scientific reports

OPEN

Ensuring reliable cave temperature data for climate change research

Alice Cimenti[1,2,9], Lorenzo Cresi[3,9]✉, Marco Isaia[2,4], Elena Piano[2], Anna Piquet[1], Giuseppe Nicolosi[5], Stefano Mammola[4,6,7], Antonella Senese[3,10] & Fiorella Acquaotta[1,8,10]

Caves are unique natural laboratories for studying climate change and its ecological impacts. However, analyzing air temperature in these environments is challenging due to stable microclimatic conditions and high humidity. Collecting reliable data requires dedicated devices and protocols. We developed a standardized quality control procedure—Cave Air Temperature Quality Control (CAT-QC)—to assess the reliability of temperature data collected inside caves. The protocol consists of four main steps: (i) assessing data completeness; (ii) identifying physically implausible values; (iii) detecting statistical outliers using three progressively sensitive methods; and (iv) conducting a final manual check. We tested CAT-QC on a dataset from 19 caves in the Piedmont region (Northwest Italy), recorded with iButton devices. The protocol effectively identified gaps, absurd values, and abrupt temperature changes, many of which were due to human interference or sensor issues. Data flagged through CAT-QC can be further reviewed to address biases and rerun through the process if needed. Designed for broad applicability, CAT-QC is dynamic and can be tailored to local series characteristics, making it suitable for diverse subterranean environments. This tool provides a robust framework for ensuring data quality and comparability in cave climate studies, supporting research and conservation efforts in the context of climate change.

Keywords Air temperature, Long-term monitoring, Quality control, Subterranean ecosystem, Climate change

Caves are unique natural laboratories to study climate change processes and their broader ecological impacts[1,2]. Their semi-closed nature makes them relatively insulated from rapid external atmospheric fluctuations[3,4]. The temperature of the air inside the cave becomes increasingly stable the more one moves away from the entrance[5] attaining values typically hovering around the external mean annual temperature of the surface area surrounding the cave[3,6,7]. As the insulation is recognized as affecting the thermal regime along the caves, in this study, we refer to cave insulation as the thermal buffer of the air inside the cave from the external atmosphere.

Notwithstanding these insulation mechanisms, theoretical and empirical studies indicate that caves are sensitive to external climatic variations over longer time frames[8–10]. In fact, due to their ability to buffer surface temperature changes, caves act as important proxies for understanding past and present climate dynamics, offering a critical perspective that complements traditional atmospheric climate studies[11]. In the same vein, these systems could be used as models to anticipate the effect of current anthropogenic climate change[2,12].

Moreover, caves serve as essential research sites for a range of disciplines, including ecology, biology, hydrogeology, and conservation science. One key area of study is the unique adaptations of cave-dwelling organisms, which thrive in stable conditions characterized by darkness, high humidity, and constant temperatures. Even minor alterations in these conditions could have profound effects on subterranean biodiversity[13]. Consequently, understanding climate change impacts on cave ecosystems is essential for preserving these often-endemic species and maintaining ecological balance[14–16].

[1]Dipartimento di Scienze della Terra, Università degli Studi di Torino, Torino, Italy. [2]Dipartimento di Scienze della Vita e Biologia dei Sistemi, University of Torino, Torino, Italia. [3]Department of Environmental Science and Policy (ESP), Università degli Studi di Milano, Via Celoria 10, Milano 20133, Italy. [4]National Biodiversity Future Center, Palermo, Italy. [5]Department of Biological, Geological and Environmental Sciences, University of Catania, Catania, Italy. [6]Molecular Ecology Group (MEG), Water Research Institute (IRSA), National Research Council (CNR), Verbania, Italy. [7]Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland. [8]Natrisk, Centro Interdipartimentale sui Rischi Naturali in Ambiente Montano e Collinare, University of Turin, Turin, Italy. [9]These authors contributed equally to this work: Alice Cimenti and Lorenzo Cresi. [10]These authors jointly supervised this work: Antonella Senese and Fiorella Acquaotta. ✉email: lorenzo.cresi@unimi.it

Despite the ecological and climatological importance, existing compilations of cave temperature data from around the world remain highly uneven in both geographic and temporal coverage (e.g[3].,. For example, data is particularly scarce in tropical regions[17]. The lack of long-term monitoring programs and uniform methodologies hinders our ability to evaluate the impact of climate change on subterranean environments[2]. Without a consistent data collection framework, it is difficult to develop accurate models to predict future climatic and biological shifts, particularly at regional and local scales, where methodologies are highly sensitive to definition, scale, and estimation[14]. Indeed, most climate change modelling exercises for subterranean species are based on far-from-ideal proxy variables of interpolated external temperature conditions, leading to high uncertainties (e.g[18,19]. Furthermore, caves exhibit delayed equilibrium between internal and external temperatures[12], reinforcing their role as valuable climate proxies while highlighting the necessity for precise and continuous monitoring.

Several challenges contribute to the lack of standardized cave climate data. Firstly, caves are often remote and difficult to access, requiring substantial resources for regular monitoring. Secondly, the challenging conditions within caves, particularly high relative humidity levels with values often close to saturation[5], increase the risk of sensor failure and data loss for temperature sensors[2]. Thirdly, variations in lithology, altitude, and outside climate can significantly influence cave microclimates. For example, caves formed in high-thermal-conductivity rocks experience greater fluctuations in response to external air temperature changes, while altitude influences external climatic conditions, subsequently affecting subterranean temperatures[5].

Given these challenges, implementing a rigorous quality control protocol is essential to ensure the reliability and comparability of cave temperature datasets. High-quality, standardized climate records are fundamental for accurately assessing climatic trends and ecological impacts in subterranean environments. A well-designed quality control framework can help eliminate inconsistencies caused by sensor errors, site-specific anomalies, or data gaps, thereby strengthening the validity of long-term cave climate research. This study presents a novel Cave Air Temperature Quality Control (CAT-QC) protocol (Fig. 1), developed through the monitoring (up to 4 years) of 19 small to medium caves (from 10 to 3500 m of development) in northwestern Piedmont (Italy) using iButton devices. Our approach establishes a standardized methodology to obtain reliable and comparable temperature series in subterranean environments, addressing key data gaps, physically implausible values, statistical outliers, and anomalies caused by human interference or sensor malfunctions.

Methods

Cave air temperature quality control (CAT-QC)

In this study, we developed a structured and automated quality control workflow in R (version 4.3.1[20];, leveraging the packages lubridate (v1.9.3[21];, ggplot2 (v3.5.1[22];, and zoo (v1.8.12[23]; (Fig. 1). The procedure involves complementary checks to identify anomalies and data inconsistencies through different steps: (i) assessment of data completeness and the presence of missing values; (ii) detection of physically non-sense temperature values; (iii) identification of different types of statistical outliers based on three criteria of increasing sensitivity; and

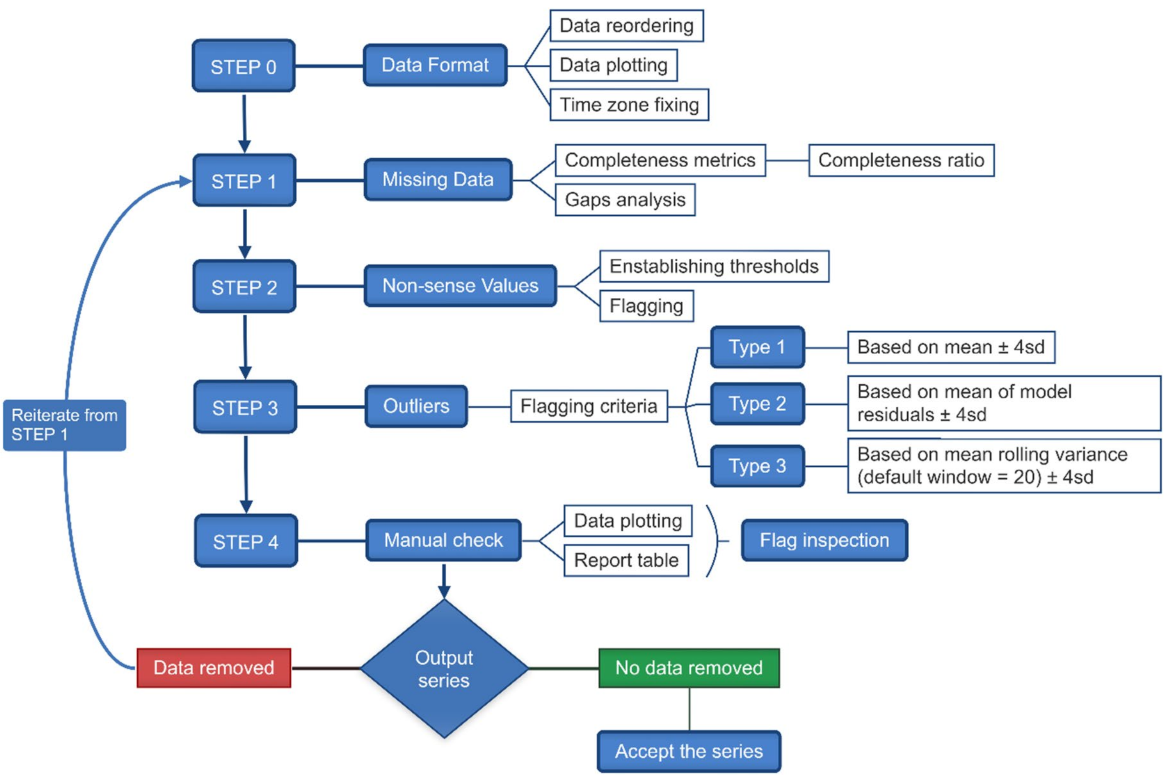

Fig. 1. Schematic summary of the quality control for cave air temperature (CAT-QC) workflow.

lastly (iv) manual check aimed to examine outliers and decide whether they have to be removed from the series. In case of data removal, an additional run of the entire quality control procedure is necessary. Additionally, we implemented visualization functions (plot_data) to facilitate visual inspection and interpretation of outliers. The custom R scripts, functions, plots and implementations described herein are publicly available alongside this paper (Online Resource 1; Online Resource 2), ensuring transparency and reproducibility of the quality control procedure. These components provide a solid framework for flagging, documenting, and addressing anomalies, enhancing data quality for downstream ecological and climatological analyses.

Step 0: data format

Before any advanced quality control step, an appropriate data formatting process is needed (Fig. 1). The CAT-QC achieves this by chronologically ordering the yearly series of data and plotting them to obtain the complete temperature series available for each cave, allowing the user to verify if the temperatures are continuous, and to visually see where non-sense values are located. This shows the seasonality pattern as well, something to be inspected to spotlight any suspicious seasonal trend given by a device's error. Moreover, to provide temporal consistency, the UTC standard time zone is set in every series. This approach makes sure to avoid issues related to local time shifts that could happen during the data recording phase.

Step 1: missing data

The first step of the CAT-QC is the detection of missing data, if any. The CAT-QC assesses data completeness based on the expected number of records and the length and distribution of the possible gaps in the series. In most climatic analyses, monthly mean values can be computed when at least 80% of daily data is available, with gaps not exceeding six non-consecutive days per month[24]. For annual averages, a series is considered complete when at least 94% of daily data is available, with gaps limited to 20 non-consecutive days[25–27]. Based on the acquisition rate of the time series being analysed, days with missing readings exceeding the 80% threshold are excluded from the calculations.

Step 2: non-sense values

Secondly, the CAT-QC evaluates the presence of non-sense values[28] (Fig. 1; Table 1). Since the interest is in the signal of external air temperature inside caves, to exclude any other source of subterranean heat, we used the thresholds based on the device temperature range and on European temperature records which led to thresholds of ± 40 °C[29–31]. When non-sense values are identified, they are flagged as "non-sense value" and shown in a report table. The identified non-sense values will not be cut nor converted in NA (i.e. not available data) until the CAT-QC is completed.

Step 3: outliers

We refer to outliers as any record that significantly deviates from the average range of the dataset, based on three different criteria (see below). A typical series of temperatures measured in a cave lies within a narrow temperature range[6], which is likely to become narrower as the cave deepens. On the contrary, the temperature signals from a shallow cave or other superficial subterranean habitats are much more similar to the external air temperature, with seasonal and daily variations[9,32]. These characteristics lead the distribution of temperature series data to have both lower (deep caves) and higher (shallow caves) variance. Therefore, depending on the location of the recording sensor within a cave, outliers could be erroneous data caused by a device malfunction or natural values. Furthermore, an outlier in a cave temperature series can also have other natural explanations, such as an abrupt decrease in temperature due to water infiltration, such as melting snow[5], or an abrupt increase due to visitors entering a tourist cave[33]. For this reason, the highlighted "outlier" data is not blindly removed; instead, the decision is left to the user via an appropriate subjective control (i.e. step 4). Because of this, the CAT-QC uses three methods with increasing sensitivity to flag outliers:

A. The first outlier detection method is based on the overall mean ± 4 standard deviations of the series, defined as the generally used threshold for outliers[34]. When this kind of outlier is identified, it is flagged as "Type 1 outlier".
B. For a more accurate procedure, the CAT-QC fits the temperature series with a model, extracting residuals, and then targeting outliers among those. We use a harmonic model to calculate the residuals, effectively de-seasoning the time series. Specifically, we fit a single-frequency sinusoidal function to capture the dominant annual cycle, which is the primary source of variance in cave air temperature data. The model is defined as: $T(t) = A \cdot \sin(2\pi/365.25 \cdot t + \phi) + C$, where $T(t)$ is the temperature at time t, A is the

Flag	Description
Non-sense values	Values above + 40 °C or under − 40 °C.
Type 1 outliers	Outliers identified based on the overall mean ± 4 times the standard deviation.
Type 2 outliers	Outliers identified based on ± 4 times the standard deviation from the mean of the residuals calculated with a harmonic model.
Type 3 outliers	Outliers identified based on ± 4 times the standard deviation from a rolling variance calculated on a 20 records window.

Table 1. List of flags presented in the quality control for cave air temperature (CAT-QC) with their descriptions.

amplitude, ϕ is the phase shift, and C is the mean temperature. We acknowledge that this single-frequency approach may not perfectly capture asymmetric seasonal peaks/troughs or higher-frequency variations (e.g., diurnal or weather-front related). However, the primary purpose of this step is to remove the dominant seasonal signal to isolate anomalies. The specific implementation of this harmonic model, including the use of the sine and cosine terms to capture the annual cycle, as well as the performance metrics for the harmonic models, including the R-squared and coefficients for each time series, are calculated in the R script (Online Resource 1) and shown in the models' summary table (Online Resource 3). The residuals, R(t) = T_measured (t)-T(t), represent the deviation from this expected seasonal pattern. Outliers are then flagged ("Type 2 outlier") when a given value varies ± 4 standard deviations from the mean of the residuals.

C. Lastly, the CAT-QC evaluates the amount of variance between records to flag "Type 3 outlier". To do this, it calculates the rolling variance choosing a window of 20 records and then identifying outliers based on the mean variance along the window, using a threshold of ± 4 standard deviation. Notice that this window is an adjustable parameter that the user can set according to the sensor's acquisition rate and to the expected variance of the series to increase or decrease the sensitivity of this step.

The output of this step of the CAT-QC procedure is a new extended dataset reporting all preexisting information from the original series (no data is cut or converted into NA) plus additional logical columns dedicated to the flags based on outlier types for each record. Finally, the total number of each outlier type is automatically computed and shown in a new object in the R environment. This allows for precise identification of each flagged data across the dataset, so they can be properly analysed, counted, or removed based on the type of flag in Step 4.

Table 1 shows all the outlier types with their description. The great added value of this step is to provide a dynamic outliers' detection, which is strongly linked to local series features since Step 3 C is based on a rolling statistic which can be adjusted depending on the considered data series.

Step 4: manual check

To identify possible erroneous values and non-climatic biases in series, each step of the quality control needs to be manually reviewed to make sure that the revealed record is truly a measurement error. For the last manual control of the CAT-QC (Fig. 1), graphs are plotted to allow the user to visualize the flagged data and to decide whether to remove them. This step is crucial for differentiating between true sensor errors and real, albeit extreme, natural events (e.g., sudden water infiltration or human presence). The manual check also allows for the removal of large sections of data that, while not every single point is flagged by the automated steps, are clearly erroneous based on expert knowledge and field notes (e.g., periods where the sensor was known to be temporarily moved). In such cases, the automated QC often flags a high density of outliers within the affected section, indicating a systemic issue. This is a deliberate feature of the protocol, as the automated steps are designed to be sensitive but not overly aggressive, and large-scale systematic errors may not register as typical statistical outliers. The decision to remove unflagged data is always documented and based on external evidence. Since outliers are coloured differently to point out the different outlier types, one record can be flagged as a non-sense value and as an outlier based on more than one criterion; in this case, it would appear in the plot with a different colour indicating multiple flags. The graphs enable to identify seasonal patterns of the series and abrupt temperature changes and to link the events with the date and time.

Each flag designated as NA during the final manual check is treated as missing data. Consequently, once the manual check is complete, the CAT-QC process should be repeated to identify any outliers that were previously deemed acceptable. When no erroneous data are detected and the number of missing values does not compromise the integrity of the dataset, the series can be considered reliable.

Application of CAT-QC on a case study

To test our quality control protocol, we analysed 38 air temperature series recorded in caves located across the Western Alps between autumn 2020 and winter 2024–2025. To establish a reliable dataset and remove erroneous records, we applied the quality control process two times. The first run was performed on the raw data. After the first quality control and erroneous data were removed, the QC-CAT was run again to further refine the data.

Data collection and dataset

We analysed 19 different caves distributed across the Western Alps, Piedmont, Italy (Table 2). We measured air temperatures at two locations inside the cave: one at the cave opening and the other as deep inside as possible (called entrance and internal, respectively). The measurements were carried out continuously from autumn 2020 until winter 2024–2025 (Table 2), using I-Button devices (model DS1922L and DS1923) with an accuracy of ± 0.5 °C and a resolution of 0.0625 °C[31]. They combine a thermometer and a datalogger in a single compact unit. The temperatures were recorded every 6 h each day (i.e. at 0 am, 6 am, 12 pm, and 6 pm).

They rely on a battery for self-sustaining and energy supply, that should allow for a time of use of multiple years of records. Storage-wise, the devices are capable of recording 8 kb in 8 or 16 bit[35]. To ensure continuous data collection within these limits, the instruments were replaced annually, resulting in a total of 133 individual annual series. We decided to use the I-Button because of their data recording capacity, their compact size to be hidden in caves, and their cost-efficiency, as they are considered expendable materials. To create a single, continuous record for analysis, the annual series were merged for each of the two monitoring points (entrance and internal) in every cave prior to applying the CAT-QC protocol.

Cave name	Inventory number	WGS84 coordinates (cave opening)	Elevation of the cave opening (m a.s.l.)	Development (m)	Distance of the entrance sensor from the cave opening (m)	Distance of the internal sensor from the cave opening (m)	Monitored period
Grotta delle Arenarie	PI2509	45.711919 ° N 8.3145207° E	770	3000	1	70	12 Nov 2020 28 Oct 2024
Buco dell'Aria Calda	PI1102	44.3484026° N 7.4623675° E	852	115	1	20	21 Oct 2020 14 Nov 2024
Grotta Occidentale del Bandito	PI1003	44.290002° N 7.427431° E	714	690	1	50	21 Oct 2020 14 Nov 2024
Grotta di Bercovei	PI2503	45.6606509° N 8.2654061° E	415	170	1	150	12 Nov 2020 04 Nov 2024
Grotta di Bossea	PI0108	44.241548° N 7.8398498° E	836	2800	1	1700	16 Oct 2020 17 Jan 2025
Grotta inferiore del Caudano	PI121	44.2930025° N 7.7905788° E	778	3440	2	800	16 Oct 2020 17 Jan 2025
Grotta di Chiabrano	PI1621	44.9470661° N 7.1060462° E	1080	27	1	6	20 Oct 2020 14 Oct 2024
Grotta La Custreta	PI1593	45.4463179° N 7.5458605° E	1350	180	1	26	17 Nov 2020 21 Oct 2024
Tuna dal Diaou	PI1591	45.0262921° N 7.1218405° E	1392	140	1	50	20 Oct 2020 14 Oct 2024
Tana della Dronera	PI151	44.3440187° N 7.8459223° E	525	134	2	80	27 Oct 2020 17 Oct 2025
Balma Fumarella	PI1597	45.1270119° N 7.0319789° E	864	47	1	15	19 Oct 2022 09 Oct 2024
Buca del Ghiaccio della Cavallaria	PI1609	45.5182779° N 7.7952332° E	1550	24	2	16	17 Nov 2020 21 Oct 2024
Ghieisa d'la Tana	PI1538	44.8492344° N 7.2226677° E	841	50	1	12	27 Oct 2020 21 Oct 2024
Buco del Maestro	PI1148	44.6845875° N 7.2363044° E	750	17	1	8	20 Oct 2020 11 Dec 2024
Grotta Testa di Napoleone-1	PI1569	45.1144036° N 7.2605897° E	450	50	1	25	9 Oct 2020 25 Oct 2022
Buco del partigiano	PI1315	44.506897° N 7.2932269° E	1170	13	2	11	17 Oct 2020 28 Sep 2024
Grotta del Pugnetto	PI1501	45.2722721° N 7.4124256° E	820	765	1	500	29 Oct 2020 29 Nov 2024
Borna del Servais B	CAPI1756	45.322528° N 7.327619° E	1387	17	2	12	29 Oct 2020 29 Nov 2024
Sotterranei del forte (A) di Vernante	artificial (not included in the speleological inventory)	44.2522208° N 7.527921° E	784	91	3	50	21 Oct 2020 14 Nov 2023

Table 2. List of monitored caves (in Italian "grotte") and their characteristics. All the reported data are granted by the speleological inventory of the region[37].

A detailed description of the dataset is available in[36] (under preparation), and the database itself is going to be accessible as well in Figshare.

Results

Missing data

In our case study, all devices initially recorded without interruption, yielding a total of 133 annual time series, including data from device replacements. These yearly series were then merged into longer multi-year time series, creating one pair per cave (entrance and internal). However, due to occasional device failures, some series were incomplete, typically missing about one year of data. In some instances, short overlaps between consecutive device records occurred, providing an opportunity to test the CAT-QC protocol on particularly raw and fragmented data. In summary, 5 internal series and 6 entrance series—29% of the 38 multi-year series—had complete 4-year records. Another 6 internal and 8 entrance series (37%) spanned 3 to 4 years, typically missing about one year of data. The remaining 34% of the series covered 2 years or less.

Non-sense values and outliers

Initially, the CAT-QC was applied to a total of 185,488 raw records. This allowed us to recognize the erroneous data, which were then automatically removed. Thus, the records were reduced to 160,460. With the second run of CAT-QC, no non-sense data was identified, and the number of outliers greatly decreased (see Tables 3 and 4).

In the first run of CAT-QC, punctual non-sense values emerged in five internal series (i.e. PI2503, PI1621, PI1593, PI1569, Sotterranei del forte di Vernante) and in one entrance series (i.e. PI2509). The occurrence of anomalous temperature readings, such as – 40 °C, may be attributed to battery depletion, leading to insufficient power for accurate sensor operation, or a system reset triggered by power interruptions, as indicated by the

Caveme	Internal series first run						Internal series second run					
	Total records	Non-sense values	Type 1 outliers	Type 2 outliers	Type 3 outliers	Total outliers	Total records	Non-sense values	Type 1 outliers	Type 2 outliers	Type 3 outliers	Total outliers
Grotta delle Arenarie	5144	0	4	72	139	215	4368	0	3	81	43	127
Buco dell'Aria Calda	3481	0	31	59	21	111	3176	0	0	0	17	17
Grotta Occidentale del Bandito	1647	0	25	27	17	69	1616	0	0	38	16	54
Grotta di Bercovei	6932	265	265	265	22	552	5796	0	0	0	32	32
Grotta di Bossea	4999	0	134	247	24	405	4632	0	0	0	34	34
Grotta inferiore del Caudano	6797	0	77	292	76	445	6153	0	0	0	34	34
Grotta di Chiabrano	7212	1	2	27	20	49	5800	0	0	47	69	116
Grotta La Custreta	3323	1	1	17	49	67	2792	0	0	3	37	40
Tuna dal Diaou	5251	0	0	27	96	123	4280	0	0	1	15	16
Tana della Dronera	5138	0	113	229	85	427	4804	0	0	0	70	70
Balma Fumarella	1863	0	5	33	19	57	1448	0	0	6	18	24
Buca del Ghiaccio della Cavallaria	3316	0	0	18	51	69	2792	0	0	18	51	69
Ghieisa d'la Tana	3489	0	0	82	91	173	3076	0	0	3	24	27
Buco del Maestro	6876	0	0	181	44	225	6032	0	0	12	27	39
Grotta Testa di Napoleone-1	3657	40	40	41	61	142	2976	0	0	22	11	33
Buco del partigiano	5517	0	1	28	67	96	4296	0	0	22	62	84
Grotta del Pugnetto	3128	0	58	84	46	188	2851	0	0	3	0	3
Borna del Servais B	5111	0	0	92	0	92	4484	0	0	28	37	65
Sotterranei del forte (A) di Vernante	7187	2	8	141	22	171	5924	0	0	33	189	222

Table 3. Comparison between the results emerged from the CAT-QC first and second runs with the internal series. For each cave is shown from the left to the right the total records, the number of records flagged as non-sense values, as type 1 outliers, Tas type 2 outliers, as type 3 outliers, and the total number of outliers. The total records shown in the second run is the number of records that underwent manual check at the end of the first run and therefore cleaned from non-sense values and records found to be erroneous among the outliers.

Battery-On Reset Alarm. Additionally, these extreme values could result from sensor malfunctions, calibration errors, or the device defaulting to predefined failsafe values when proper measurements cannot be obtained.

Specifically, the second iteration of the CAT-QC identified totally three type 1 outliers in the internal series and five type 1 outliers in the entrance series (Tables 3 and 4); type 2 and type 3 outliers greatly decreased (Fig. 2).

Various outliers were found by the CAT-QC. Some outliers highlighted as multiple flags could be grouped because they all show significant temperature changes at the beginning or end of the series. All series showed this behaviour because the devices started recording temperatures as soon as they were set up a day before fieldwork, or they needed some time to adjust to the cave condition after installation. Those records are likely to be flagged as Type 1 outliers, or by multiple flags since major anomalies indicate the exact times when the instruments were set up before being placed in the cave or when they were removed from the cave and brought back to the laboratory.

Although we suspected that the CAT-QC could also mistakenly indicate as type 1 outliers the temperature peaks that can normally be recorded outside the caves due to the meteorological variability, this case did not occur, thanks to the natural insulation of caves. On the other hand, only truly anomalous values caused by indoor recordings were found to be type 1 outliers.

The type 2 outliers allowed us to check the seasonal pattern of the series and highlight those data points that deviate from the normal residual distribution by identifying smaller breaks than the type 1 outliers. Instead, type 3 outliers highlight rapid decreases or increases in temperature. The comparison of each series plots (internal series first run and second run, entrance series first run and second run) is available in the supplementary materials (Online Resource 2).

Finally, the temperature series underwent two iterations of the CAT-QC, in which we included a manual check of the flagged outliers, which were sufficient for our unique database, spanning four consecutive years across 19 caves in the Western Alps.

Discussion

The CAT-QC protocol represents a simple-to-use, time-efficient tool for semi-automatically assessing the reliability and consistency of air temperature records collected in subterranean environments. The protocol provides a standardized framework for data validation, which is crucial for enhancing the comparability of datasets across different studies. By implementing a multi-step approach, the CAT-QC process aims to detect erroneous data that could arise due to sensor malfunctions, external or human disturbances. However, the protocol is not without limitations and could be fine-tuned for broader applications to different geological, climatic and ecological contexts.

One of the primary challenges associated with the CAT-QC protocol is the subjectivity involved in the final manual check of flagged data. While automated processes help identify anomalies, researchers must make individual judgments on whether specific values should be retained or discarded, introducing potential inconsistencies. The process is also time-consuming, as multiple steps, including visual inspections, must be performed to ensure the accuracy of the dataset. Additionally, the protocol's sensitivity to environmental variability means that natural variations in cave temperatures may sometimes be misclassified as outliers, particularly in caves with highly dynamic temperature regimes. Another factor affecting its effectiveness is the quality and placement of sensors within caves. If devices suffer technical failures, biases may be introduced that the quality control process may fail to fully mitigate. Finally, as the protocol has only been tested in the Piedmont region, its applicability to caves with different morphological and climatic conditions remains to be fully validated. Moreover, as each cave exhibits unique morphology, development, and environmental conditions, these factors must also be carefully considered in the decision-making process.

For example, to find an explanation for type 2 and 3 outliers, a direct comparison with the air temperature series recorded by automatic weather stations in the immediate surroundings, with similar geographic context (i.e. altitude, land use, topographic coherence), should be conducted. In addition, morphological and environmental characteristics of the caves (e.g. cave's development) should be considered. In fact, in our case study, 14 out of 19 caves have a planimetric development of less than 500 m, therefore, they could be more influenced by the external air temperature fluctuation and atmospheric events.

The application of the CAT-QC protocol highlighted several challenges associated with long-term cave air temperature monitoring, particularly those related to sensor stability and accessibility. Amongst the recommendations that we can suggest to minimize these challenges in future monitoring efforts we encourage deploying multiple sensors in parallel at each monitoring location to provide immediate cross-validation, allowing for the early detection of sensor drift or failure.

In future protocols we are incorporating rigorous pre- and post-deployment calibration checks against a certified reference thermometer. This practice would quantify sensor drift over the monitoring period, allowing for more accurate data correction during the QC process.

To minimize non-climatic biases, future protocols should require standardized sensor shielding and placement guidelines (e.g., minimum distance from cave walls or floor) to ensure that recorded temperatures accurately reflect the air temperature rather than the rock surface temperature.

Despite these challenges, the CAT-QC demonstrated effectiveness on series with different characteristics, making it adaptable to diverse temperature regimes with a marked or lower degree of seasonal fluctuation. Therefore, it was crucial to apply the protocol to both the more stable internal series and to the entrance series which show a higher temperature fluctuation. Its structured methodology ensures a standardization of cave temperature datasets, making them suitable for large-scale climate studies. The protocol's ability to identify and rectify anomalies enhances the overall integrity of the collected data, reducing errors and improving the confidence researchers can place in their analyses. Moreover, its flexibility enables adaptation to various cave environments, helping to shed light on the intricate interactions between cave microclimates and external

Cave name	Entrance series first run						Entrance series second run					
	Total records	Non-sense values	Type 1 outliers	Type 2 outliers	Type 3 outliers	Total outliers	Total records	Non-sense values	Type 1 outliers	Type 2 outliers	Type 3 outliers	Total outliers
Grotta delle Arenarie	6928	13	13	21	50	84	5768	0	0	36	44	80
Buco dell'Aria Calda	5527	0	1	174	5	180	4696	0	0	35	40	75
Grotta Occidentale del Bandito	3693	0	41	164	67	272	3136	0	0	32	21	53
Grotta di Bercovei	6939	4	4	46	62	112	5768	0	0	11	35	46
Grotta di Bossea	6797	0	0	186	0	186	6152	0	0	7	53	60
Grotta inferiore del Caudano	4999	0	2	238	39	279	4632	0	1	48	63	112
Grotta di Chiabrano	5252	0	1	37	93	131	4280	0	0	9	55	64
Grotta La Custreta	1381	0	1	9	4	14	1372	0	0	16	24	40
Tuna dal Diaou	7210	3	4	28	44	76	5800	0	0	26	79	105
Tana della Dronera	5011	29	29	34	86	149	4624	0	0	11	48	59
Balma Fumarella	5711	0	0	56	46	102	4336	0	0	40	23	63
Buca del Ghiaccio della Cavallaria	3316	0	0	17	50	67	2792	0	0	9	34	43
Ghieisa d'la Tana	6931	0	0	89	0	89	6008	0	0	6	69	75
Buco del Maestro	3433	4	4	71	42	117	3100	0	4	4	42	50
Grotta Testa di Napoleone-1	1698	0	0	14	20	34	1496	0	0	0	23	23
Buco del partigiano	5516	0	1	22	103	126	5748	0	0	49	66	115
Grotta del Pugnetto	4717	0	7	170	103	280	4368	0	0	25	55	80
Borna del Servais B	4860	0	0	155	124	279	4364	0	0	0	35	35
Sotterranei del forte (A) di Vernante	5501	0	0	196	47	243	4724	0	0	38	18	56

Table 4. Comparison between the results emerged from the CAT-QC first and second runs with the entrance series. For each cave is shown from the left to the right the total records, the number of records flagged as non-sense values, as type 1 outliers, as type 2 outliers, as type 3 outliers, and the total number of outliers. The total records shown in the second run is the number of records that underwent manual check at the end of the first run and therefore cleaned from non-sense values and records found to be erroneous among the outliers.

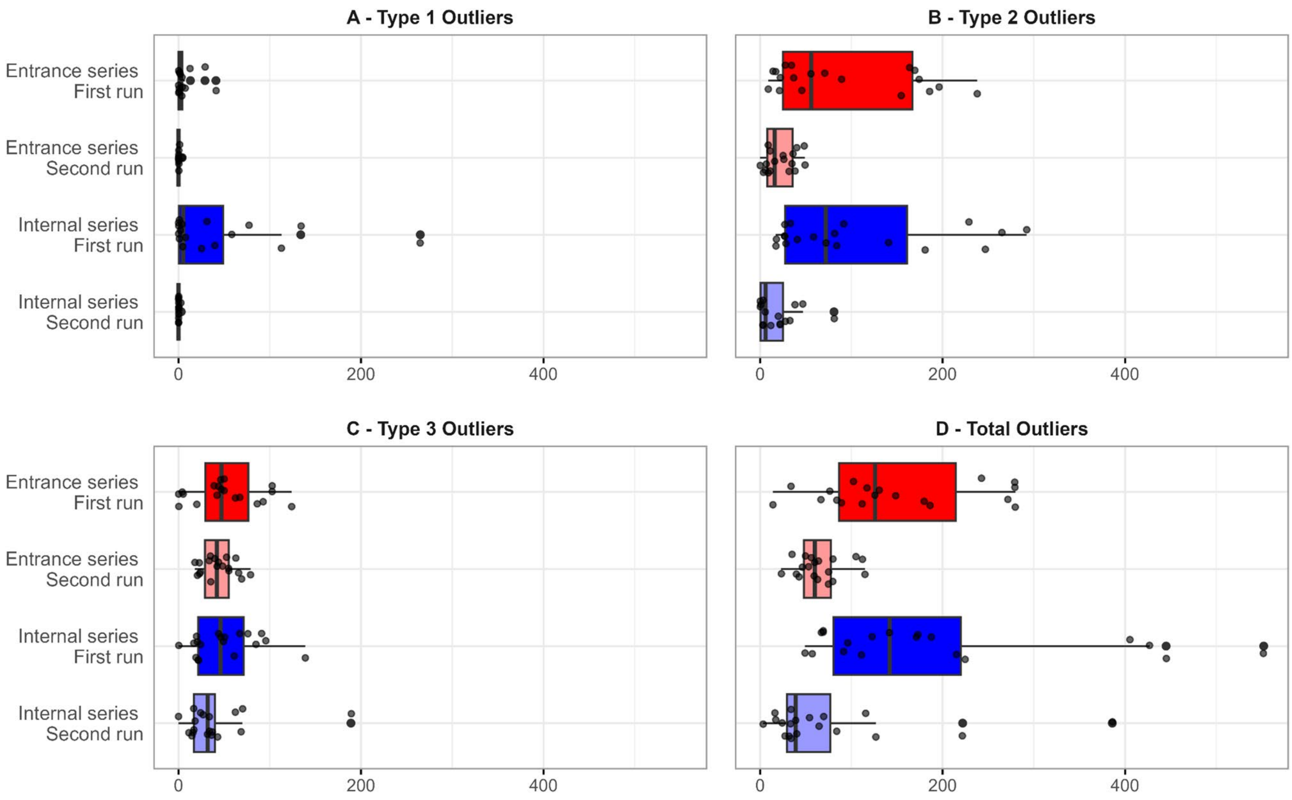

Fig. 2. Comparison of the number of records of type 1 **(A)**, type 2 **(B)**, type 3 **(C)** outliers and of the total outliers' number **(D)** between the outlier detection results of the CAT-QC first and second run for the entrance (red) and internal (blue) series.

climatic conditions. While further improvements, including automation and expanded testing, would enhance its efficacy, the CAT-QC protocol remains a powerful tool for advancing knowledge in this specialized field.

Conclusions

The stability of the subterranean climate offers unique opportunities for ecological and climatological research. This study developed and applied a comprehensive quality control protocol (CAT-QC) to assess the reliability of cave air temperature data, which we tested on data collected in the Piedmont region (Italy) from autumn 2020 to winter 2024–25. The protocol was designed to identify data gaps, non-sense values and outliers, ensuring a standardized approach for future climatological studies in subterranean environments.

The CAT-QC proved to be an effective method for refining cave temperature datasets with greater or lesser seasonality, providing a reproducible approach that can be applied to other climatologically significant underground settings worldwide. By systematically identifying and addressing data inconsistencies, the protocol enhances the accuracy and comparability of cave temperature records, facilitating broader adoption of systematic cave climate monitoring. This, in turn, contributes to interdisciplinary climate change research by improving the reliability of long-term climate datasets. Indeed, we believe that implementing standardized quality control methodologies is essential to support conservation efforts and climate change mitigation strategies on a broader scale.

Data availability

The cave air temperature time series used in this study will be made publicly available in Figshare and linked to a forthcoming data paper currently in preparation. The R script implementing the CAT-QC protocol is provided as Online Resource 1 (Supplementary Material) to ensure transparency and reproducibility of the quality control process. The data is also available upon request by directly contacting Marco Isaia (marco.isaia@unito.it).

Received: 10 July 2025; Accepted: 28 December 2025
Published online: 16 January 2026

References

1. Sánchez-Fernández, D. et al. Subterranean environments as model systems. *Subterr. Biol.* **25**, 1–7. https://doi.org/10.3897/subtbiol.25.23530 (2018).
2. Mammola, S. et al. Climate change going deep: alterations in cave ecosystems. *Anthropocene Rev.* **6** (1–2), 98–116. https://doi.org/10.1177/2053019619851594 (2019).

3. Medina, M. J. et al. Temperature variation in caves and its ecological relevance. *Sci. Rep.* **13**, 20735. https://doi.org/10.1038/s41598-023-47728-9 (2023).
4. Culver, D. C. & Pipan, T. *The Biology of Caves and Other Subterranean Habitats* 2nd edn. (Oxford University Press, 2019).
5. Badino, G. Underground meteorology: what's the weather underground? *Acta Carsol.* **39** (3), 427–448 (2010).
6. Moore, G. W. & Sullivan, G. N. Out of phase seasonal temperature fluctuations in Cathedral cave. *Geol. Soc. Am. Bull.* **76**, 313 (1964).
7. Smithson, P. A. Inter-relationships between cave and outside air temperatures. *Theor. Appl. Climatol.* **44** (1), 65–73 (1991).
8. Obleitner, F., Trüssel, M. & Spötl, C. Climate warming detected in caves of the European alps. *Sci. Rep.* **14**, 27435. https://doi.org/10.1038/s41598-024-78658-y (2024).
9. Domínguez-Villar, D., Lojen, S., Krklec, K., Baker, A. & Fairchild, I. J. Is global warming affecting cave. *temperatures? Clim. Dyn.* **45**, 569–581. https://doi.org/10.1007/s00382-014-2226-1 (2015).
10. Pipan, T., Petrič, M., Šebela, S. & Culver, D. C. Climate change and subsurface interactions in Postojna–Planina cave system. *Reg. Environ. Change.* **19** (2), 379–389. https://doi.org/10.1007/s10113-018-1349-z (2019).
11. Fairchild, I. J. & Baker, A. *Speleothem Science: From Process to Past Environments* (Wiley-Blackwell, 2012).
12. Badino, G. Cave temperatures and global Climatic change. *Int. J. Speleol.* **33** (1–4), 103–113 (2004).
13. Vaccarelli, I. et al. Multilevel and context-dependent climate effects in subterranean ecosystems. *One Earth.* **6** (11), 1510–1522. https://doi.org/10.1016/j.oneear.2023.09.001 (2023).
14. Malard, F. et al. Diversity patterns of stygobiotic crustaceans across Europe. *Freshw. Biol.* **54** (4), 756–776. https://doi.org/10.1111/j.1365-2427.2009.02180.x (2009).
15. Sánchez-Fernández, D., Galassi, D. M. P., Wynne, J. J., Cardoso, P. & Mammola, S. Subterranean ecosystems in climate change agendas. *Nat. Clim. Chang.* **11** (6), 458–459. https://doi.org/10.1038/s41558-021-01057-y (2021).
16. Ferreira, R. L., Howarth, F. G. & Mammola, S. Moving forward: Mitigating the effect of climate change in subterranean ecosystems. In: Innovative Conservation Techniques and Perspectives. Academic Press, pp 265–275 (2025).
17. Mejía-Ortíz, L., Christman, M. C., Pipan, T. & Culver, D. C. What's the temperature in tropical caves? *PLoS One.* **15** (12), e0237051. https://doi.org/10.1371/journal.pone.0237051 (2020).
18. Sánchez-Fernández, D. et al. Thermal niche estimators and poor dispersers under climate change. *Sci. Rep.* **6**, 23381. https://doi.org/10.1038/srep23381 (2016).
19. Mammola, S., Goodacre, S. L. & Isaia, M. Climate change May drive cave spiders to extinction. *Ecography* **41** (1), 233–243. https://doi.org/10.1111/ecog.02902 (2018).
20. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2023). https://www.R-project.org/
21. Grolemund, G. & Wickham, H. Dates and times made easy with lubridate. *J. Stat. Softw.* **40** (3), 1–25. https://doi.org/10.18637/jss.v040.i03 (2011).
22. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
23. Zeileis, A. & Grothendieck, G. Zoo: S3 infrastructure for regular and irregular time series. *J. Stat. Softw.* **14** (6), 1–27. https://doi.org/10.18637/jss.v014.i06 (2005).
24. Klein Tank, A. M. G. et al. Daily dataset of 20th-century surface air temperature and precipitation series. *Int. J. Climatol.* **22**, 1441–1453. https://doi.org/10.1002/joc.773 (2002).
25. Acquaotta, F. & Fratianni, S. The importance of the quality and reliability of the historical time series for the study of climate change. *Rev. Bras. Climatol.* **14** https://doi.org/10.5380/abclima.v14i1.38168 (2014).
26. Baronetti, A., Fratianni, S. & Acquaotta, F. High-resolution temperature climatology of an alpine region using local-scale regression models. *Theor. Appl. Climatol.* **137**, 45–58. https://doi.org/10.1007/s00704-018-2595-5 (2019).
27. Giaccone, E., Fratianni, S. & Acquaotta, F. A new methodology for quality control of temperature data: application to the meteorological station of Cuneo (NW Italy). *Geograph Fis. Dinam Quat.* **38** (1), 41–50. https://doi.org/10.4461/GFDQ.2015.38.04 (2015).
28. Dunn, R. J. H. et al. HadISD: A quality-controlled global synoptic report database. *Clim. Past.* **8** (5), 1649–1679. https://doi.org/10.5194/cp-8-1649-2012 (2012).
29. European Commission, Centre, J. R. & Dosio, A. Mean and extreme climate in Europe under 1.5, 2, and 3°C global warming. Publications Office. https://data.europa.eu/doi/ (2020). https://doi.org/10.2760/826427
30. World Meteorological Organization. WMO verifies new European temperature record. Accessed 30 Jan 2024 (2024).
31. Maxim Integrated DS1922L/DS1922T iButton temperature loggers with 8KB datalog memory. (2015). https://www.analog.com/media/en/technical-documentation/data-sheets/ds1922l-ds1922t.pdf
32. Mammola, S. et al. Ecology and sampling of the milieu souterrain superficiel (MSS). *Sci. Nat.* **103**, 1–24. https://doi.org/10.1007/s00114-016-1413-9 (2016).
33. Piano, E., Mammola, S., Nicolosi, G. & Isaia, M. Advancing tourism sustainability in show caves. *Cell. Rep. Sustain.* **1** (3). https://doi.org/10.1016/j.crsust.2024.100020 (2024).
34. Domonkos, P. Efficiencies of inhomogeneity-detection algorithms: comparison of different detection methods and efficiency measures. *J. Climatol.* **2013**, 1–15. https://doi.org/10.1155/2013/390945 (2013).
35. WDSEN. iButton temperature loggers DS1921G DS1922 DS1923. (2022). http://www.wdsen.com/DS1922L-en.html
36. Cimenti, L. et al. (eds) (submitted) Western Alps Cave Air Temperature Dataset. Submitted manuscript.
37. Catasto Speleologico Piemontese e Valdostano. Speleological Inventory. Accessed January 2025. (2025). https://catastogrotte-piemonte.net/

Acknowledgements

This research was supported by P.R.I.N 2022 DEEP CHANGE (2022MJSYF8), funded by the Ministry of Universities and Research (Italy) through the EU-Next Generation EU funds, and projects CAVEAT (CSTO162355) and CAVELAB (ORTO11T92F), funded by the Compagnia di San Paolo and University of Turin. Additional support was provided by Biodiversa+ (project "DarCo"), the European Biodiversity Partnership under the 2021–2022 BiodivProtect joint call for research proposals, co-funded by the European Commission (GA N°101052342) and with the funding organisations Ministry of Universities and Research (Italy), Agencia Estatal de Investigación – Fundación Biodiversidad (Spain), Fundo Regional para a Ciência e Tecnologia (Azores, Portugal), Suomen Akatemia – Ministry of the Environment (Finland), Belgian Science Policy Office (Belgium), Agence Nationale de la Recherche (France), Deutsche Forschungsgemeinschaft e.V. (Germany), Schweizerischer Nationalfonds (Grant N° 31BD30_209583, Switzerland), Fonds zur Förderung der Wissenschaftlichen Forschung (Austria), Ministry of Higher Education, Science and Innovation (Slovenia), and the Executive Agency for Higher Education, Research, Development and Innovation Funding (Romania).

Author contributions

All authors contributed to the study conception and design. Marco Isaia, Elena Piano, Anna Piquet, Stefano Mammola, Giuseppe Nicolosi, Alice Cimenti, and Lorenzo Cresi carried out fieldwork. Material preparation, data collection were performed by Marco Isaia, Elena Piano, Anna Piquet, Stefano Mammola, Giuseppe nicolosi and Alice Cimenti and Lorenzo Cresi. Lorenzo Cresi, Alice Cimenti, Antonella Senese, and Fiorella Acquaotta carried out analyses and prepared figures. Alice Cimenti and Lorenzo Cresi wrote the first draft. Software creation and data analysis were performed by Lorenzo Cresi, Alice Cimenti, Fiorella Acquaotta and Antonella Senese.All authors contributed to the writing with comments and additions to the text. All authors approved the final submission.

Funding

The project "DEEP CHANGE" was supported by the EU-Next Generation EU funds which funded the Italian Research Program PRIN 2022 (Grant number: 2022MJSYF8).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-34366-9.

Correspondence and requests for materials should be addressed to L.C.