0.5 °C from −10 to +85 °C). The reference system for calibration was based on a Julabo DYNEO DD thermostatic bath [27] (manufacturer-declared stability ±0.01 °C; resolution 0.01 °C) used to generate isothermal setpoints. The temperature was verified using two Pt-100 probes calibrated against standards traceable to the Spanish Centre of Metrology (CEM), connected to a high-resolution digital reader (0.001 °C). These probes were used as redundant working standards to assess the temporal stability and spatial uniformity of the bath at each setpoint.

2.2. Environmental Conditions

The experiments were conducted in the Thermal Metrology Laboratory of the Department of Energy Engineering at ETSII-UPM under environmental conditions continuously monitored using a Testo Saveris 2 H1 data logger (range: −30 to +50 °C; 0–100% RH; accuracy: ±0.5 °C and ±2% RH; resolution: 0.1 °C and 0.1% RH). This monitoring enabled verification, within the instrument's stated accuracy, that no relevant environmental anomalies occurred during the experimental period that could compromise the metrological validity of the tests. In accordance with the TH-001 procedure, temperature and relative humidity were recorded within the laboratory's normal operating conditions; during the experiments, values typically remained within 20–25 °C and 30–50% RH, without significant fluctuations.

2.3. Experimental Setup

The eight DS18B20 probes and two Pt-100 reference probes were immersed in a thermostatic bath at a uniform depth to minimise vertical thermal gradients. To ensure maximum data integrity and avoid bus contention, each DS18B20 sensor was connected to a dedicated digital input pin on the Arduino microcontroller, avoiding a multidrop topology with a shared bus. Furthermore, the sensors were configured in active power mode (using a common VCC and GND bus with a dedicated 4.7 kΩ pull-up resistor per data line) instead of parasitic power mode. This setup was chosen to prevent self-heating issues and voltage drops during simultaneous temperature conversions, common limitations in shared parasitic lines, and to reduce communication errors associated with line capacitance.

The probes were arranged in a dedicated holder that ensured adequate separation and reproducible positioning, preventing probe-to-probe contact and potential local disturbance. The typical cable lengths were <1 m, consistent with short-range 1-Wire configurations. The calibration points were strategically selected to cover the manufacturer's full high-accuracy range (−10 to +85 °C), ensuring the correction models remain valid under the diverse thermal conditions that an IoT edge node might encounter.

In this regard, the validation supports the direct applicability of these sensors across multiple ongoing deployments, such as subsurface environmental monitoring (MESEME project) [28], preventive conservation of historical university library collections (Smart Heritage ETSIME-UPM project) [29], and industrial applications like concrete setting studies (Smart Concrete project) [30], in collaboration with the Quality Control Laboratory for Construction Products (LOEMCO).

The experimental setup, including the detailed electronic interface and the physical distribution of the sensors, is shown in Figure 1, which illustrates the thermostatic bath, the Pt-100 based reference system, the Arduino UNO microcontroller, and the control computer running LabVIEW [30].

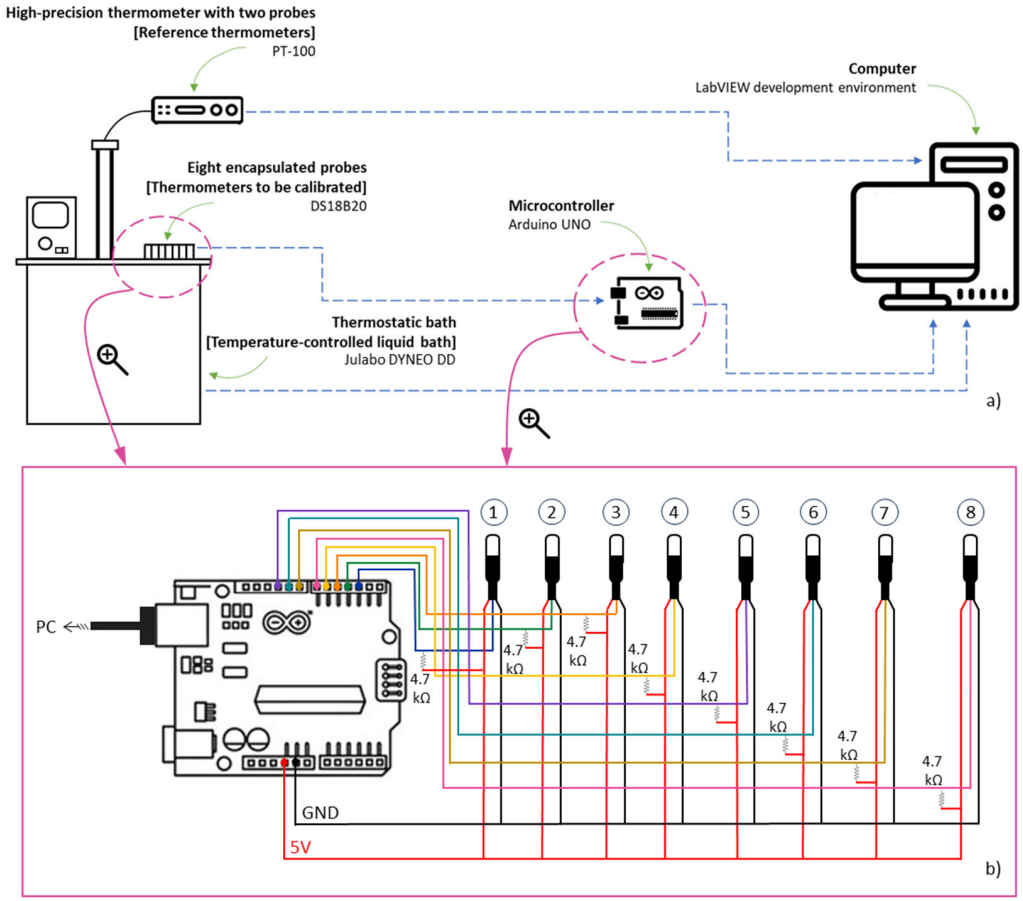


Figure 1. Detailed architecture of the experimental setup. (**a**) Functional schematic illustrating the interconnection between the thermostatic bath, the Pt-100 reference standards, and the sensor batch. (**b**) Electronic interface diagram of the Arduino-DS18B20 configuration: the setup of dedicated digital pins (D2–D9) and independent pull-up resistors (4.7 kΩ) is detailed. This design ensures signal integrity and fault isolation, preventing systematic anomalies (such as that observed in Sensor 8) from affecting the remainder of the batch. Cabling lengths of less than 1 m ensure electrical stability and signal clarity against capacitive effects. Furthermore, the integration with the control computer is shown; it simultaneously manages data acquisition and the circuit's power supply, thereby ensuring full process traceability.

2.4. Acquisition and Processing

Data acquisition was performed on a laboratory computer using a LabVIEW application, which enabled the synchronisation of the readings from the DS18B20 sensors (via the Arduino UNO) and Pt-100 probes (via a digital thermometer). The data were logged in real time and stored in a .txt file for subsequent analyses.

Data pre-processing included: (i) computation of the reference temperature, T_{ref}, as the arithmetic mean of the two Pt-100 readings, in accordance with the TH-001 procedure; (ii) filtering of anomalous values; and (iii) verification of the stability and uniformity of the isothermal medium at each set point. Once stable conditions were achieved, ten readings were recorded per temperature level for subsequent statistical analysis.

2.5. Calibration Procedure

The TH-001 comparison calibration procedure for direct-reading thermometers was used. It comprised:

- Selection of calibration points within the target nominal range (−10 to +85 °C): a 5 °C step was used in the first test and a 2 °C step in subsequent tests to increase point density and improve model fitting. Extended tests were performed using additional points outside the nominal range.
- Bath stabilisation was performed at each setpoint until stable conditions were reached (variation in $T_{\text{ref}} < 0.02$ °C over 5 min), where T_{ref} is the mean of the two Pt-100 readings.
- Recording of synchronised reading series: Once stability was achieved, 10 readings were acquired per temperature level for subsequent statistical treatment.
- The individual correction for each sensor is computed (Equation (1)), defined as the difference between the reference and the sensor under calibration:

$$C = T_{\text{ref}} - T_{\text{DS}} \tag{1}$$

  Therefore, the corrected temperature is obtained as $T_{\text{corr}} = T_{\text{DS}} + C$. In practice, C in equation, is modelled as a function of T_{DS} for affine correction implies $T_{\text{corr}} = a + bT_{\text{DS}}$ (equivalently $C = a + (b-1)T_{\text{DS}}$).
- Fitting of correction models (Section 2.7).

2.6. Uncertainty Estimation

Uncertainty evaluation was performed in accordance with the TH-001 procedure, considering a comparison calibration model in which the reference temperature was defined as $t_{\text{ref}} = \frac{t_1+t_2}{2}$, where t_1 and t_2 were the readings from the two Pt-100 probes. The contributions considered (standard uncertainties) were as follows:

- u_{res}: uncertainty associated with the DS18B20 resolution (Equation (2)), assuming a rectangular distribution.

$$u_{\text{res}} = \frac{0.0625}{\sqrt{12}} = 0.018\ ^{\circ}\text{C} \tag{2}$$

- u_{rep}: repeatability of the DS18B20 at each setpoint, calculated as the standard deviation of the readings acquired under stable conditions (ten readings).
- u_{estab}: temporal stability of the isothermal medium during acquisition, estimated as the standard deviation of T_{ref} (mean of the Pt-100 readings) over the stable interval at each setpoint.
- u_{uniform}: As an independent spatial uniformity map of the bath was not available, a conservative approach was adopted in this study to represent the nonuniformity

of the isothermal medium within the calibration volume. Rather than directly estimating u_{uniform} (spatial uniformity of the isothermal medium) (Equation (3)), an effective non-uniformity of the calibration volume (u_{vol}) was defined and used. It is computed from the observed difference between the two Pt-100 probes placed at different positions within the bath under the same stabilisation conditions. This difference, $\Delta T = |T_{\text{Pt-100,1}} - T_{\text{Pt-100,2}}|$, is interpreted as a representative upper bound on the maximum thermal gradient within the volume occupied by the full assembly (DS18B20 probes + reference), and it is converted to a standard uncertainty by assuming a rectangular distribution (conservative case), such that:

$$u_{\text{uniform}} = \frac{\Delta \text{T}}{2\sqrt{3}} \tag{3}$$

This approach fulfils the intent of TH-001, as it explicitly introduces into the uncertainty budget a term that captures the effects of spatial gradients and potential thermal heterogeneities of the isothermal medium during the calibration process, thereby avoiding an underestimation of uncertainty when a full spatial characterisation of the bath is not available.

- $u_{\text{cert,Tref}}$: component arising from the Pt-100 calibration certificate. From the expanded uncertainty reported in the certificate, U_{cert} (with k = 2), the standard uncertainty of each Pt-100 was obtained as $u_{\text{cert},i} = \frac{U_{\text{cert}}}{2}$ (Equation (4)). As T_{ref} is defined as the mean of two independent standards, this contribution is propagated as follows:

$$u_{cert,T\text{ref}} = \frac{u_{cert,i}}{\sqrt{2}} \tag{4}$$

Other contributions considered in TH-001 (drift and influence quantities) were assumed to be either encompassed by the calibration certificate or negligible relative to the dominant components under the test conditions.

The combined standard uncertainty was calculated as follows Equation (5):

$$u_{\text{c}} = \sqrt{u_{\text{res}}^2 + u_{\text{rep}}^2 + u_{\text{estab}}^2 + u_{\text{uniform}}^2 + u_{\text{cert,Tref}}^2} \tag{5}$$

The expanded uncertainty is expressed as follows (Equation (6)):

$$U = k\, u_c, k = 2 \text{ (approx. 95\%)} \tag{6}$$

2.7. Evaluation of Correction Models

Although linear models are the most commonly used in the literature [31–33], this study evaluated four models with different levels of complexity to identify the best trade-off between computational simplicity and goodness of fit. This is particularly relevant for solutions in which such adjustments are implemented directly on IoT nodes with limited processing power. In all cases, the corrected temperature is denoted by T_{corr}.

The following models were evaluated:

2.7.1. Constant Offset

A uniform correction was applied to all readings to compensate for potential systematic bias relative to the reference (Equation (7)).

$$T_{\text{corr}} = T_{\text{DS}} + c_0 \tag{7}$$

2.7.2. Linear Model

This model assumes a linear relationship between the sensor reading and reference, allowing the correction of both offset and slope errors (Equation (8)).

$$T_{\text{corr}} = a + b\,T_{\text{DS}} \tag{8}$$

Although traceability in thermometry is underpinned by normative frameworks such as the ITS-90 [34] and, for resistance thermometers, by resistance-temperature relationships, the present work performs a direct correlation between the measured and reference temperatures. Within the considered range, the linear fit is regarded as an adequate approximation for deriving operational corrections.

2.7.3. Second-Order Polynomial Model

This model introduces a quadratic term to capture potential non-linearities within the tested range (Equation (9)).

$$T_{\text{corr}} = a + b\,T_{\text{DS}} + c\,T_{\text{DS}}^2 \tag{9}$$

2.7.4. Segmented Model (Three Intervals)

This approach considers independent regressions over predefined intervals to improve the fit in regions where more pronounced deviations are observed (Equation (10)), consistent with the segmentation applied in the Results section: $[-12, 0)$, $[0, 60)$, and $[60, 86]$ °C. The model was defined as a piecewise function over these three distinct intervals.

$$T_{\text{corr}} = \begin{cases} a_1 + b_1\,T_{\text{DS}}, & -12 \le T_{\text{DS}} < 0 \\ a_2 + b_2\,T_{\text{DS}}, & 0 \le T_{\text{DS}} < 60 \\ a_3 + b_3\,T_{\text{DS}}, & 60 \le T_{\text{DS}} \le 86 \end{cases} \tag{10}$$

Model evaluation was performed using MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) as performance metrics, complemented by a residual analysis to assess error structure (distribution, residual bias, and potential unmodelled patterns). This allowed the evaluation of the accuracy, robustness, and practical applicability of each model.

2.8. Assessment of Inter-Test Reproducibility

Reproducibility was assessed across four independent tests by analysing the consistency of the mean bias and its variability between experiments for each probe, together with an evaluation of the overall error metrics (MAE/RMSE) across tests. Because the tests covered different thermal ranges, these assessments were performed consistently by scenario (mid-range versus extended range) and, where appropriate, between equivalent tests (Tests 3 and 4). This analysis enabled the identification of sensors exhibiting stable behaviour across tests and the detection of potential trends consistent with systematic drift, providing practical criteria for reliable deployment in IoT networks and long-term applications.

2.9. Experimental Design, Sample Size, and Data Cleaning

The tests were designed to provide progressive coverage from a mid-range temperature interval, representative of environmental monitoring applications, to an extended range within the interval of interest (approximately −10 to 85 °C), incorporating set points outside the nominal range.

- Test 1 (mid-range, 5 °C step): exploratory in nature, aimed at assessing the overall behaviour of the batch over a range representative of typical applications.

- Test 2 (densified mid-range, 2 °C step): increased point density to examine the linearity and temperature dependence of the error in greater detail.
- Test 3 (extended range, 2 °C step): included thermal conditions close to the interval of interest and additional points to investigate the batch behaviour towards the tested extremes.
- Test 4 (repeat of the extended range): Test 3 protocol was replicated to assess inter-test reproducibility and explore potential trends consistent with drift.

At each temperature setpoint, sufficient time was allowed to record approximately ten consecutive readings per sensor during the steady-state interval used for data acquisition, yielding a sample size of nearly 10,000 records for the complete experiment across the seven probes. This volume of data enables robust statistical analyses, estimation of repeatability across temperature setpoints, and assessment of subtle error trends as a function of the temperature.

Prior to statistical treatment, the dataset underwent a basic quality review.

- The internal consistency was verified for each setpoint (i.e., no appreciable drift during the steady-state acquisition interval).
- Only records associated with transient bath phases or isolated acquisition errors (incomplete readings or values markedly inconsistent with the remainder of the setpoint) were identified and discarded.
- No significant data losses were observed, except for probe 4 at setpoints above approximately 68 °C, where the sensor ceased to provide reliable readings; therefore, these points were excluded from the subsequent analyses for that unit. In addition, probe 8 was excluded from the study because of a systematic failure.

3. Results

This section presents the results of the four tests summarised in Table 1, which were conducted on the batch of DS18B20 sensors described in Section 2. As a prerequisite, the behaviour of the isothermal medium and reference system (stability and uniformity) was assessed. Subsequently, sensor errors, filtering of anomalous readings, and the performance of correction models were analysed. Finally, uncertainty analysis and inter-test reproducibility were integrated.

Table 1. Summary of the thermal ranges, calibration points, and sample sizes for each test.

Test	Nominal Temperature Range (°C)	Nominal Step (°C)	No. of Calibration Points	No. of Records Per Probe
1	20 to 75	5	12	120
2	20 to 72	2	27	270
3	−12 to 86	2	50	510
4	−12 to 86	2	50	510

To ensure clarity across the extensive metrological evaluation presented in this section, the results are structured into five key stages: (i) characterization of the isothermal medium stability; (ii) assessment of the DS18B20 raw performance; (iii) inter-test reproducibility analysis; (iv) evaluation of the four proposed correction models; and (v) final uncertainty estimation and overall performance metrics.

Probe 8 was excluded from the analysis due to a systematic failure, and the data exclusions described in Section 2 were applied.

3.1. Behaviour of the Isothermal Medium and Reference Standards

The correct operation of the isothermal medium and reference system is a prerequisite for rigorously interpreting the behaviour of DS18B20 sensors. This subsection evaluates (i) the temporal stability of the bath during data acquisition, (ii) the spatial uniformity of the isothermal medium based on the Pt-100 standards, and (iii) the metrological consistency of these standards as a traceable reference.

3.1.1. Bath Stability and Spatial Uniformity of the Isothermal Medium

The Julabo DYNEO DD 601F thermostatic bath was operated in the internal control mode. The probes were placed within the isothermal zone characterised by the laboratory (uniform immersion depth and away from the walls) to minimise vertical gradients and boundary effects. At each temperature setpoint, the system was allowed to reach a steady state, and in accordance with the TH-001 procedure, 10 consecutive readings per channel were recorded during the stable interval.

Temporal stability (within set point). The stability during acquisition was evaluated using the standard deviation of T_{ref} at each setpoint, where T_{ref} is the mean of the two Pt-100 readings. In Test 3 (Figure 2), the standard deviation of T_{ref} remained on the order of thousandths of a degree across the range, with no appreciable trends within the steady-state temperature window, which is consistent with the stable behaviour of the isothermal medium during data collection.

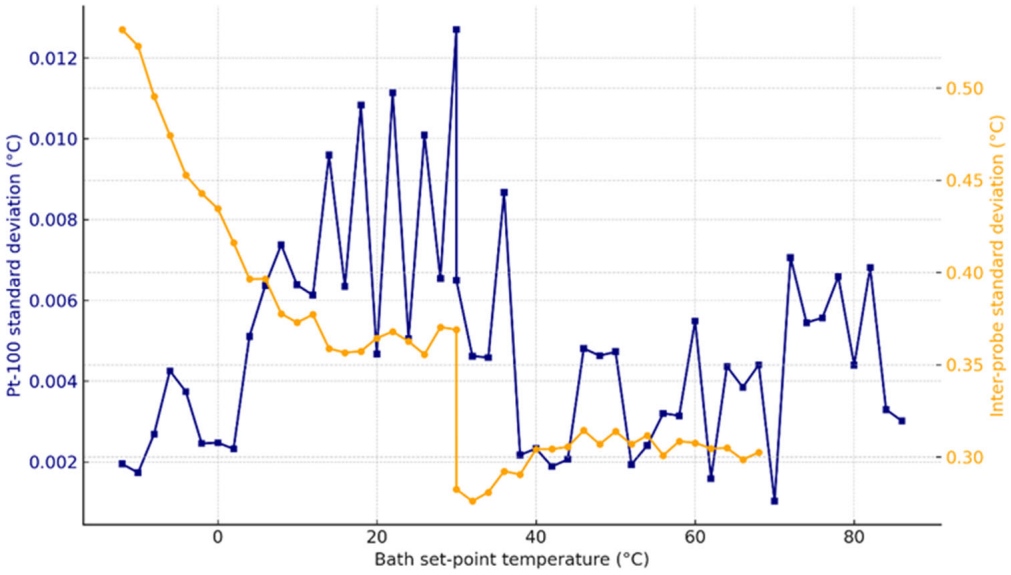


Figure 2. Stability (SD of T_{ref}, left axis) and inter-probe variability of DS18B20 relative to T_{ref} (right axis) for Test 3.

In Test 4, the maximum standard deviation (SD) of T_{ref} (0.0673 °C) corresponded to a setpoint for which the calculation included part of the transient phase. When the SD was recomputed over the steady-state temperature window, the indicator was consistent with the remaining setpoints.

During the preliminary analysis, an anomalously high standard deviation of the Pt-100 readings at 30 °C was detected because 20 readings were recorded (including part of the transient). When the calculation was restricted to the 10 readings from the stable interval, the value returned to levels consistent with the other set points. Hence, the stability indicators were computed exclusively over the steady-state window in accordance with TH-001.

In this study, the spatial uniformity was estimated from the difference between the two Pt-100 probes placed at different positions within the bath. This difference provides

an experimental bound on the nonuniformity within the measurement zone and allows its contribution to be expressed as a standard uncertainty under a rectangular assumption (Equation (11)).

$$u_{uniform} = \frac{|\, t_1 - t_2 \,|}{2\sqrt{3}} \tag{11}$$

where t_1 y t_2 are the corrected readings of the Pt-100 probes at the steady-state setpoint. This approach is consistent with the treatment of stability and uniformity of the isothermal medium in TH-001.

In addition, the dispersion of the residuals of the seven DS18B20 probes with respect to T_{ref} was computed for each setpoint. This quantity is reported as the batch's inter-probe variability (rather than bath uniformity), since it integrates the intrinsic dispersion of the sensors under calibration together with minor effects related to positioning/immersion.

In the extended-range tests (Tests 3 and 4), this variability increased towards the thermal extremes, consistent with the larger DS18B20 deviations near the boundaries of the extended interval of interest.

The bath-uniformity contribution used in the uncertainty calculation was obtained from $|t_1 - t_2|$ between the Pt-100 probes, whereas the "inter-probe SD" was used here to describe the batch dispersion prior to correction.

3.1.2. Performance of the Pt-100 Standards

Two laboratory Pt-100 probes were used as reference thermometers, calibrated with traceability to the national realisation of the ITS-90 (Spanish version: EIT-90) maintained by the Spanish Centre of Metrology (CEM). The calibration certificate provides, for the tested points, the correction to be applied and an expanded certificate uncertainty of U_{cert} = 0.14 °C with k = 2, estimated in accordance with EA-4/02 M:2013. Accordingly, the standard uncertainty associated with the certificate for each Pt-100 was taken as $u_{\text{cert},i} = \frac{U_{\text{cert}}}{k} = 0.07$ °C. Because the reference temperature in this study was defined as $T_{\text{ref}} = \frac{t_1+t_2}{2}$ (mean of both PT-100 readings after applying the certificate corrections), the certificate contribution to the reference was propagated as $u_{\text{cert,Tref}} = \frac{u_{\text{cert},i}}{\sqrt{2}}$, assuming independence between standards.

In all tests, T_{ref} was computed as the mean of the two Pt-100 readings over the steady-state acquisition interval. The differences between the Pt-100 probes remained small (on the order of hundredths of a degree) and showed no systematic trends across the thermal range, confirming their mutual consistency and supporting their suitability as a traceable reference against which deviations in the DS18B20 batch were evaluated and corrected.

3.2. Behaviour of the DS18B20 Sensors

The behaviour of the DS18B20 sensors was analysed based on their deviations relative to the reference temperature T_{ref} (mean of two Pt-100 probes), both in absolute terms and through an analysis of the residuals and their statistical distribution. Performance was assessed prior to applying the correction models by examining batch consistency, the presence of outliers, and the degree of compliance with the manufacturer's specifications within the range of interest. With this convention, $r \equiv C$, and $T_{\text{corr}} = T_{\text{DS}} + r$.

Figure 3 shows the relationship between the reference temperature T_{ref} (mean of the two Pt-100 probes) and the temperature recorded by the DS18B20 sensors across the four tests (cleaned data). In this representation, each point corresponds to the mean of the 10 steady-state readings recorded at a given temperature set point for a given probe. Given the wide thermal range considered, the point cloud appears visually close to the ideal 1:1 line; therefore, the performance evaluation relies primarily on residual analysis and error metrics (Figure 4 and Table 2).

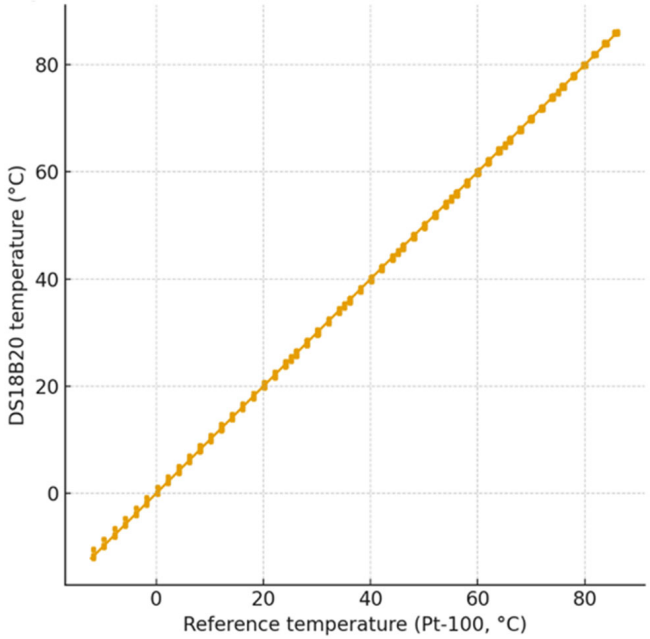


Figure 3. Agreement between the reference temperature (T_{ref}) (mean of two Pt-100 probes) and the temperature recorded by the DS18B20 sensors in Tests 1–4 (filtered data). Each point represents the mean of ten steady-state readings per setpoint and probe. The approximating line corresponds to an ideal 1:1 relationship. Spurious readings identified for probe 3 in Test 4 and setpoints above 68 °C with no valid readings from probe 4 were excluded from the analysis.

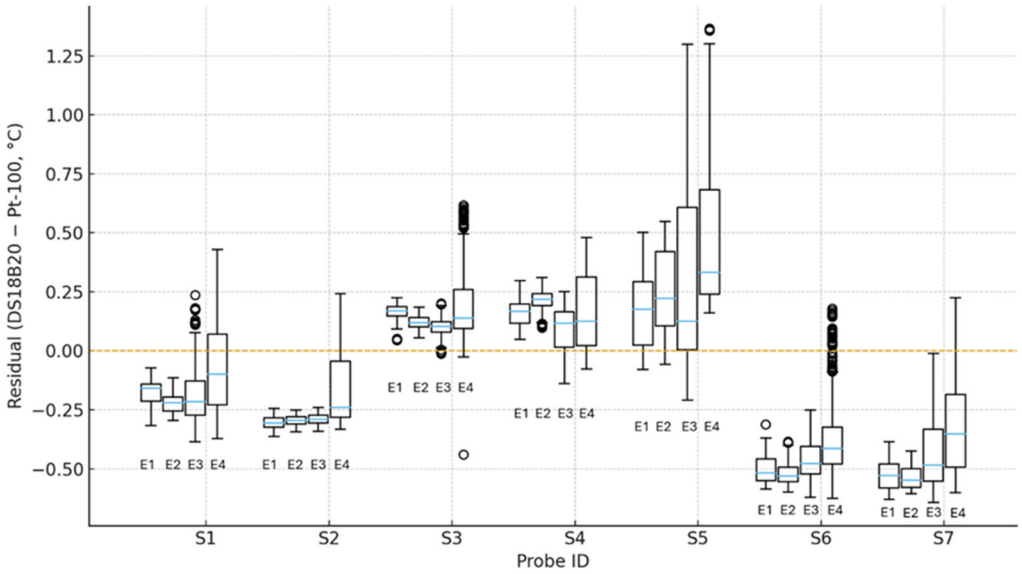


Figure 4. Distribution of residuals by probe (S1–S7) and test (E1–E4). The box boundaries represent the interquartile range (IQR), from the first to the third quartile (Q_1 and Q_3); the blue horizontal line indicates the median residual, representing the central tendency or bias; the whiskers (error bars) extend to the minimum and maximum values within 1.5 × IQR; and the small circles represent outliers (data points beyond the whiskers). The dashed yellow line indicates a zero residual, serving as the reference for an ideal measurement.

Figure 4 shows the distribution of the residuals, $r = T_{ref} - T_{DS}$, using boxplots by probe and test, facilitating the analysis of bias and dispersion across experiments. Table 2 summarises the main descriptive statistics for each test (mean error, standard deviation, MAE, RMSE, and, where reported, mean relative error).

Table 2. Overall descriptive statistics by test (residuals: $r = T_{ref} - T_{DS}$) after data cleaning.

Test	Temperature Range (°C)	No. of DS18B20 Readings Used	Mean Error (°C)	Standard Deviation (°C)	Mean Absolute Error, MAE (°C)	Mean Relative Error, MRE (%)
1	20 to 75	820	−0.15	0.30	0.29	0.74
2	20 to 72	1860	−0.15	0.32	0.31	0.79
3	−12 to 86	3470	−0.13	0.33	0.28	4.42
4	−12 to 86	3450	−0.02	0.35	0.28	4.42

3.2.1. Descriptive Statistics by Test

The statistical analysis of the four tests indicated an overall near-linear behaviour of the batch and an appreciable inter-probe dispersion, as expected for low-cost sensors. In Tests 1 and 2 (mid-range: 20 to 75 °C and 20 to 72 °C), the batch exhibited a moderate overall mean bias (on the order of −0.15 °C), with MAE typically between 0.15 and 0.30 °C and standard deviations on the order of 0.15 to 0.30 °C. In Tests 3 and 4 (extended range: −12 to 86 °C), the mean error remains of similar or smaller magnitude, but dispersion tends to increase towards the thermal extremes, where the largest discrepancies are concentrated; within this range, MAE generally lies around 0.20 to 0.40 °C, and RMSE can approach 0.5 °C for the least favourable probes and at extreme temperatures.

Across all tests, the coefficient of determination (R^2) between the DS18B20 readings and T_{ref} was very high, reflecting a strong linear dependence on temperature over the range considered. However, because R^2 can be weakly discriminative when the thermal range is wide, performance evaluation relies primarily on MAE/RMSE and residual analysis, which quantifies bias, dispersion, and potential unmodelled patterns. Table 2 summarises these results and shows that, although variability exists between probes, no structural changes are observed between tests beyond the dispersion increase associated with the extended-range.

3.2.2. Residual Distribution, Data Cleaning, and Outliers

Figure 4 presents the distribution of the residuals, using boxplots for the probe and test. Overall, the residuals showed moderate probe-dependent biases (medians that may deviate from zero), interquartile ranges consistent with the standard deviations summarised in Table 2, moderate tails, and a limited number of outliers, predominantly concentrated at the coldest and hottest temperatures in Tests 3 and 4.

During the preliminary analysis, a very small set of anomalous readings was detected for probe 3 in Test 4, with residuals on the order of ±2 °C, which was clearly larger than the typical behaviour of the batch. A detailed inspection showed that, at the same instants, T_{ref} and the readings from the other probes remained clustered within a narrow interval (±0.2 to 0.3 °C), with no evidence of bath instability or common perturbations. Only probe 3 exhibited isolated jumps of approximately 2 °C, a pattern consistent with sporadic reading errors (e.g., faults in the 1-Wire line for that probe or transient conversion failures) rather than a genuine temperature change.

To avoid distorting the statistical characterisation and applying an explicit and reproducible criterion, the following data-cleaning procedure was adopted:

- The residuals were computed as $r_{s,i} = T_{ref,i} - T_{DS,s,i}$ for each probe (s) and observation (i).

For probe 3, candidate outliers were flagged as points simultaneously satisfying: $|r_{3,i}| > 1.0$ °C, and consistency of the remaining probes and the reference at the same observation (mutual differences < 0.5 °C).

- Nine observations in Test 4 met these criteria and were classified as spurious. These nine readings represented 1.8% of probe 3 readings in Test 4 (and approximately 0.6% of the total readings for that probe across the study).
- These observations were removed prior to recomputing descriptive statistics and generating the final versions of Figures 3 and 4.

After this filtering step, the residual distribution of probe 3 becomes consistent with that of the most stable probes in the batch, with no anomalous tails. This filtering does not materially affect the overall results (MAE, RMSE, R^2), but it prevents a very limited number of spurious readings from biasing the interpretation of the unit's behaviour.

For the remainder of the batch, probes 1 and 2 show more symmetric distributions and a lower incidence of outliers, whereas probes 5, 6, and 7 tend to exhibit greater dispersion in the extended range, particularly towards the thermal extremes. Probe 4 remains homogeneous within the interval in which it provides valid readings; however, it ceases to return reliable values above approximately 68 °C, and those setpoints were therefore systematically excluded from the analysis and from figure preparation.

3.2.3. Compliance with the Manufacturer's Specifications

The DS18B20 datasheet specifies a typical accuracy of ±0.5 °C over the −10 °C to +85 °C range under standard operating conditions. In light of the results obtained (after removing spurious readings and excluding setpoints with no valid readings):

- Mid-range (approximately 20 to 70 °C): the batch generally meets the specification, with absolute errors rarely exceeding ±0.4 °C and, over much of the range, remaining below ±0.3 °C.
- Extended range (near −10 °C and +85 °C): isolated cases are observed in which the error approaches the ±0.5 °C limit or slightly exceeds it, particularly for probes 5, 6, and 7, consistent with the expected degradation near operational boundaries.
- Probe 4: compliance is satisfactory within the interval in which reliable readings are produced; however, its inability to operate above 68 °C implies that, for applications requiring the full range, this unit should be discarded or assigned to restricted use.

The mean error and standard deviation were computed over all residuals ($T_{\text{DS18B20}} - T_{\text{Pt-100}}$) from the seven probes included in each test, applying the indicated data cleaning (exclusion of spurious readings from probe 3 in Test 4 and exclusion of setpoints above 68 °C with no valid readings from probe 4).

- MAE: mean of the absolute value of the residual.
- RMSE: square root of the mean of the squared residual.
- MRE: computed as the mean value of $\left(\frac{|r|}{|T_{\text{ref}}|}\right) \cdot 100$; in Tests 3 and 4 it increases because temperatures close to 0 °C and negative values are included, where small absolute errors translate into larger relative errors.

3.3. Cross-Test Analysis

The joint analysis of the four tests enables assessment of the robustness of the calibration procedure and the consistency of DS18B20 sensor behaviour across different thermal ranges and experimental runs. Whereas Tests 1 and 2 were conducted over a mid-range temperature interval (20 to 75 °C and 20 to 72 °C), Tests 3 and 4 extended the study to a wider range within the interval of interest (−10 to 85 °C), incorporating additional points (−12 to 86 °C).

Table 2 summarises the overall statistics by test (mean error, standard deviation, MAE, and, where reported, mean relative error), while Figures 3 and 4 provide visualisation of the overall linearity against the reference T_{ref} and the residual distributions by probe and

test, respectively. In general, the four tests show: (i) a strongly linear relationship between T_{DS} and T_{ref}, (ii) very similar behaviour between the mid-range tests, and (iii) a moderate increase in dispersion in the extended-range tests, concentrated at the thermal extremes.

3.3.1. Within-Test Consistency

Within-test consistency was assessed by analysing, for each test, the behaviour of the seven probes relative to the reference T_{ref} across the different temperature setpoints. Overall, the batch exhibits low-to-moderate within-test dispersion and probe-dependent biases, with a stable error structure within each test.

In Tests 1 and 2 (mid-range), several probes show particularly stable behaviour (low dispersion around their bias), whereas others exhibit more pronounced negative biases that are nonetheless consistent. This suggests that a substantial portion of the error is systematic in nature (offset and/or slope), which is favourable for subsequent correction using linear models.

In Tests 3 and 4 (extended range), within-test consistency is maintained, although dispersion increases towards the thermal extremes (near −10 °C and above 70 °C), where the sensors operate closer to the boundaries of the interval of interest. Probe 5 tends to show the most marked increase in dispersion under extreme conditions, while probe 4 remains consistent within the range in which it provides valid readings (with systematic exclusion of setpoints above 68 °C).

The values discussed in this subsection describe probe-level trends (within-probe variability within a given test) and are not directly equivalent to the "global" standard deviation in Table 2, which aggregates residuals across multiple probes and therefore also incorporates inter-probe variability.

3.3.2. Between-Test Reproducibility

Inter-test reproducibility was analysed by evaluating, for each probe, the error metrics obtained across the different tests. In the mid-range (Tests 1 and 2), differences in MAE/RMSE by probe are generally small, indicating high coherence between tests. Likewise, probes with pronounced negative biases (6 and 7) tend to preserve this pattern consistently, which is compatible with repeatable systematic errors and therefore highly reliable correction.

In the extended range (Tests 3 and 4), reproducibility remains notable: variations between tests for most probes remain bounded, and the overall pattern is also preserved for the least favourable units. In particular, the error increase is consistently concentrated under extreme conditions, without abrupt structural changes between tests.

3.3.3. Thermal Drift Between Tests

To assess the potential presence of drift over time, a specific analysis was performed contrasting Tests 3 and 4, which share the same extended range (−12 to 86 °C) and therefore represent the most demanding scenario. Differences in MAE and RMSE between these tests, by probe, generally remain moderate and, overall, provide no evidence of systematic drift. For the least favourable probes (probe 5), the increase in error in Test 4 is interpreted as greater sensitivity at the thermal extremes rather than as a monotonic shift consistent with drift.

At the global level, the differences observed between Test 3 and Test 4 are consistent in magnitude with the method uncertainty in the extended range (Section 3.5); therefore, they should be interpreted as expected variations within the measurement system rather than as significant drift.

3.4. Evaluation of Correction Models

After characterising the uncorrected behaviour of the DS18B20 sensors, several correction models were evaluated with the aim of reducing systematic error and dispersion while maintaining a simple and robust implementation suitable for IoT monitoring systems.

Models were fitted on a probe-by-probe basis using the cleaned dataset (exclusion of spurious readings identified for probe 3 in Test 4 and exclusion of setpoints above 68 °C with no valid readings from probe 4). The reference was T_{ref}, defined as the mean of two Pt-100 probes after applying the certificate corrections. In all cases, unweighted ordinary least squares (OLS) fitting was used. This choice was made for simplicity and because no dominant heteroscedasticity was observed within the range of interest that would justify a specific weighting scheme; accordingly, results should be interpreted as in-sample performance over the combined dataset, complemented by the between-test coherence assessment described in Section 3.3.

Four approaches were analysed:

- Constant offset
- Affine linear model
- Second-order polynomial model
- Segmented model (independent linear fits over thermal intervals)

Model quality was assessed using MAE and RMSE as the primary performance metrics, complemented by R^2 (for descriptive purposes) and inspection of residual structure. Figure 5 shows the evolution of MAE and RMSE before and after correction for each probe, and Table 3 summarises the final probe-level performance after applying the selected model.

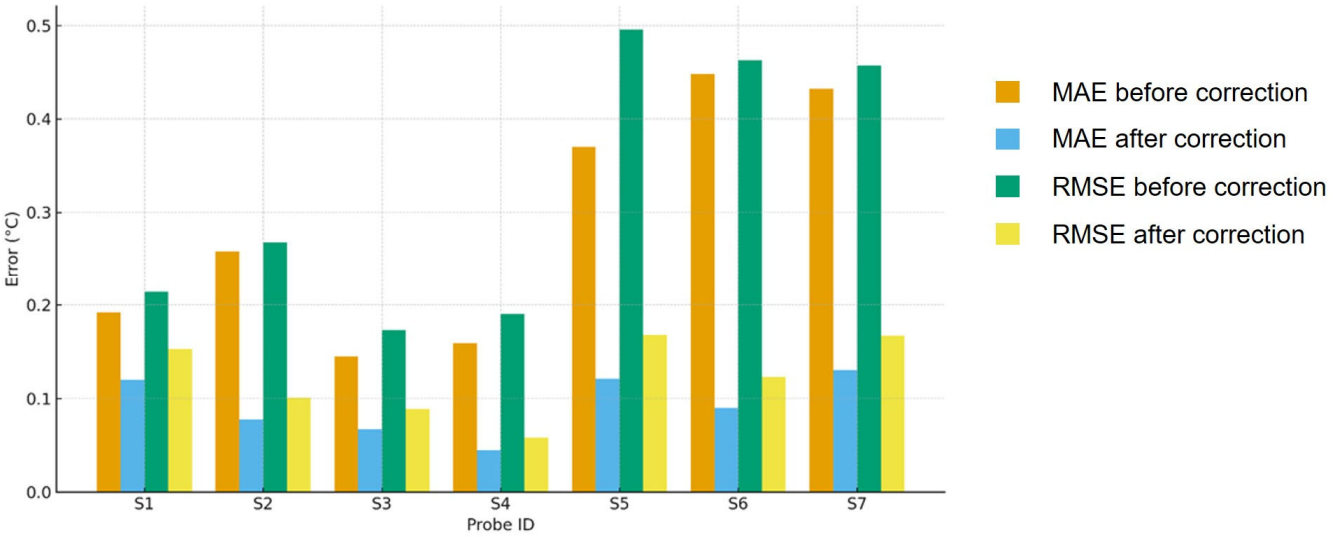


Figure 5. Evaluation of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) before and after probe-specific correction (combined data from Tests 1–4). For each probe (S1–S7), the four bars represent (from left to right): MAE before correction (orange), MAE after correction (blue), RMSE before correction (green), and RMSE after correction (yellow). The reduction in bar height across all probes demonstrates the effectiveness of the selected models in minimizing both systematic and random error components. While RMSE (yellow) is mathematically bound to be greater than or equal to MAE (blue), the close proximity between these two metrics after correction indicates a lack of significant outliers or uncontrolled random errors.

Table 3. Final probe-level performance after applying the affine linear model (combined data from Tests 1–4, cleaned dataset). Ranking established using RMSE as the primary criterion (and MAE as a secondary criterion in the case of very close values).

Rank	Probe	Corrected MAE (°C)	Corrected RMSE (°C)	R^2 (T_{corr} vs. Pt-100)
1	4	0.046	0.057	0.99999
2	3	0.063	0.089	0.99999
3	2	0.076	0.102	0.99998
4	6	0.090	0.123	0.99998
5	1	0.118	0.154	0.99997
6	7	0.130	0.168	0.99996
7	5	0.121	0.169	0.99996

3.4.1. Constant-Offset Model

The constant-offset model corrects only the sensor's mean bias (Equation (12)).

$$T_{corr} = T_{DS} + c_0 \tag{12}$$

where c_0 is estimated as the mean residual of the probe relative to T_{ref} over the calibration dataset.

This model yields a moderate reduction in MAE and a more limited improvement in RMSE, but it does not correct gain (slope) errors. Consequently, a residual temperature dependence (residual slope) tends to remain, particularly evident in the extended range, making the model insufficient when good performance is required across the full interval of interest.

3.4.2. Affine Linear Model

The affine linear model introduces two parameters (slope and intercept):

$$T_{\text{corr}} = a\,T_{\text{DS}} + b \tag{13}$$

Allowing simultaneous correction of both bias and gain (Equation (13)). This model was fitted on a probe-by-probe basis using the combined set of available setpoints (four tests, within each unit's effective operating range).

In contrast to the constant-offset model, the affine linear model provides a substantial improvement:

- It markedly reduces MAE and RMSE for all probes.
- It generally removes systematic structure in the residuals, which become approximately centred around zero with no clear trend with temperature.
- It retains a simple implementation (two coefficients per probe), compatible with microcontrollers.

Practically, this model reduces residual errors to the same order of magnitude as the calibration-method uncertainty, such that post-correction performance becomes increasingly limited by the reference system (standards and isothermal medium) rather than by the sensor itself.

Visual inspection of the residuals confirmed that the errors were randomly distributed around the zero-baseline. No discernible cyclic patterns or trends related to the bath's stabilisation phases were observed, indicating that the affine linear model sufficiently captured the sensors' systematic behaviour.

3.4.3. Higher-Complexity Models (Polynomial and Segmented)

To explore potential further improvements, two higher-complexity approaches were also tested:

- Second-order polynomial model.

$$T_{\text{corr}} = aT_{\text{DS}}^2 + bT_{\text{DS}} + c \tag{14}$$

  Within the studied range, this model provides only marginal improvements relative to the affine linear model (Equation (14)) (typically small changes in MAE/RMSE) and may introduce additional sensitivity at the range extremes, increasing the risk of overfitting without clear operational benefits.
- Segmented model by thermal intervals. Independent linear fits were evaluated over three intervals: −12 to 0 °C, 0 to 60 °C, and 60 to 86 °C. Although this approach can slightly reduce error at the extremes, it requires managing multiple coefficients per probe, complicates implementation in embedded systems, and may introduce discontinuities at interval boundaries unless additional constraints are imposed. Overall, the gains do not offset the increased complexity for routine use in IoT networks.

3.4.4. Synthesis and Selected Model

From the overall evaluation of models, based on MAE/RMSE (Figure 5) and final probe-level performance (Table 3), the following conclusions are drawn:

- The constant-offset model is simple but insufficient over wide ranges, as it does not correct gain and retains temperature-dependent residual structure.
- The affine linear model offers the best balance between accuracy, robustness, and simplicity: it corrects both bias and gain, significantly reduces MAE and RMSE, and yields residuals with no dominant systematic patterns.
- The polynomial and segmented models provide only marginal improvements over the linear model while increasing complexity and the risk of overfitting.

Accordingly, the probe-specific affine linear model was adopted as the definitive correction. The resulting parameters (a, b) were used in the uncertainty analysis and in the overall batch-performance assessment in the subsequent sections.

3.5. Analysis of Uncertainty Estimation

A rigorous measurement-uncertainty evaluation is essential for metrologically validating the comparison calibration process between DS18B20 sensors and the reference temperature T_{ref}, defined as the mean of two traceable Pt-100 probes. In this work, the analysis is conducted in accordance with the CEM TH-001 procedure and the GUM/VIM framework, using an uncertainty calculation that identifies the dominant contributions from the reference system and the sensor under calibration.

The aims of this subsection are to: (i) identify the main uncertainty components, (ii) estimate the combined and expanded uncertainty of the process, and (iii) relate this uncertainty to the residual errors obtained after applying the selected affine linear model (Section 3.4). In all cases, the cleaned dataset is used (exclusion of spurious readings from probe 3 in Test 4 and exclusion of setpoints above 68 °C with no valid readings from probe 4), such that the described calculation represents the normal behaviour of the experimental setup and the valid probes.

3.5.1. Uncertainty Components

Five main contributions were considered. To avoid double counting, DS18B20 repeatability was estimated from its readings within the steady-state temperature window (rather than from residuals), and reference-system contributions were evaluated separately.

- Pt-100 certificate uncertainty (Type B). The calibration certificate of each Pt-100 reports an expanded uncertainty U_{cert} = 0.14 °C with k = 2 (Equation (15)). Therefore, the standard uncertainty per Pt-100 is:

$$u_{\text{cert},i} = \frac{U_{\text{cert}}}{k} = \frac{0.14}{2} = 0.07\ ^{\circ}\text{C} \tag{15}$$

Since the reference is defined as $T_{\text{ref}} = (T_{\text{Pt-100,1}}) + T_{\text{Pt-100,2}})/2$, and assuming independence between certificates, the standard certificate contribution to T_{ref} is propagated as (Equation (16)):

$$u_{\text{cert, Tref}} = \frac{u_{\text{cert},i}}{\sqrt{2}} = \frac{0.07}{\sqrt{2}} \approx 0.050\ ^{\circ}\text{C} \tag{16}$$

- Effective non-uniformity of the measurement volume (Type B, conservative approach). Comparison calibration requires control of spatial variation within the fluid volume in which the Pt-100 and DS18B20 probes are simultaneously immersed. In this work, because an independent uniformity map was not available at each setpoint, a conservative approach was adopted based on the experimental indicators of the system for each test (Table 4), interpreted as an "effective non-uniformity" (bath + reproducible positioning) of the measurement volume.

Table 4. Isothermal medium stability parameters and inter-probe variability by test (°C). The "inter-probe SD" is reported as DS18B20 batch variability with respect to T_{ref}, rather than the spatial uniformity of the bath.

Test	Range	Mean SD T_{ref}	Max SD T_{ref}	Mean Inter-Probe SD (DS18B20 vs. T_{ref})	Max Inter-Probe SD
1	20 to 75	0.0034	0.0067	0.298	0.331
2	20 to 72	0.0044	0.0078	0.317	0.346
3	−12 to 86	0.0053	0.0148	0.307	0.496
4	−12 to 86	0.0115	0.0673	0.321	0.514

Under this approach, a representative half-width (a_{vol}) (Equation (17)) was selected and converted to a standard uncertainty assuming a rectangular distribution:

$$u_{\text{vol}} = \frac{a_{\text{vol}}}{\sqrt{3}} \tag{17}$$

- Temporal stability of T_{ref} over the steady-state temperature window (Type A). The residual temporal variation in the reference during steady-state acquisition was quantified as the standard deviation of T_{ref} at each setpoint, (Equation (18)) computed over the steady-state reading window:

$$u_{\text{stab}} = SD_{T_{\text{ref}}} \tag{18}$$

In practice, this contribution is secondary relative to u_{vol}, but it is retained explicitly to complete the calculation.

- DS18B20 repeatability within the steady-state temperature window (Type A). Probe repeatability was estimated from the standard deviation of DS18B20 readings over the steady-state window (Equation (19)) (prior to averaging per setpoint):

$$u_{\mathrm{rep},s} = SD_{T_{\mathrm{DS},s}} \tag{19}$$

- DS18B20 resolution (Type B). With 12-bit configuration, the quantisation step is $\Delta = 0.0625\ ^\circ\mathrm{C}$. Assuming a rectangular distribution for the quantisation error (Equation (20)), the associated standard uncertainty is:

$$u_{\mathrm{res}} = \frac{\Delta}{\sqrt{12}} = \frac{0.0625}{\sqrt{12}} \approx 0.018\ ^\circ\mathrm{C} \tag{20}$$

3.5.2. Combined and Expanded Uncertainty

For each probe s, the combined standard uncertainty associated with the comparison calibration (within that probe's effective operating range) (Equation (21)) was estimated as

$$u_{c,s} = \sqrt{u^2_{\mathrm{cert},T_{\mathrm{ref}}} + u^2_{\mathrm{vol}} + u^2_{\mathrm{stab}} + u^2_{\mathrm{rep},s} + u^2_{\mathrm{res}}} \tag{21}$$

Using representative values (Equations (22) and (26)):

$$u_{\mathrm{cert},T_{\mathrm{ref}}} = 0.050\ ^\circ\mathrm{C} \tag{22}$$

$$u_{\mathrm{vol}} = 0.17 \text{ to } 0.19\ ^\circ\mathrm{C} \tag{23}$$

u_{stab} on the order of 0.001 to 0.01 °C (steady-state window)

$$u_{\mathrm{rep},s} = 0.03 \text{ to } 0.11\ ^\circ\mathrm{C}\ (\text{probe-dependent}) \tag{24}$$

$$u_{\mathrm{res}} = 0.018\ ^\circ\mathrm{C} \tag{25}$$

yields

$$u_{c,s} = 0.18 \text{ to } 0.23\ ^\circ\mathrm{C} \tag{26}$$

and, assuming $k = 2$ (approximately 95% coverage), the expanded uncertainty:

$$U_s = k \cdot u_{c,s} = 0.36 \text{ to } 0.46\ ^\circ\mathrm{C} \tag{27}$$

For probe 4, uncertainty is reported only within its effective operating range (up to 68 °C); above this threshold, higher uncertainty is not assigned, but rather the corresponding setpoints are considered invalid due to the absence of reliable readings.

3.5.3. Relation to Residual Errors and Traceability

Assessing method uncertainty against post-correction residual errors allows the following to be established:

- Dominance of the reference system. The calculation is dominated by the contribution associated with the effective non-uniformity of the measurement volume (u_{vol}) and, secondarily, by the propagated certificate uncertainty in T_{ref}. DS18B20 resolution and the temporal stability of T_{ref} over the steady-state window are clearly secondary contributions.
- Consistency with corrected performance. After applying the affine linear model (Section 3.4), the probe-specific corrected RMSE values (Table 3) typically lie below the

estimated expanded uncertainty U_s of the process, which is consistent with a scenario in which final accuracy is mainly limited by the reference system and the isothermal stability of the calibration medium.

- Metrological traceability. Traceability is ensured by the use of Pt-100 probes calibrated with traceability to CEM and by application of a comparison calibration procedure in accordance with TH-001; consequently, corrected results inherit this traceability through T_{ref} and the uncertainty evaluation.
- Practical limit to model improvement. Because method uncertainty is constrained by the reference system, further marginal improvements through more complex models (polynomial or segmented) would tend to be masked by the process uncertainty itself; thus, the affine linear model represents an optimal trade-off between simplicity and performance within the experimental setup used.

3.6. Overall Performance Analysis

The integrated analysis of the four tests provides an overall assessment of the behaviour of the DS18B20 sensor batch, both in its uncorrected state and after applying the probe-specific affine linear correction model selected in Section 3.4. This combined view is relevant for evaluating the suitability of the sensors as calibrated thermometers for continuous monitoring applications and for assessing the robustness of the implemented calibration procedure.

Table 2 summarises the overall statistics by test prior to correction (mean error, standard deviation, MAE, and, where reported, mean relative error), whereas Table 3 reports the final probe-level performance after linear correction. Figures 3–5 complement this analysis by showing, respectively, the overall relationship between T_{ref} (mean of two Pt-100 probes) and DS18B20 readings, the characterisation of residual distributions by probe and test, and the performance evolution of MAE and RMSE before and after correction.

3.6.1. Overall Accuracy Before Correction

Taken together, the four tests show that DS18B20 sensors exhibit reasonable uncorrected performance consistent with the manufacturer's typical specification, yet with clear room for improvement and with appreciable inter-probe variability:

- The overall MAE per test typically lies around 0.28 to 0.31 °C (Table 2), with increased dispersion in the extended range (−12 to 86 °C) concentrated at the thermal extremes.
- At the probe level, uncorrected errors vary substantially between units (MAE and RMSE over broad intervals), indicating that the batch is not interchangeable without individual calibration: some probes exhibit moderate biases, whereas others show more pronounced and stable negative biases.
- The overall relationship between T_{ref} and T_{DS} is strongly linear (Figure 3), but the boxplots (Figure 4) reveal probe-dependent systematic biases, particularly for units whose residuals indicate consistent overestimation or underestimation relative to the reference.

Accordingly, although the uncorrected batch meets the typical ±0.5 °C tolerance over much of the interval of interest, inter-probe variability justifies individual calibration when interchangeability and metrological consistency across units are required.

3.6.2. Overall Performance After Affine Linear Correction

Applying the probe-specific affine linear model yields a marked improvement in performance:

- Corrected MAE is reduced to 0.046 to 0.130 °C (Table 3).

- Corrected RMSE falls within 0.057 to 0.169 °C, with substantial reductions relative to the raw condition, as shown in Figure 5.
- The coefficient of determination (R^2) between corrected temperature and T_{ref} is virtually unity for all probes (Table 3), consistent with the high linearity of the sensor-model combination over the studied range.

Probe 4 shows the best overall performance (MAE = 0.046 °C; RMSE = 0.057 °C), although its use is explicitly limited to its effective operating range (−12 to 68 °C) due to the loss of valid readings above 68 °C. At the other end, probes 5 and 7 exhibit the highest corrected RMSE values (0.17 °C), yet these remain low for post-correction environmental monitoring applications.

3.6.3. Batch Homogeneity and Probe Classification

Based on corrected performance (Table 3), the batch can be operationally classified as follows:

- Best performance: probes 4, 3, and 2 (corrected RMSE < 0.11 °C; MAE < 0.08 °C).
- Intermediate performance: probes 6 and 1 (RMSE 0.12 to 0.15 °C; MAE 0.09 to 0.12 °C).
- Less favourable performance: probes 5 and 7 (RMSE 0.17 °C). Nevertheless, their residual errors remain bounded and are compatible with post-correction use in most continuous monitoring scenarios.

This reduction in inter-probe dispersion after correction facilitates integration of the batch into distributed networks by decreasing systematic differences between units.

3.6.4. Relation to Method Uncertainty

Section 3.5 indicates that the expanded uncertainty associated with the calibration process typically lies within U_s = 0.36 to 0.46 °C (k = 2). Against this, the probe-specific corrected RMSE values (Table 3) remain within 0.057 to 0.169 °C, i.e., below the method's expanded uncertainty.

This result has two direct implications:

- Final performance is conditioned primarily by the reference system and the calibration environment (effective non-uniformity of the calibration volume and traceability of the standards), rather than by DS18B20 resolution.
- The affine linear model effectively exploits the available improvement margin compatible with the metrological infrastructure; further refinements would tend to be masked by process uncertainty.

3.7. Individual Sensor Performance and Operational Classification

Beyond the overall batch analysis (Section 3.6), describing probe-level performance is useful to support selection and allocation decisions in operational deployments. Table 3 summarises the quality indicators after affine linear correction (corrected MAE, corrected RMSE, and R^2), and Figure 5 demonstrates the reduction in MAE and RMSE relative to uncorrected data. Stability was qualitatively verified through the consistency of MAE/RMSE across tests (Section 3.3).

Based on corrected RMSE and the consistency observed between tests, three functional groups are proposed:

- High performance (internal control/operational reference): probes 4, 3, and 2.
- Standard-precision monitoring: probes 1 and 6.
- General monitoring: probes 5 and 7. In addition, probe 8 is considered excluded due to systematic failure (out of scope of the analysis)

This classification does not modify the metrological traceability of the process (determined by T_{ref}, derived from traceable Pt-100 standards, and by the isothermal medium), but it facilitates the design of networks with redundancy and a quality hierarchy.

3.7.1. High-Performance Group (Probes 4, 3, and 2)

Probes 4, 3, and 2 exhibit the lowest post-correction residual errors (Table 3). Probe 4 shows the best performance within its effective operating range, which is explicitly limited to −12 to 68 °C due to loss of valid readings above 68 °C. Probes 3 and 2 show centred residuals and consistent performance after removal of spurious readings identified for probe 3 in Test 4.

3.7.2. Standard-Precision Group (Probes 1 and 6)

Probes 1 and 6 provide robust and stable performance for environmental monitoring after linear correction (Table 3), with moderate residual errors and no evidence of dominant systematic structures in the residuals (Figure 4), consistent with the between-test coherence described in Section 3.3.

3.7.3. General Monitoring Group (Probes 5 and 7)

Probes 5 and 7 show the highest corrected RMSE values in the batch (Table 3), although these remain bounded and are suitable for applications where typical residual errors on the order of 0.17 to 0.20 °C are acceptable. In particular, probe 5 exhibits higher sensitivity under extreme conditions relative to the rest.

3.7.4. Practical Criteria for Selection and Deployment

Based on this classification, the following simple operational criteria are recommended:

- Include at least one probe from the high-performance group per subset of nodes for internal control and early detection of anomalous behaviour.
- Reserve the best-performing probes for critical locations or cross-validation points.
- Assign probes from the general monitoring group to areas where spatial measurement density is prioritised over minimisation of residual error.

In all cases, each probe's calibration coefficient pair (a, b) and functional group should be recorded alongside its physical identifier.

4. Discussion

The four tests show that encapsulated DS18B20 sensors, calibrated by comparison against a traceable reference T_{ref} (mean of two calibrated Pt-100 probes) in a controlled isothermal medium and corrected using a probe-specific affine linear model, can substantially improve their performance relative to the manufacturer's typical specification. Under uncorrected conditions, the overall per-test errors (MAE = 0.28 to 0.31 °C; RMSE < 0.5 °C) are consistent with ±0.5 °C over the range of interest (approximately −10 to 85 °C), but unit-specific biases and inter-probe variability limit interchangeability when metrological consistency is required. After individual correction, performance improves markedly (MAE = 0.046 to 0.130 °C; RMSE = 0.057 to 0.169 °C), with excellent linearity (R^2 close to 1) and residuals showing no discernible structure.

From a metrological perspective, once correction is applied, the residual error is constrained by the calibration infrastructure: $u_{c,s}$ = 0.18 to 0.23 °C and U_s = 0.36 to 0.46 °C (k = 2). The fact that corrected RMSE values remain below U is consistent with a scenario in which the practical limit is imposed by the reference system and the isothermal conditions of the calibration volume (traceability of the Pt-100 probes and effective non-uniformity), rather than by DS18B20 resolution or the complexity of the correction model.

Under controlled or indoor conditions, the literature on low-cost digital sensors reports substantial error reductions with sensor-specific calibration and simple (often linear) models, typically achieving tenths of a degree over comfort ranges [35–37]. Outdoors, particularly for air temperature, errors are often dominated by radiation and ventilation effects, and may remain on the order of 0.5–1 °C or higher even after correction [14]. In this context, the main contribution of this work is to extend the analysis to an expanded range (−12 to 86 °C), demonstrate inter-test reproducibility (Section 3.4), and place the procedure within a comprehensive metrological framework (TH-001, GUM/EA-4/02), while also providing an operational usage criterion based on corrected performance and each probe's effective range.

Regarding temporal stability, no evidence of significant drift was observed throughout the experimental campaign, which involved 72 h of intensive continuous measurements (Tests 3–4) including thermal cycling from −12 to 86 °C and changes in immersion media. This short-term stability remains well within the manufacturer's specified nominal drift of ±0.05 °C (typically rated after 1000 h of stress at 125 °C). Furthermore, although the experimental duration is shorter than the standard 1000 h rating period, the repeated exposure to thermal stress, cycling across the full operational range, serves as a robust indicator of the encapsulation's integrity and the sensors' metrological resilience. These findings confirm that the sensors maintained their characteristics throughout the validation process.

The study was conducted in a laboratory with continuously monitored ambient conditions, but without active climate control; no explicit environmental corrections were applied, nor were external effects independently characterised (e.g., air currents). While this study provides a robust thermal characterization, it is important to note that the experiments were conducted in a controlled laboratory setting. External factors such as electromagnetic interference (EMI), mechanical vibrations, or extreme humidity fluctuations were not independently characterized. In industrial deployments where these factors are prevalent, additional anti-interference testing would be required. However, for the target applications of this work, such as indoor climate control and heritage monitoring, thermal accuracy and reproducibility remain the primary metrological concerns.

Geometric/positional contributions were incorporated globally into the effective non-uniformity of the calibration volume. The representativity of the results is constrained by the sample size and provenance (a single batch and encapsulation); therefore, transfer to other batches, manufacturers, or encapsulations requires equivalent verification. Models were fitted using unweighted ordinary least squares without strict out-of-sample validation; consequently, performance should be interpreted within the tested domain. Finally, incidents were managed in a traceable manner (exclusion of spurious readings, delimitation of probe 4 above 68 °C, and exclusion of one unit due to systematic failure), reinforcing the need to document cleaning criteria and effective operating ranges when scaling the procedure or transferring it to field conditions.

5. Conclusions

The reference system provided adequate stability and uniformity for comparison calibration and for quantifying the uncertainty associated with the process. Without data pre-processing and outlier removal, the DS18B20 sensors exhibited errors consistent with the commercial specification (±0.5 °C over −10 to 85 °C). However, inter-probe variability and systematic biases argue against treating them as interchangeable sensors without individual adjustment. Applying a probe-specific affine linear model reduced the residual errors to MAE = 0.046 to 0.130 °C and RMSE = 0.057 to 0.169 °C, with R^2 close to unity and no appreciable structure in the residuals.

Within this framework, a decision rule with a target tolerance of ±0.5 °C was established. After individual correction, all probes comply with this criterion with a high level of confidence; the residual error (RMSE $\leq$ 0.17 °C) remains significantly below the tolerance limit, even when accounting for the expanded uncertainty of the system (U_s). This confirms that the proposed correction is sufficient to describe the sensor–reference relationship over the studied range.

The uncertainty analysis yields a typical combined standard uncertainty of $u_{c,s}$ = 0.18 to 0.23 °C and an expanded uncertainty of U_s = 0.36 to 0.46 °C (k = 2, level of confidence of approx. 95%). Consequently, after correction, final performance is conditioned primarily by the reference system and the thermal characteristics of the calibration volume (traceability of the Pt-100 standards and the effective non-uniformity of the measurement volume), rather than by the DS18B20 resolution or the complexity of the correction model.

Consistency across Tests 3–4, particularly in the extended range demonstrates good reproducibility of the calibration parameters, error metrics, and residual profiles. No evidence of significant temporal drift was observed throughout the experimental campaign, which involved two consecutive days of continuous measurements for Test 3 (covering 30 to 86 °C in distilled water, followed by −12 to 30 °C in isopropyl alcohol with 2 °C increments) and a subsequent 24 h period for the identical execution of Test 4. Nevertheless, repeating the evaluation after several months of operation under real conditions is planned in order to confirm long-term stability. Incident management was documented transparently: one unit was excluded due to a systematic failure (probe 8, omitted from the analysis), isolated faulty readings were removed (probe 3, Test 4) using an objective criterion, and the effective operating range of probe 4 was limited to 68 °C due to loss of valid readings at higher temperatures, an outcome that will inform improvements in future hardware and software versions of the Arduino UNO-DS18B20 system.

Specifically, future work will focus on verifying long-term stability through evaluations conducted after six months of real-world operation, following the same methodology as in Tests 3–4. This will allow for a direct comparison between the observed field-aging drift and the theoretical stability limits specified in the sensor's datasheet.

Overall, this study demonstrates and validates that low-cost encapsulated temperature probes can be individually calibrated with documented traceability, yielding instruments suitable for evidence-based decision-making in applications where reliability and consistency are required but extreme metrological refinement is not. In this respect, the proposed approach is particularly relevant for continuous IoT monitoring deployments in domains such as preventive heritage conservation, smart campus initiatives, and indoor environmental monitoring, provided that each probe's effective operating range is respected and that verification is replicated whenever the batch, encapsulation, or experimental conditions change, thereby ensuring reliable historical datasets. For outdoor deployments, further field studies and longer-term analyses are still required to assess real-world performance.

Author Contributions: Conceptualization, J.A.R.-R., D.A.M.-S. and L.P.M.; methodology, J.A.R.-R., A.M.L.; software, A.M.L., J.M.L. and J.A.R.-R.; validation, D.A.M.-S. and J.A.R.-R.; formal analysis, J.L.C.M. and J.A.R.-R.; investigation, J.A.R.-R., L.P.M. and A.G.L.; resources, A.M.L. and J.M.L.; data curation, J.A.R.-R. and A.M.L.; writing—original draft preparation, J.A.R.-R. and A.M.L.; writing—review and editing, A.G.L., L.P.M., A.M.L., J.M.L., D.A.M.-S. and J.L.C.M.; visualization, J.A.R.-R. and A.M.L.; supervision, D.A.M.-S. and J.L.C.M.; project administration, J.A.R.-R. and A.M.L.; funding acquisition, D.A.M.-S. and J.A.R.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available on request due to restrictions. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

Acknowledgments: We express our sincere gratitude to the Thermal Metrology Laboratory of the Department of Energy Engineering at the School of Industrial Engineering of the Technical University of Madrid (ETSII-UPM) for their support and collaboration in the experimental development of this work. We also extend our thanks to the ESCE-EELISA community for fostering an academic environment of cooperation that has significantly enriched this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Paramasivam, S.; Medda, R.; Losito, M.; Nataraj, D.; Subramanium, M.; Majumder, A.; Kumar, A.; Gatto, G. Mobility-Integrated Sensor Networks for Smart Environmental Monitoring in Urban Ecosystems. In Proceedings of the 2025 23rd Mediterranean Communication and Computer Networking Conference (MedComNet), Cagliari, Italy, 25–27 June 2025; pp. 1–6. [CrossRef]
2. Hamel, P.; Ding, N.; Cherqui, F.; Zhu, Q.; Walcker, N.; Bertrand-Krajewski, J.-L.; Champrasert, P.; Fletcher, T.D.; McCarthy, D.T.; Navratil, O.; et al. Low-cost monitoring systems for urban water management: Lessons from the field. *Water Res. X* **2024**, *22*, 100212. [CrossRef]
3. Presa Madrigal, L.; Rodríguez Rama, J.A.; Martín Sánchez, D.A.; Costafreda Mustelier, J.L.; Sanjuán, M.Á.; Parra y Alfaro, J.L. Cost-Effective Temperature Sensor for Monitoring the Setting Time of Concrete. *Appl. Sci.* **2024**, *14*, 4344. [CrossRef]
4. Palomeque-Gonzalez, J. A Modular, Low-Cost IoT System for Environmental and Behavioural Monitoring in Cultural Heritage Sites. *arXiv* **2025**, arXiv:2508.00849. [CrossRef]
5. Fisher, D.K.; Gould, P.J. Open-Source Hardware Is a Low-Cost Alternative for Scientific Instrumentation and Research. *Mod. Instrum.* **2012**, *1*, 8–20. [CrossRef]
6. Heimann, I.; Bright, V.; McLeod, M.W.; Mead, M.I.; Popoola, O.A.M.; Stewart, G.; Jones, R.L. Source attribution of air pollution by spatial scale separation using high spatial density networks of low-cost air quality sensors. *Atmos. Environ.* **2015**, *113*, 10–19. [CrossRef]
7. Marcelli, M.; Piermattei, V.; Gerin, R.; Brunetti, F.; Pietrosemoli, E.; Addo, S.; Boudaya, L.; Coleman, R.; Nubi, O.A.; Jojannes, R.; et al. Toward the widespread application of low-cost technologies in coastal ocean observing (Internet of Things for the Ocean). *Mediterr. Mar. Sci.* **2021**, *22*, 255–269. [CrossRef]
8. Giordano, M.R.; Malings, C.; Pandis, S.N.; Presto, A.A.; McNeill, V.; Westervelt, D.M.; Beekmann, M.; Subramanian, R. From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. *J. Aerosol Sci.* **2021**, *158*, 105833. [CrossRef]
9. Eichstädt, S.; Werhahn, O. Metrology for sensor networks: Metrological traceability and measurement uncertainties for air quality monitoring. *Tm-Tech. Mess.* **2024**, *91*, 419–429. [CrossRef]
10. Sekhar, P.K.; Billey, W.; Begay, M.; Thomas, B.; Woody, C.; Thiagarajan, S. Sensor Reproducibility Analysis: Challenges and Potential Solutions. *ECS Sens. Plus* **2024**, *3*, 046401. [CrossRef]
11. Ionascu, M.-E.; Gruicin, I.; Marcu, M. Towards Wearable Air Quality Monitoring Systems—Initial Assessments on Newly Developed Sensors. In Proceedings of the Telecommunications Forum (TELFOR), Belgrade, Serbia, 20–21 November 2018; pp. 1–4. [CrossRef]
12. Abdinoor, J.A.; Hashim, Z.K.; Horváth, B.; Zsebő, S.; Stencinger, D.; Hegedüs, G.; Bede, L.; Ijaz, A.; Kulmány, I.M. Performance of Low-Cost Air Temperature Sensors and Applied Calibration Techniques—A Systematic Review. *Atmosphere* **2025**, *16*, 842. [CrossRef]
13. Lin, J.J.; Buehler, C.; Datta, A.; Gentner, D.R.; Koehler, K.; Zamora, M.L. Laboratory and field evaluation of a low-cost methane sensor and key environmental factors for sensor calibration. *Environ. Sci.* **2023**, *3*, 683–694. [CrossRef] [PubMed]
14. Elyounsi, A.; Kalashnikov, A.N. Evaluating suitability of a DS18B20 temperature sensor for use in an accurate air temperature distribution measurement network. *Eng. Proc.* **2021**, *10*, 56. [CrossRef]
15. Vovna, O.; Laktionov, I.; Koyfman, O.; Stashkevych, I.; Lebediev, V. Study of metrological characteristics of low-cost digital temperature sensors for greenhouse conditions. *Serbian J. Electr. Eng.* **2020**, *17*, 1–20. [CrossRef]
16. Permana, A.N.; Wibawa, I.M.S.; Putra, I.K. DS18B20 sensor calibration compared with fluke hart scientific standard sensor. *Int. J. Phys.* **2021**, *4*, 1–7. [CrossRef]
17. Rama, J.A.R.; Lázaro, A.M.; Sánchez, D.A.M.; Lorenzo, J.M.; Barrio-Parra, F.; del Álamo, L.J.F.G. 3D printing as an enabler of innovation in universities: Tellus UPM ecosystem case. In *Smart Cities (ICSC-Cities 2023)*; Nesmachnow, S., Callejo, L.H., Eds.; Communications in Computer and Information Science; Springer: New York, NY, USA, 2024; Volume 1938. [CrossRef]

18. Rodríguez Rama, J.A.; Martín Sánchez, D.A. Ecosistema Tellus UPM Para el Desarrollo de un Smart Campus Sostenible (N.º 10). e-Politécnica Sostenible. 15 October 2024. Available online: https://sostenibilidad.upm.es/eps10-tellus (accessed on 2 December 2025).
19. Comunidad EELISA Ethics. Social Commitment & Entrepreneurship (ESCE). (s. f.). Available online: https://blogs.upm.es/eelisa-ethicssocialentrepreneurship (accessed on 2 December 2025).
20. Centro Español de Metrología. TH-001: Procedimiento Para la Calibración de Termómetros Digitales (Procedimientos de Medida). 3 February 2011. Available online: https://www.cem.es/es/divulgacion/documentos/th-001-procedimiento-calibracion-termometros-digitales (accessed on 11 December 2025).
21. Oficina Internacional de Pesas y Medidas Organización Intergubernamental de la Convención del Metro. Sistema Internacional de Unidades (SI). 2ª Edición en Español. 2008. Available online: https://www.cem.es/sites/default/files/siu8edes.pdf (accessed on 11 December 2025).
22. Centro Español de Metrología (CEM). Vocabulario Internacional de Metrología (VIM). Conceptos Fundamentales y Generales, y Términos Asociados. 3ª Edición en Español. 2012. Available online: https://www.cem.es/sites/default/files/vim-cem-2012web_0.pdf (accessed on 11 December 2025).
23. Cooperación Europea para la Acreditación. EA-4/02 M: 2013. Evaluación de la Incertidumbre de Medida en Las Calibraciones. 2013. Available online: https://fedaoc.online/privado/pluginfile.php/2473/mod_data/content/2024/EA_4_02_M_2013Rev1.pdf (accessed on 11 December 2025).
24. Dallas Semiconductor. DS18B20 Datasheet. (s. f.). Available online: https://www.alldatasheet.com/datasheet-pdf/pdf/58557/DALLAS/DS18B20.html (accessed on 3 December 2025).
25. Manzueta, R.; Martín-Gómez, C.; Gómez-Olagüe, L.; Zuazua-Ros, A.; Dorregaray-Oyaregui, S.; Ariño, A.H. A Living Lab for Indoor Air Quality Monitoring in an Architecture School: A Low-Cost, Student-Led Approach. *Buildings* **2025**, *15*, 2873. [CrossRef]
26. Schuurman, D.; De los Ríos White, M.I.; Desole, M.; Santonen, T.; Campodonico, G.; Vervoort, K. Living Lab Origins, Developments, and Future Perspectives. European Network of Living Labs. 2025. Available online: https://living-in.eu/sites/default/files/files/enoll-2025.-living-lab-origins-developments-and-future-perspectives.pdf (accessed on 11 December 2025).
27. Termostatos de Inmer Inmersión, Baños y Criotermostatos. Manual de Operación Julabo. 1.956.1300-V1. Available online: https://primametrology.com/wp-content/uploads/Manual-usuario-Dyneo-Julabo-ES.pdf (accessed on 11 December 2025).
28. Rodríguez-Rama, J.A.; Martín-Sánchez, D.A.; Maroto Lorenzo, J.; Presa Madrigal, L.; García Laso, A.; Costafreda Mustelier, J.L.; Marín Lázaro, A. Soluciones IoT de bajo coste para monitorización del efecto de isla de calor subterránea en entornos urbanos. In *Proceedings of the Annual International Congress of Doctoral Students 2025 (CAIED 2025)*; Poster PO-235; UMH Editorial Electrónica: Elche, Spain. Available online: https://editorial.umh.es/2026/03/04/annual-international-congress-of-doctoral-students-2025/ (accessed on 11 December 2025).
29. Rodríguez-Rama, J.A.; Presa Madrigal, L.; Marín Lázaro, A.; Maroto Lorenzo, J.; García Laso, A.; Costafreda Mustelier, J.L.; Martín-Sánchez, D.A. Innovación digital para la preservación patrimonial: El proyecto Smart Heritage ETSIME-UPM. In *Proceedings of the IV Annual International Congress of Doctoral Student (CAIED 2024)*; Abstract CO-103; UMH Editorial Electrónica: Elche, Spain, 2024. Available online: https://innovacionumh.es/editorial/Abstracts+CAIED+4.pdf (accessed on 11 December 2025).
30. LabVIEW (National Instruments). Available online: https://www.ni.com/es/shop/labview.html (accessed on 13 December 2025).
31. Nguyen, N.H.; Nguyen, H.; Le, T.T.B.; Vu, C.D. Evaluating Low-Cost Commercially Available Sensors for Air Quality Monitoring and Application of Sensor Calibration Methods for Improving Accuracy. *Open J. Air Pollut.* **2021**, *10*, 1–17. [CrossRef]
32. Si, M.; Si, M.; Xiong, Y.; Du, S.; Du, K. Evaluation and calibration of a low-cost particle sensor in ambient conditions using machine-learning methods. *Atmos. Meas. Tech.* **2020**, *13*, 1693–1707. [CrossRef]
33. Malings, C.; Tanzer, R.; Hauryliuk, A.; Kumar, S.P.N.; Zimmerman, N.; Kara, L.B.; Presto, A.A.; Subramanian, R. Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. *Atmos. Meas. Tech.* **2019**, *12*, 903–920. [CrossRef]
34. Ministerio de Industria, Turismo y Comercio, Centro Español de Metrología (CEM). Escala Internacional de Temperatura de 1990 (EIT-90). Available online: https://idinnova.wordpress.com/wp-content/uploads/2012/07/3-eit-90revisado.pdf (accessed on 11 December 2025).
35. Koestoer, R.A.; Saleh, Y.A.; Roihan, I.; Harinaldi. A simple method for calibration of temperature sensor DS18B20 waterproof in oil bath based on Arduino data acquisition system. *AIP Conf. Proc.* **2019**, *2062*, 020006. [CrossRef]

36. Young, D.T.; Chapman, L.; Muller, C.L.; Cai, X.; Grimmond, C.S.B. A Low-Cost Wireless Temperature Sensor: Evaluation for Use in Environmental Monitoring Applications. *J. Atmos. Ocean. Technol.* **2014**, *31*, 938–944. [CrossRef]
37. Zaszewski, D.; Gruszczyński, T. A Low-cost Automatic System for Long-term Observations of Soil Temperature. *Geomat. Environ. Eng.* **2022**, *17*, 75–101. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.