



Impact of the measurement uncertainty on the monitoring of thermal comfort through AI predictive algorithms

Nicole Morresi¹, Sara Casaccia¹, Marco Arnesano², Gian Marco Revel¹

¹ Dipartimento di Ingegneria Industriale e Scienze Matematiche, Università Politecnica delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy

² Università Telematica eCampus, 22060 Novedrate, CO, Italy

ABSTRACT

This paper presents an approach to assess the measurement uncertainty of human thermal comfort by using an innovative method that comprises a heterogeneous set of data, made by physiological and environmental quantities, and artificial intelligence algorithms, using Monte Carlo method (MCM). The dataset is made up of heart rate variability (HRV) features, air temperature, air velocity and relative humidity. Firstly, MCM is applied to compute the measurement uncertainty of the HRV features: results have shown that among 13 participants, there are uncertainty values in the measurement of HRV features that ranges from $\pm 0.01\%$ to $\pm 0.7\%$, suggesting that the uncertainty can be generalized among different subjects. Secondly, MCM is applied by perturbing the input parameters of random forest (RF) and convolutional neural network (CNN) algorithm, trained to measure human thermal comfort. Results show that environmental quantities produce different uncertainty on the thermal comfort: RF has the highest uncertainty due to the air temperature (14 %), while CNN has the highest uncertainty when relative humidity is perturbed (10.5 %). A sensitivity analysis also shows that air velocity is the parameter that causes a higher deviation of thermal comfort.

Section: RESEARCH PAPER

Keywords: Monte Carlo Simulation; Artificial Intelligence; thermal comfort measurement; measurement uncertainty

Citation: Nicole Morresi, Sara Casaccia, Marco Arnesano, Gian Marco Revel, Impact of the measurement uncertainty on the monitoring of thermal comfort through AI predictive algorithms, Acta IMEKO, vol. 10, no. 4, article 34, December 2021, identifier: IMEKO-ACTA-10 (2021)-04-34

Section Editor: Carlo Carobbi, University of Florence, Gian Marco Revel, Università Politecnica delle Marche and Nicola Giaquinto, Politecnico di Bari, Italy

Received October 10, 2021; **In final form** December 6, 2021; **Published** December 2021

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partially done in the framework of the project “RenoZEB - Accelerating Energy renovation solution for Zero Energy buildings and Neighbourhoods”, funded by the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No.768718.

Corresponding author: Sara Casaccia, e-mail: s.casaccia@staff.univpm.it

1. INTRODUCTION

In these modern times, it is commonly accepted by the scientific community that the term thermal comfort refers to a very subjective concept, that differs from one individual to another [1]. In fact, thermal comfort is affected by several factors that include personal, psychological, physical and environmental diversities and therefore, the personalization of thermal comfort measurement is increasingly required to provide more tailored and customized indoor environments, that start with the single individual and end with a comfortable and satisfactory environment [2], [3], [4].

Physiological, psychological and personal parameters are progressively included in the measurement setup of thermal comfort, to encourage a more customized environment carefully tailored to the distinct preferences of the occupants that live in it.

To put together all these specific parameters, personalized thermal comfort measurement typically requires an extensive range of sensing devices that make up a sensors network. Since each sensor is properly characterized by measurement uncertainty, it is significant to include in the measurement process how this uncertainty, associated with each parameter, affects the final measurement of thermal comfort.

This aspect is even more justified since, in all research fields, it is routinely accepted by the scientific community that the result of a measurement process promptly loses its meaning if an uncertainty value is not associated with it [5]. The traditional concept of measurement uncertainty must therefore be associated and applied to modern techniques, as in the case of artificial intelligence (AI) algorithms, which are finding more and more space in the measurement process. In support of this claim, literature reports that when small variations are applied to data, and variations are represented by the measurement uncertainty

associated with the data collected, AI provides completely misleading results [6].

The reason behind the use of AI for personalized thermal comfort measurement is that the measured output is dependent on several parameters, i.e., physiological and environmental parameters, that compose the measurement system. Currently, the thermal sensation related to thermal comfort, is expressed using psychological scales, the most prominent of which is the ASHRAE 7-points scale, which is used in many comfort studies [7]. This scale is used to rate the thermal sensation vote (TSV), with 7 different verbal terms which are: “cold”, “cool”, “slightly cool”, “neutral”, “slightly warm”, “warm”, and “hot”. Each term is associated with a categorical number in the range from -3 (“cold”) to +3 (“hot”).

The application of AI in the measurement of human thermal comfort in the built environment, is becoming preferable because the adopted dataset is composed of two categories of parameters, which comprise the environmental and physiological parameters. A third category is also added to this dataset, made up of subjective parameters, expressed by the TSV. The union of these quantities generates a complex and heterogeneous dataset, making necessary the employment of AI, which comprises models comparable to a black box [8]. There are strong non-linear relationships among the parameters included in this heterogeneous dataset, which suggests that this relationship should be explored with more complex and high-level algorithms, included in the AI field [9], [10].

AI measures its performance in terms of accuracy, which is expressed through different metrics, depending on the type of algorithm adopted, and the type of data being measured; for example, when the measured quantity is binary, or is made by discrete values, or class, performances are measured using accuracy, recall or precision; in case of AI used for measuring continuous quantities, the most used metrics are the mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared error (MSE). Of course, the accuracy of AI-based models, is strictly linked to the selected algorithm as well as the quality of the dataset, which is deeply connected to the uncertainty of the measured data [8] [11].

The method for the estimation of the uncertainty is described in the Guide to the expression of uncertainty in measurement (GUM) [12], which is based on the law of propagation of uncertainties (LPU). Given the strong non-linearity of the relationship that exists between the input quantities and the output quantities of AI algorithms, the evaluation of the uncertainty associated with the output quantity is hard to assess if the LPU should be used.

When it comes to non-linear relationship, the GUM has proposed in its supplement [13], the evaluation of the uncertainty by using the Monte Carlo method (MCM), which allows assessing the uncertainty even if the relationship among quantities is not linear, and the analytical equation between quantities is not known, but it can be assumed as a black box, as happens in the context of AI models. The MCM represents a more practical alternative to the conventional LPU method, when it is not possible to effectively verify the hypothesis assumed by the LPU [11]. For this reason, this work studies the impact of the measurement uncertainty of a sensors network to measure thermal comfort, using AI models.

To quantify and access the factors that affect the outcome of a measurement process through MCM, there are two methodologies that can provide this information: sensitivity analysis and a measurement uncertainty analysis [14], [15], [16],

[17]. Both of them are now applied to AI models, the first one to analyze how much each uncertainty weights on the measurement of the model's outcome and the second to identify and quantify what are the main sources of uncertainties in the measurement [18].

As mentioned before, thermal comfort personalization is based on the measurement of a heterogeneous set of data, (i.e., environmental, physiological and personal parameters).

Regarding physiological quantities, the object of this study is the addition of Heart Rate Variability (HRV) into the process for the measurement of thermal comfort: in fact, literature has repeatedly found the relationship that exists between some measures (or indices) derived from HRV and the thermal discomfort of the participant [19]–[22]. In particular, HRV is used to extract additional features, such as a specific quantity named LF/HF, which has a relationship with the thermal comfort of the user [9], [23]. More in detail LF/HF is the ratio between two HRV-derived parameters which are the low frequency components (LF) and the high frequency components (HF). Literature has found that these two quantities represent the activity of the autonomic nervous system, responsible for the management of thermoregulation, and therefore are used to evaluate human thermal comfort. According to a previous study [23], [9], [24], physiological features of HRV and environmental quantities can be used in combination with AI models, to predict the TSV of user exposed to discomfort environmental conditions. Results have shown that the TSV can be measured using two AI algorithms, which are the Random Forest (RF), that belongs to the ML class and Convolutional Neural Networks (CNN), which are part of DL branch, with a MAPE of 20 % and 21 % respectively [23]. The further step is therefore to know how much the measurement uncertainty of each parameter involved in the measurement of thermal comfort, impacts the final measurement of TSV.

To this aim, the paper is structured as follows:

- a) First of all, MCM is applied to the raw HRV signals coming from different participants, to see how the uncertainty in the measurement of HRV impacts the computation of the HRV features, since they are fed into AI algorithms, to extract the TSV. The standard uncertainty, associated to the measurement of HRV, chosen for this analysis is ± 4 ms, computed from a previous study in which the smartwatch was compared to a reference method while participants were sitting at rest [25]. This preliminary analysis can be useful to establish if the measurement uncertainty of the smartwatch, used to assess HRV and derive its features, has the same impact on each participant's data, by having similar measuring uncertainty for each.
- b) Secondly, MCM is applied to both physiological and environmental parameters used as input to the AI algorithms (RF and CNN), to evaluate how the measurement uncertainty propagates to the output, in black box methodology such as AI algorithms. The uncertainty used for the environmental quantities is the standard uncertainty provided from the datasheet of each device, while HRV is perturbed with different values of uncertainty that range from ± 4 ms to ± 100 ms. This range of values is chosen because literature has highlighted that, depending on the activity level performed by the participants, the

uncertainty associated with the measurement of the HRV in resting condition is ± 4 ms, while is greater than or equal to 100 ms when the user is performing a motion test [25], [26]. The research described in [25] and [26] were necessary since commercial smartwatches are not provided with precise datasheets that define the measurement uncertainties of the measured quantities. This is one of the problems being encountered in literature when it comes to working with low-cost sensors and often sensors that are not designed for research purposes.

Finally, the results are used to perform a sensibility analysis related to ± 4 ms of uncertainty, to examine the contribution of the uncertainties associated to the TSV measurement, in relation to the uncertainties of the input parameters.

The paper is organized as follows: Section 2 firstly describes how MCM is applied to compute the uncertainty in the measurement of HRV features, using a wearable smartwatch; secondly, it is explained the process to compute the measurement uncertainty of human thermal comfort, using a heterogeneous dataset and trained AI algorithms. Section 3 provides the results associated with the two methodologies described in Section 2, and the result of the sensitivity analysis. Section 4 presents the discussion of results and Section 5 provides the conclusion and the innovative aspect of this work.

2. MATERIALS AND METHODS

In this paper, a procedure to evaluate the impact of the uncertainty of the input data on the thermal sensation measurement output is applied. The aim is to introduce and combine a traditional technique that is the MCM for the estimation of the measurement uncertainty, with AI models.

The methodology described in Figure 1 is adopted: the first assumption is that each data coming from a measurement process is characterized by two quantities which are the data itself and the associated uncertainty. Collected data from sensors network are merged together to build the heterogeneous dataset, that comprises the set of data and the associated uncertainties; therefore, when AI is applied in the measurement process, the model takes as input also the uncertainty of the collected data. Therefore, it is expected that the results of the AI model derive from two main aspects: the first one is the intrinsic structure of

the algorithm, while the second one is uncertainty associated with the input data [6]. The evaluation of the impact of the uncertainty associated with the input data in the measurement output, when AI is applied, is a paramount result. There is one intermediate aspect that should be considered, which is the impact of the uncertainty of the AI models, that will be combined with the uncertainty of the measurement instrument. The last part of Figure 1 deals with two types of analysis which are the sensitivity analysis (SA) and the uncertainty analysis, performed by MCM.

The procedure of the MCM consists in using data, acquired through real experiments, and perturbing them by assigning different measurement uncertainties or perturbations. Each input parameter is modified, by adjusting it with a different perturbation, one at a time, while the other input variables are kept unchanged. Based on the obtained results, the effect of different measurement uncertainties on the prediction of the TSV is observed through the simulation of perturbed data.

The described methodology summarizes the steps to evaluate the measurement uncertainty of different parameters, used as input variables (or predictors) in AI algorithms, applied for the measurement of personalized thermal comfort, expressed through the TSV. The GUM supplement provides the step necessary to perform the MCM. From a general point of view, MCM provides a general approach to numerically approximate the *cumulative density function (cdf)* of the output of a certain quantity $y = f(x)$. The main concept behind MCM is that every sample of the input quantity x_i , chosen from a predetermined distribution can be used. In this way, by taking a random sample of each input x_i from its related probability distribution function (*pdf*), it is possible to estimate a possible result of the output y and the associated uncertainty.

To explain how MCM is adapted in the context of this research, the following steps were applied:

- 1) The number of trials M of Monte Carlo trials is set to 200,000. M is the number of output quantity values that need to be selected; in this study, it is chosen a priori. Usually, a number of trials equal to 10^6 are considered, which is the number of trials that should provide a coverage interval of 95 %, as reported in the GUM supplement [13]. Since it is commonly accepted that, the higher the number of trials, the higher is expected the convergence of the results, authors have set the number of Monte Carlo trials (M) to 200,000, as reported from the literature [14];
- 2) M vectors x_i , $i = 1, \dots, M$, were generated, by selecting randomly from the probability density function (PDF) of each input quantity $[x_{HRV}, x_{t_a}, x_{RH}, x_{v_a}]$ in order to realize a set of possible input that can be associated to the input quantity. The random sample is obtained from a gaussian distribution, with uncertainties described in Table 1.
- 3) For each vector generated in step 2, the corresponding output y (or TSV in this case) is computed, yielding to M vector output quantity values;
- 4.1) Point 3 is applied to estimate the uncertainty associated with the measurement of the HRV features, described in Section 2.1;
- 4.2) In addition, point 3 is applied, perturbing the features used as input in the AI models, described in Section 2.2;
- 5) The representation G of the distribution function for Y is computed, starting from the set of M output of Y ;

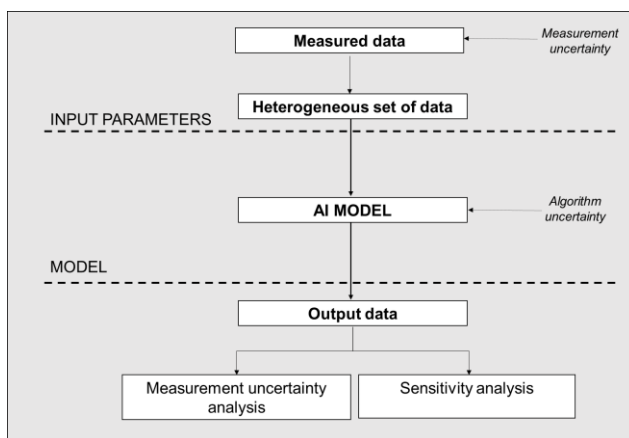


Figure 1. Conceptual description of the procedure adopted to study the impact of the measurement uncertainty in the context of AI model.

Table 1. Characteristics of the measurement instrument used in the test.

Input Quantities	Manufacturer	Model	Standard Uncertainty	Distribution
Air temperature (t_a)	Thermo Sensor GmbH	PT100 (4 wired)	$\pm 0,1$ °C	Normal
Relative humidity (RH)	Ahlborn	FHAD46C41A	$\pm 2\%$ of reading	Normal
Air velocity (v_a)	Ahlborn	FVAD05TOK300	$\pm 3\%$ of reading	Normal
Heart rate variability (HRV)	Samsung	Samsung Galaxy Watch	± 4 ms	Normal

- 6) G is used to compute an estimate $\mathbf{y}$ of Y and the covariance matrix $\mathbf{u}_i$ associated with $\mathbf{y}$;
- 7) G is used to compute the appropriate coverage region for Y , for a stipulated coverage probability p ;

In this paper, we will refer to a generical model, with a number i of input $\mathbf{x}_i$ and one output $\mathbf{y}$, which is the measured quantity. The input parameters used in this study to measure TSV, were chosen on the basis of a previous study explained below: 13 participants were recruited to carry out a test consisting in exposing each participant in a semi-controlled room, in which environmental quantities were changed by varying the air temperature of the room.

The sample population involved in the test is made of healthy people, not suffering from heart disease in order not to have perturbation on the collected HRV; the age of the sample population is 31 years $\pm$ 6 years old and participants were 7 females and 6 males. As reported in previous works [9], [23], the personal characteristics of the user (e.g., gender) are not included in this research since literature has shown that the analysis of environmental variables and personal characteristics in the prediction of user thermal comfort, does not lead to statistically significant results. Moreover, authors did not consider differences in age, since literature reports that higher temperatures are preferred by elderly people, while lower temperatures are more suitable for the younger population; the sample population of this experiment is not made of older people and therefore authors do not deepen this aspect [9].

More in detail, the profile temperature of the test room was created as follows: the test room was previously set at 15 °C for 5 minutes, then the temperature set-point was set to 26 °C, then temperature was set back to 15 °C. The total duration was lengthened up to 100 minutes. The experiment was performed both in January 2020 and January 2021; the outdoor daily average temperatures were 8 °C in January 2020 and 4 °C in January 2021. During the experiment, the participants were sitting at a workstation and could perform light office activities (e.g., reading, working on the laptop).

During the test, different quantities were monitored and used to train and test AI models for predicting the TSV of the users, which was collected during the test [23].

The monitored quantities for measuring the TSV to express human thermal comfort in this study were:

- Air temperature (t_a);
- Relative humidity (RH);
- Air velocity (v_a);
- Heart Rate Variability (HRV).

The list above contains both environmental quantities (t_a , RH, v_a) and physiological quantities (HRV).

The procedure with which the physiological and environmental data were acquired consisted in varying the environmental conditions of a semi-controlled room, while the sensors network acquired continuously these parameters. Participants were informed of the purpose of the study and gave their informed consent; in addition, the study was validated by

the ethical committee at Università Politecnica delle Marche. The study was carried out in compliance with the principles laid down in the Declaration of Helsinki, in accordance with the Guidelines for Good Clinical Practice.

The following paragraphs explain the methodology used in this research: in particular, Section 2.1 aims at defining the measurement uncertainty associated with each HRV feature, derived from HRV, using MCM. Section 2.2 explains how to implement MCM in combination with AI models, to assess the TSV of participants, using environmental and physiological data.

2.1 Monte Carlo approach on HRV features

In this first part of the analysis, MCM is used to estimate how the uncertainty of the device through which the HRV is collected, influences the computation of the HRV features, that will be later used to evaluate human thermal comfort. HRV signal, which is made by the distance in time of two subsequent RR peaks of the ECG trace, is divided into time frames from which it is possible to extract some indices (or features). Each time frame is built as follows: the first time frame corresponded to 5 minutes of the HRV signal, which is considered the minimum time duration recommended for computing HRV spectral analysis. After the extraction of the first window, a new window was computed by appending a new HRV sample interval of the signal, while the oldest sample was removed from the beginning of the window; the process was repeated until the end of the signal.

According to the literature, several HRV features, to be extracted from the HRV signal, were identified. Time-domain HRV features $f(HRV_t)$ are a collection of statistical and geometrical indices for the measurement of the variability in the HRV sequence that act as indices to interpret the oscillations of cardiac cycles. The $f(HRV_t)$ statistical indices computed in this study are: standard deviation of RR intervals (SDANN), root mean square of RR (RMSSD), mean value of RR (MEAN), median of RR (MEDIAN), percentage of the difference between adjacent RR intervals that differs more than 50 ms (PNN50), percentage of the difference between adjacent RR intervals that differs more than 25 ms (PNN25). In addition, HRV studies imply the use of frequency-domain features $f(HRV_f)$, which are useful for the understanding of the stationarity or stability of the HRV signal. To obtain the frequency-domain analysis, the first power spectral density (PSD) was computed through the autoregression modelling-based method that has proven to provide better resolution. Each frequency band was then computed: LF (0.04 Hz - 0.15 Hz) and HF (0.15 Hz - 0.4 Hz), LF/HF, HF/LF and the total power spectrum (TP). Non-linear features $f(HRV_{nl})$ were also computed through the Poincaré plot. The Poincaré plot is a graphical representation of an HRV time series along the cartesian plane: the X-axis contains one HRV sample, while the Y-axis contains the following HRV sample. The Poincaré plot provides two additional features obtained by adjusting the point cloud of the figure formed into an ellipse; the first index (SD1) represents the dispersion of points perpendicular to the identity line, while the second one

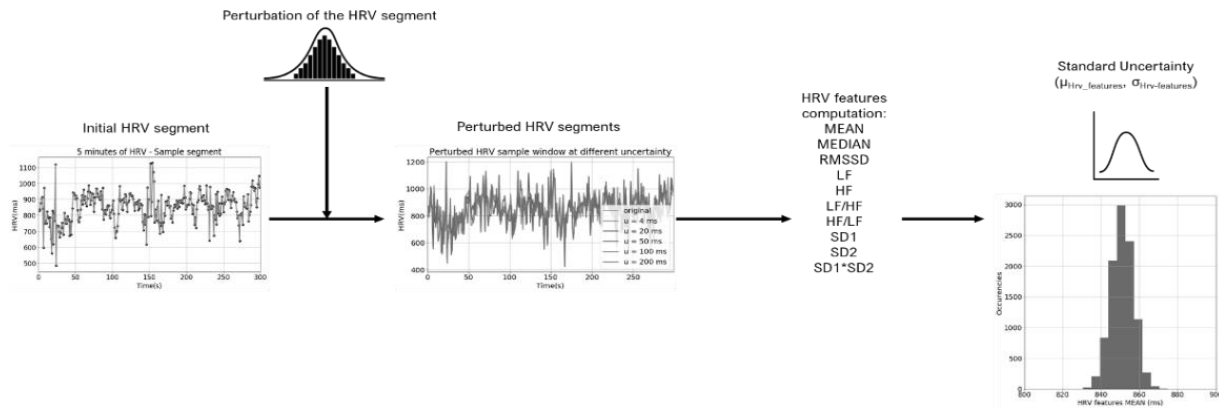


Figure 2. Description of the MCM adopted to evaluate the impact of the smartwatch uncertainty on the HRV features, that will be used to measure and assess human thermal comfort.

(SD2) is the dispersion of point along the identity line of the Poincaré plot.

In this section, there is a first analysis made specifically on the individual features of the HRV. In practice, we want to assess how it is the uncertainty measurement of each feature, when the smartwatch is affected by a specific uncertainty in the measurement of HRV; the procedure is summarized in Figure 2.

First of all, one HRV window of 5 minutes is considered. This window will be perturbed with a random sample coming from a normal distribution with mean equal to 0, and a standard deviation equal to U , for 200,000 iterations. The perturbed HRV segment will be used to compute HRV features in order to establish the impact of the uncertainty. The final uncertainty is computed as the standard deviation of the resulting HRV, by a coverage factor $k = 2$.

For the simulation study, 13 segments of HRV of a duration of 5-minutes, were extracted from the set of experimental procedures described previously. The duration of 5 minutes is required because is the minimum time required to compute short term HRV features.

2.2 Monte Carlo approach on AI models for measuring human thermal comfort

For this second part of the analysis, the set of HRV features, explained in Figure 3, was chosen, [23]. Figure 3 shows the approach used to apply MCM and AI: a set of input parameters was used to train RF regression algorithm and a CNN, to predict

the TSV, output of the algorithm. Once the models are trained, the procedure consists in perturbing the input parameters (HRV , t_a , RH , v_a) one at a time, with different uncertainties, maintaining constant the other quantities, and applying the trained model to each perturbed set of features, to obtain the final measurement of the TSV. It is worth noting that the current analysis is conducted locally, by choosing one observation of the whole dataset; the final result is therefore associated to the chosen observation.

Characteristics of the sensors used to collect environmental parameters, and the measurement uncertainty employed to perturb the different parameters are shown in Table 1.

Sensitivity Analysis

A sensitivity analysis (SA) is performed to identify and quantify what are the main sources of uncertainties in the measurement. Ideally, SA is the methodology for studying how the uncertainty of the output provided by a model, can be associated, qualitatively or quantitatively, to different uncertainties of the input parameters of the model [27], [28].

The GUM Supplement also provides the specifications for assessing the sensitivity coefficients; more in detail, it is explained that MCM is not sufficient to fully compute sensitivity coefficients but provides a methodology to compute the influence of each input quantity on the output quantity [13].

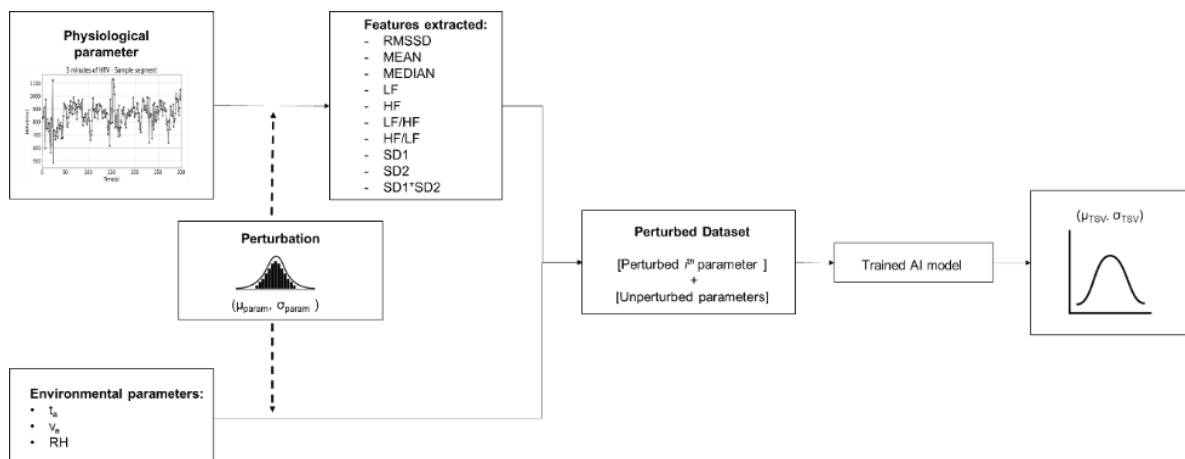


Figure 3. Description of the procedure adopted to apply MCM to the measurement of thermal comfort, with a heterogeneous set of data.

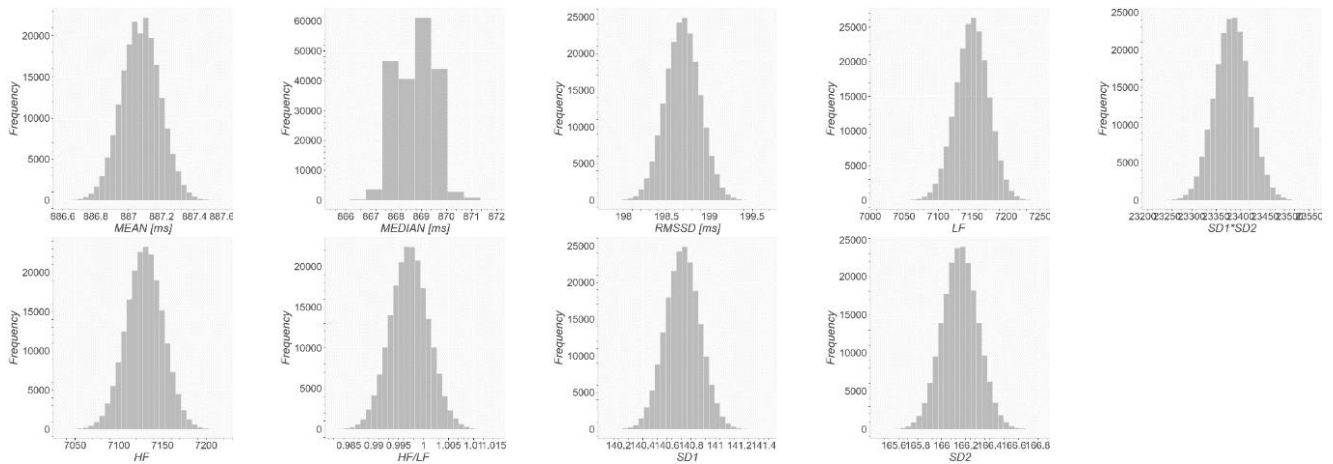


Figure 4. Histogram obtained from MCM with 200,000 iterations applied to the HRV features.

The procedure implies that one input quantity should be perturbed, and the remaining input quantities should be kept at their best estimates, in order to obtain the *pdf* of the output quantity, depending only on the variable perturbed. By using this procedure, the GUM supplement proposes an approach that can be representative as a generalization of the approximate partial-derivative formula; in particular, it is reported that the sensitivity coefficients can be approximated as the ratio of the standard deviation of the resulting model values and the standard uncertainty associated with the best estimate of relevant input quantity, as reported in Eq. 1:

$$c_i = \frac{u_i(y)}{u(x_i)}, \quad (1)$$

where c_i is the sensitivity coefficient, $u(x_i)$ is the standard uncertainty associated with the i th input estimate x_i , that contributes to the standard uncertainty $u_i(y)$.

3. RESULTS

3.1 Impact of the uncertainty of measurement on the HRV features

In this section, results of the MCM used to analyse the propagation of the uncertainty in the measurement of HRV that propagates to the computation of the HRV features, are shown.

The simulated results are used to obtain the frequency histogram of each HRV features that are computed with the simulation, to obtain the related uncertainty. For each HRV features the values of the standard uncertainty computed with the MCM are shown. The frequency histograms that come out of one simulation for one participant, are displayed in Figure 4.

Table 2 contains the uncertainty associated with the computation of each HRV feature, expressed in percentage, among the 13 participants. It can be seen that the minimum measurement uncertainties are obtained when computing the MEAN feature (± 0.01 % of reading), while the highest values are associated with the measurement of LF/HF and HF/LF (± 0.7 % of reading).

Physiological quantities, such as HRV, can vary among participants and therefore it is interesting to understand the impact of the uncertainty on each feature, divided according to the related domain, among the different users, to see if it is possible to establish a generalized standard uncertainty associated to each HRV features, in relation to the value of the uncertainty of the device which measures it.

The uncertainty associated with the features in the frequency domain is higher than the features in the time or the features in the non-linear domain. A reason that can explain this result is that the frequency components are more sensitive to the uncertainty of the device (± 4 ms) because they physically reflect

Table 2. Measurement uncertainty computed for each HRV features used to estimate the TSV, among the 13 participants.

ID	Uncertainty (%)									
	MEAN	RMSSD	MEDIAN	LF	HF	LF/HF	HF/LF	SD1	SD2	SD1*SD2
1	0.01	0.10	0.09	0.34	0.30	0.4	0.4	0.10	0.1	0.2
2	0.01	0.08	0.09	0.31	0.25	0.4	0.4	0.08	0.1	0.1
3	0.01	0.29	0.07	1.30	0.81	1.5	1.5	0.29	0.3	0.5
4	0.01	0.11	0.08	0.53	0.55	0.6	0.6	0.11	0.1	0.2
5	0.01	0.25	0.06	0.93	0.58	1.1	1.1	0.25	0.2	0.4
6	0.01	0.10	0.08	0.30	0.28	0.4	0.4	0.10	0.1	0.2
7	0.01	0.16	0.10	0.57	0.59	0.6	0.6	0.16	0.1	0.2
8	0.01	0.04	0.13	0.26	0.12	0.3	0.3	0.04	0.0	0.1
9	0.01	0.12	0.08	0.31	0.32	0.5	0.5	0.12	0.1	0.2
10	0.01	0.24	0.06	0.92	0.58	1.1	1.1	0.24	0.2	0.4
11	0.01	0.09	0.07	0.31	0.25	0.4	0.4	0.09	0.1	0.1
12	0.01	0.30	0.08	0.66	0.79	1.0	1.0	0.30	0.2	0.4
13	0.01	0.06	0.10	0.26	0.19	0.3	0.3	0.06	0.1	0.1
μ	0.01	0.1	0.1	0.5	0.4	0.7	0.7	0.1	0.1	0.2

Table 3. Results of MCM applied to compute the measurement uncertainty $u_i(y)$ of the measurement of the TSV, both for the RF and the CNN algorithm. The uncertainty was computed considering a coverage factor k equal to 2.

	RF				CNN			
	HRV	t_a	RH	v_a	HRV	t_a	RH	v_a
Standard uncertainty	± 4 ms	$\pm 0,1$ °C	$\pm 2\%$ of reading	$\pm 3\%$ of reading	± 4 ms	$\pm 0,1$ °C	$\pm 2\%$ of reading	$\pm 3\%$ of reading
TSV Estimate (μ)	1	1.4	1.4	1.4	-0.5	-0.5	-1.2	-1.2
$u_i(y)$ [TSV unit] (σ)	7.46E-14	0.2	1.78E-13	0.0105	0.015	0.015	0.126	0.029
$u_i(y)$ [%]	7.24E-12	14	1.28E-11	0.8	2.825	2.9	10.5	2.5

the oscillatory activity of the HRV; for this reason, it can be verified that a small variation of HRV can lead to a greater variation of features in the frequency domain.

3.2 Impact of the uncertainty on AI models for measuring human thermal comfort

This section reports the results used to compute the measurement uncertainty of the TSV, using MCM applied to AI models, which are RF and CNN. Table 3 contains the results which represent the contribution of the uncertainty of each input quantity (HRV, t_a , RH, v_a) into the AI model, to estimate the associated uncertainty in the output TSV. Each *pdf* of the input quantity is used, one at a time, according to Monte Carlo simulation, to obtain the resulting pdf of the TSV. The following table contains the results of an MCM conducted by perturbing each parameter, one at a time, with the respective uncertainty $u_i(y)$, where i are the perturbed parameters.

The simulation regarding the impact of the HRV uncertainty of the TSV was made for ± 4 ms. Since HRV measurement through smartwatches is particularly prone to motion artifacts, more simulations were conducted by simulating different HRV uncertainties, which are $U = [4, 10, 20, 50, 100]$ ms. The results of the simulations are shown in Figure 5: the contribution of HRV uncertainty increases after 50 ms, suggesting that up to this value, the uncertainty of the devices that measure HRV is still acceptable for assessing the TSV.

On the other hand, environmental sensors that measure t_a , RH and v_a , are less subjected to variations in the uncertainty of the measurement of TSV. Environmental quantities produce different uncertainty on the TSV, which depends on different parameters. RF has highest uncertainty due to the t_a ($u_{t_a}(y) = 14\%$), while CNN has highest uncertainty when RH is perturbed ($u_{RH}(y) = 10.5\%$). This different result is explainable if we consider that the two algorithms perform with different rules and can be considered as black boxes.

Sensitivity Analysis

In Table 4 the partial uncertainty budget due to each parameter is given according to GUM uncertainty framework; the sensitivity coefficients were evaluated by Monte Carlo simulation. According to the table, SA highlights that air velocity and air temperature are the parameters that most affect the TSV prediction, respectively for RF and CNN.

RH, and HRV perturbed with ± 4 ms are the less sensitive parameters in the measurement of TSV, while v_a is the parameter that mostly contributes to the TSV measurement uncertainty. This outcome is consistent with the test procedure, which

implied that during the experiment, to generate thermal discomfort, the window was opened, and the participant reported experiencing greater discomfort.

4. DISCUSSION

Depending on the kind of algorithm trained than can be both ML and DL algorithms, there are different quantitative results in the uncertainty of the TSV, using an MCM, with 200,000 iterations, in relation to the environmental parameters. RF provides greater uncertainty when t_a is perturbed with an uncertainty of $\pm 0,1$ °C; on the contrary, CNN exhibits a higher uncertainty when RH is perturbed with $\pm 2\%$ of reading. Physiological parameter (HRV signal), when perturbed with ± 4 ms of uncertainty, impacts the resulting TSV with 2.8 % of uncertainty and $> 0.001\%$ when it comes with CNN and RF respectively.

In addition, MCM on AI models for measuring human thermal comfort (Section 2.2) is applied to one observation of the entire dataset, and therefore it is a local analysis that is strongly related to the trained model. The overall methodology presented in this paper can be applied to brand new models, that do not include an analytical equation, to compute the uncertainty associated with the output of the model.

MCM is a powerful tool that can be used to simulate the results of the impact of more than one measurement uncertainty, such in the case of the analysis presented in the manuscript in Section 3.2, in which HRV is perturbed with a set of uncertainty that ranges from ± 4 ms to ± 100 ms. The impact of this study is that MCM can be applied in a variety of circumstances in which the analytical model that represents the relationship between input and output is not known, such in the case of AI models, but the initial perturbation can be simulated. When the measurement uncertainty of the device is not available, the uncertainty can be therefore hypothesized and simulated, or eventually a calibration to determine the uncertainty can be performed.

5. CONCLUSION

This paper presents an approach to assess the measurement uncertainty of human thermal comfort, expressed in terms of TSV, by using a method that comprises a heterogeneous set of data, made by physiological and environmental quantities, and AI models. The objective is therefore to quantify the measurement uncertainty of the TSV, while the user is performing light-office activities, by using GUM guidelines and applying MCM.

Table 4. Sensitivity coefficients compute for the two algorithms.

	HRV	t_a	RH	v_a
Standard Uncertainty	± 4 ms	$\pm 0,1$ °C	$\pm 2\%$ of reading	$\pm 3\%$ of reading
Sensitivity Index RF	1.87E-14	1.8	4.61E-13	5.8
Sensitivity Index CNN	0.004	0.09	0.32	16.03

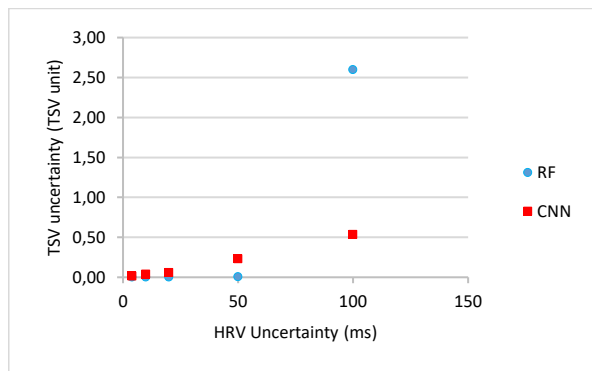


Figure 5. Trend of the measurement uncertainty of the TSV, in response to different HRV uncertainties measurement, obtained with 200,000 iterations.

A preliminary analysis was conducted to assess the impact of the measurement uncertainty of the instrument used to collect HRV, which is a commercial smartwatch. MCM was applied to compute the uncertainty associated with the features extracted from HRV, which will be later fed into an RF and CNN model. Results have shown that among 13 participants, there are uncertainty values in the measurement of features that ranges from $\pm 0.01\%$ to $\pm 0.7\%$, suggesting that among different users the uncertainty can be generalized.

Then, MCM was applied by perturbing a set of parameters (HRV , t_a , RH and v_a) to compute the uncertainty in the measurement of the TSV, using RF model and a CNN. Results have shown that environmental quantities produce different uncertainty on the TSV. RF has the highest uncertainty due to t_a uncertainty ($U = 1\%$), while CNN has the highest uncertainty when RH is perturbed ($U = 10.5\%$). On the other hand, the sensitivity analysis that expresses the relationship between the TSV and the input parameters highlights that v_a is the parameter that causes the greatest variation on the TSV.

ACKNOWLEDGEMENT

The activity presented in this work was partially carried out within the RenoZEB project, funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 768718.

REFERENCES

- [1] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, L. Xie, X. Ou, Convolutional Neural Network and Kernel Methods for Occupant Thermal State Detection using Wearable Technology, in 2018 International Joint Conference on Neural Networks (IJCNN), Jul. 2018, pp. 1–8.
DOI: [10.1109/IJCNN.2018.8489069](https://doi.org/10.1109/IJCNN.2018.8489069)
- [2] S. Tham, R. Thompson, O. Landeg, K. A. Murray, T. Waite, Indoor temperature and health: a global systematic review, *Public Health*, vol. 179, Feb. 2020, pp. 9–17.
DOI: [10.1016/j.puhe.2019.09.005](https://doi.org/10.1016/j.puhe.2019.09.005)
- [3] F. Kobiela, R. Shen, M. Schweiker, R. Dürichen, Personal thermal perception models using skin temperatures and HR/HRV features: comparison of smartwatch and professional measurement devices, in Proceedings of the 23rd International Symposium on Wearable Computers, London, United Kingdom, Sep. 2019, pp. 96–105.
DOI: [10.1145/3341163.3347737](https://doi.org/10.1145/3341163.3347737)
- [4] A. Aryal, B. Becerik-Gerber, Thermal comfort modeling when personalized comfort systems are in use: Comparison of sensing and learning methods, *Building and Environment*, vol. 185, Nov. 2020, p. 107316.
DOI: [10.1016/j.buildenv.2020.107316](https://doi.org/10.1016/j.buildenv.2020.107316)
- [5] A. Ferrero, S. Salicone, A Comparison Between the Probabilistic and Possibilistic Approaches: The Importance of a Correct Metrological Information, *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 3, Mar. 2018, pp. 607–620.
DOI: [10.1109/TIM.2017.2779346](https://doi.org/10.1109/TIM.2017.2779346)
- [6] A. Fornaser, M. D. Cecco, P. Bosetti, T. Mizumoto, K. Yasumoto, Sigma-zrandom forest, classification and confidence, *Meas. Sci. Technol.*, vol. 30, no. 2, Dec. 2018, p. 025002.
DOI: [10.1088/1361-6501/aaf466](https://doi.org/10.1088/1361-6501/aaf466)
- [7] W. Liu, X. Tian, D. Yang, Y. Deng, Evaluation of individual thermal sensation at raised indoor temperatures based on skin temperature, *Building and Environment*, vol. 188, Jan. 2021, p. 107486.
DOI: [10.1016/j.buildenv.2020.107486](https://doi.org/10.1016/j.buildenv.2020.107486)
- [8] M. Khanafer, S. Shirmohammadi, Applied AI in instrumentation and measurement: The deep learning revolution, *IEEE Instrumentation Measurement Magazine*, vol. 23, no. 6, Sep. 2020, pp. 10–17.
DOI: [10.1109/MIM.2020.9200875](https://doi.org/10.1109/MIM.2020.9200875)
- [9] I. Pigliautile, S. Casaccia, N. Morresi, M. Arnesano, A. L. Pisello, G. M. Revel, Assessing occupants personal attributes in relation to human perception of environmental comfort: Measurement procedure and data analysis, *Building and Environment*, vol. 177, June 2020, p. 106901.
DOI: [10.1016/j.buildenv.2020.106901](https://doi.org/10.1016/j.buildenv.2020.106901)
- [10] S. Casaccia, G. M. Revel, G. Cosoli, L. Scalise, Assessment of Domestic Well-Being: From Perception to Measurement, *IEEE Instrumentation Measurement Magazine*, vol. 24, no. 6, Sep. 2021, pp. 58–67.
DOI: [10.1109/MIM.2021.9513641](https://doi.org/10.1109/MIM.2021.9513641)
- [11] J. Singh, L. A. Kumaraswamidhas, N. Bura, N. Dilawar Sharma, A Monte Carlo simulation investigation on the effect of the probability distribution of input quantities on the effective area of a pressure balance and its uncertainty, *Measurement*, vol. 172, Feb. 2021, p. 108853.
DOI: [10.1016/j.measurement.2020.108853](https://doi.org/10.1016/j.measurement.2020.108853)
- [12] Evaluation of measurement data—guide for the expression of uncertainty in measurement, *JCGM 100*: 2008.
- [13] I. BIPM, I. IFCC, I. ISO, IUPAP, and OIML, Evaluation of measurement data—supplement 1 to the “Guide to the expression of uncertainty in measurement”—propagation of distributions using a Monte Carlo method. Joint Committee for Guides in Metrology, *JCGM 101*: 2008, International Organization for Standardization, Geneva, 2008.
- [14] A. Chen, C. Chen, Comparison of GUM and Monte Carlo methods for evaluating measurement uncertainty of perspiration measurement systems, *Measurement*, vol. 87, Jun. 2016, pp. 27–37.
DOI: [10.1016/j.measurement.2016.03.007](https://doi.org/10.1016/j.measurement.2016.03.007)
- [15] V. Isaiev, O. Velychko, Metrological characterisation of current transformers calibration unit for accurate measurement, *Acta IMEKO*, 10 (2021) 2, Jun. 2021, pp. 6–13.
DOI: [10.21014/acta_imeko.v10i2.918](https://doi.org/10.21014/acta_imeko.v10i2.918)
- [16] F. Ezedine, J.-M. Linares, W. M. W. Muhamad, J.-M. Sprauel, Identification of most influential factors in a Virtual Reality tracking system using hybrid method, *Acta IMEKO*, 2 (2013) 2, Jan. 2014, pp. 20–27.
DOI: [10.21014/acta_imeko.v2i2.136](https://doi.org/10.21014/acta_imeko.v2i2.136)
- [17] T. T. Nguyen, A. Amthor, C. Ament, Algorithm for a high precision contactless measurement system, *ACTA IMEKO*, vol. 2, no. 2, Art. no. 2, Jan. 2014, doi: 10.21014/acta_imeko.v2i2.82
- [18] Sensitivity Analysis for Importance Assessment - Saltelli - 2002 - Risk Analysis - Wiley Online Library. Online [Accessed 15 June 2021]
<https://onlinelibrary.wiley.com/doi/full/10.1111/0272-4332.00040>
- [19] H. Zhu, H. Wang, Z. Liu, D. Li, G. Kou, C. Li, Experimental study on the human thermal comfort based on the heart rate variability

- (HRV) analysis under different environments, *Sci. Total Environ.*, vol. 616–617, Mar. 2018, pp. 1124–1133.
DOI: [10.1016/j.scitotenv.2017.10.208](https://doi.org/10.1016/j.scitotenv.2017.10.208)
- [20] K. N. Nkurikiyeyezu, Y. Suzuki, G. F. Lopez, Heart rate variability as a predictive biomarker of thermal comfort, *J Ambient Intell Human Comput*, vol. 9, no. 5, Oct. 2018, pp. 1465–1477.
DOI: [10.1007/s12652-017-0567-4](https://doi.org/10.1007/s12652-017-0567-4)
- [21] H. Zhu, H. Wang, D. Li, Z. Xiao, H. Su, X. Kuang, Evaluation of the Human Thermal Comfort under Simulated Weightlessness: an Experimental Study Based on the Power Spectrum Analysis of the Heart Rate Variability, *Microgravity Sci. Technol.*, vol. 31, no. 1, Feb. 2019, pp. 73–83.
DOI: [10.1007/s12217-018-9669-7](https://doi.org/10.1007/s12217-018-9669-7)
- [22] W. Liu, Z. Lian, Y. Liu, Heart rate variability at different thermal comfort levels, *Eur J Appl Physiol*, vol. 103, no. 3, Jun. 2008, pp. 361–366.
DOI: [10.1007/s00421-008-0718-6](https://doi.org/10.1007/s00421-008-0718-6)
- [23] N. Morresi, S. Casaccia, M. Sorcinelli, M. Arnesano, A. Uriarte, J. I. Torrens-Galdiz, G. M. Revel, Sensing physiological and environmental quantities to measure human thermal comfort through Machine Learning techniques, *IEEE Sensors Journal*, 2021, pp. 12322–12337.
DOI: [10.1109/JSEN.2021.3064707](https://doi.org/10.1109/JSEN.2021.3064707)
- [24] A. Affanni, Dual-channel electrodermal activity and an ECG wearable sensor for measuring mental stress from the hands, *Acta IMEKO*, 8 (2019) 1, pp. 56 - 63.
DOI: [10.21014/acta_imeko.v8i1.562](https://doi.org/10.21014/acta_imeko.v8i1.562)
- [25] N. Morresi, S. Casaccia, M. Sorcinelli, M. Arnesano, G. M. Revel, Analysing performances of Heart Rate Variability measurement through a smartwatch, 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Bari, Italy, 1 June-1 July 2020, pp. 1-6.
DOI: [10.1109/MeMeA49120.2020.9137211](https://doi.org/10.1109/MeMeA49120.2020.9137211)
- [26] N. Morresi, S. Casaccia, G. M. Revel, Metrological characterization and signal processing of a wearable sensor for the measurement of heart rate variability, in 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Lausanne, Switzerland, 23-25 June 2021, pp. 1–6.
DOI: [10.1109/MeMeA52024.2021.9478713](https://doi.org/10.1109/MeMeA52024.2021.9478713)
- [27] G. M. Revel, E. Sabbatini, M. Arnesano, Development and experimental evaluation of a thermography measurement system for real-time monitoring of comfort and heat rate exchange in the built environment, *Meas. Sci. Technol.*, vol. 23, no. 3, Feb. 2012, p. 035005.
DOI: [10.1088/0957-0233/23/3/035005](https://doi.org/10.1088/0957-0233/23/3/035005)
- [28] O. Sima, M.-C. Lépy, Application of GUM Supplement 1 to uncertainty of Monte Carlo computed efficiency in gamma-ray spectrometry, *Applied Radiation and Isotopes*, vol. 109, Mar. 2016, pp. 493–499.
DOI: [10.1016/j.apradiso.2015.11.097](https://doi.org/10.1016/j.apradiso.2015.11.097)