

Uncertainty assessment in building energy performance with a simplified model

Fally Titikpina^{1,*}, Antoine Caucheteux², Abderafi Charki¹, and David Bigaud¹

¹ L'UNAM, LARIS, Systems Engineering Research Laboratory, University of Angers, 62 avenue Notre Dame du Lac, Angers, France

² Cerema, risk, environment, mobility and development research Center, 23 avenue de l'Amiral Chauvin, 49136 Les Ponts-de-Cé, France

Received: 14 January 2015 / Accepted: 7 September 2015

Abstract. To assess a building energy performance, the consumption being predicted or estimated during the design stage is compared to the measured consumption when the building is operational. When valuing this performance, many buildings show significant differences between the calculated and measured consumption. In order to assess the performance accurately and ensure the thermal efficiency of the building, it is necessary to evaluate the uncertainties involved not only in measurement but also those induced by the propagation of the dynamic and the static input data in the model being used. The evaluation of measurement uncertainty is based on both the knowledge about the measurement process and the input quantities which influence the result of measurement. Measurement uncertainty can be evaluated within the framework of conventional statistics presented in the Guide to the Expression of Measurement Uncertainty (GUM) as well as by Bayesian Statistical Theory (BST). Another choice is the use of numerical methods like Monte Carlo Simulation (MCS). In this paper, we proposed to evaluate the uncertainty associated to the use of a simplified model for the estimation of the energy consumption of a given building. A detailed review and discussion of these three approaches (GUM, MCS and BST) is given. Therefore, an office building has been monitored and multiple temperature sensors have been mounted on candidate locations to get required data. The monitored zone is composed of six offices and has an overall surface of 102 m².

Keywords: Building energy performance, Uncertainty evaluation, GUM Method, Bayesian Approach, Monte Carlo

Nomenclature

n	Number of observations
N	Number of parameters
f	Functional relationship between parameters
R_{ay}	Horizontal Solar radiations
$\Delta\Psi$	Difference between the outdoor and indoor temperature
T_{out}	Outdoor Temperature
T_{in}	Indoor Temperature
M	Number of Monte Carlo trials
$Q_{building}$	Building heat loss
Q_{cv}	Heat loss due to conduction and ventilation
Q_{inf}	Heat loss due to infiltrations
Q_I	Heat loss due to building inertia
A_o	Heat gain due to occupation
A_s	Heat gain due to solar radiations
ξ	Building Inertia coefficient

U_b	Average Conduction coefficient through the building envelop
Φ	Envelop Solar radiations absorption coefficient
β	Envelop conduction coefficient
x_i, x_j	Input parameters
$u(y)$	Output uncertainty
γ	Infiltration coefficient due to wind
v	Wind Speed
S	Surface of the zone
y	Output parameter
q	Air flow rate in the building
c_p	Air Specific Heat
ω	Week of the year
X	A selected input quantity

1 Introduction

The regulation of energy consumption in all area has become today a worldwide concern. In a context of energy resources increasing scarcity, especially oil and gas, it has become essential and vital to limit the energy demand of

* Correspondence: fally.titikpina@univ-angers.fr

some sectors especially in the area of buildings energy consumption. According to the International Energy Agency (IEA), at the current consumption trend, oil reserves will run out in less than 50 years, natural gas in 60 years and coal in 120 years while the energy demand is increasing substantially [1].

The building sector is the largest energy consumer in the European Union. In 2008, it engages 40% of the total energy consumption and gives off 36% of CO₂ emissions. Among all economic area in Europe, the construction industry is the area of huge energy demand. It is responsible of more than 20% of greenhouse gas emissions [2]. With the continuous demographic growth, these numbers are in continual progression and are making the building energy consumption matter an international issue, which is managed by public policies [3].

A building energy performance is defined as the amount of energy actually consumed or estimated to meet the different needs associated with a standardized use of the building, which may include heating, cooling, hot water, ventilation and lighting [3]. This standardized use is somewhat unpredictable and usually lead to a gap in the energy predicted during the building design stage and the real energy consumption when the building is operational [4].

The challenge in the construction industry is to ensure that the predicted energy performance of a given building is achieved with a minimum of uncertainty at the operational stage [4]. This situation comes to provide high performance in thermal comfort and reliability at lowest cost possible. This task can be achieved primarily by finding the sources of the performance gap and later by estimating the amount of uncertainty associated to the performance being predicted [5].

The parameters that describe a building such as the static parameters (wall absorption, inertia coefficient, conduction and infiltration coefficients due to wind) and dynamic parameters (weather and occupation) are relevant sources of uncertainty but does not represent the entire causes when estimating building energy consumption. De wit [6] introduced various causes of uncertainty in the building performance assessment: the specification uncertainty coming from inaccurate description of the building and its systems, the modelling uncertainty which comes from the physical model itself, the errors introduced in the model during simulation, and those arising from external conditions and occupants [7]. To better sense uncertainty and its influence on building performance, it is important that these sources of uncertainty are known. Several research works were focused on this problem, although few case studies included this aspect into their analyses. In this paper, the uncertainties coming from measured data are evaluated in different ways.

Most of time, the uncertainty associated to measured data comes not only from the random errors but also from the data itself. These parameters are affected by the measurement methods, the equipment being used, the quality of the data and also by the assumptions made during the analysis process. A set of methods are used to es-

timate uncertainty in measurements. Among them, the GUM (Guide to the expression of Uncertainty in Measurements) [8], the Monte Carlo Simulation [9], and the Bayesian inference [10, 11] are the common ones.

These methods are based on both the knowledge about the measurement process and on the impact of parameters being involved. The model equation represents the knowledge about the measurement process and the relation between the quantities while their degree of belief is expressed by probability density functions.

In this paper, we tried to estimate the energy demand of a selected zone of an office building with a simplified model, (the BEECHAM model [12] which combines physical principles and statistical method and its associated uncertainty with the three different approaches cited above. The content is organized as follows; Section 2 reviews the uncertainty estimation methods, each one with its base principles and the presentation of the BEECHAM model, followed in Section 3 by the application made and in Section 4 by the results and discussions.

2 Materials and methods

2.1 GUM principle

The Guide to the expression of Uncertainty in Measurements [8] is frequently used to evaluate measurement uncertainty. It has been published in 1993 and constitutes one of the reference documents in the area of measurement uncertainty. When repeated measurements are made on a variable, or when information is available on the data or on the measurement process, the Guide specifies the estimation of the associated uncertainty. Then, given the estimates and the associated uncertainties of the input quantities, the GUM describes how to determine the best estimation of the measurand and its associated uncertainty as well as a confidence interval.

According to the GUM, uncertainty is a parameter associated to the result of a measurement, which characterizes the dispersion of the values that could be reasonably attributed to the quantity being measured (measurand). Two types of methods are presented. Among them, the type A method of evaluation of uncertainty, carried out by the statistical analysis of series of observations. In most cases, the measurand y is not measured directly, but is determined from the measurement of N other quantities x_1, x_2, \dots, x_N through a functional relationship f , given in equation (1).

$$y = f(x_1, x_2, \dots, x_N). \quad (1)$$

The GUM method is based on the concept of a measurement equation (Eq. (1)), which represents the process to determine the best estimate of y and its associated standard uncertainty $u(y)$ from the estimates and associated standard uncertainty of the input quantities x_1, x_2, \dots, x_N [8]. In this equation the variables are the input quantities and the measurement result is obtained by introducing the estimates of the inputs in

the model. Then, the uncertainty of the output quantity is calculated from the measurement model by propagation in equation (2). The GUM Uncertainty framework is then said to be a method of propagation of uncertainties. When the input quantities are correlated, the appropriate expression for the uncertainty associated with the result of a measurement is:

$$u^2(y) = \sum_{i=1}^N \frac{\partial f}{\partial x_i}^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left(\frac{\partial f}{\partial x_i} \right) \left(\frac{\partial f}{\partial x_j} \right) u(x_i, x_j), \quad (2)$$

where x_i and x_j are the estimates of the parameters X_i and X_j and $u(x_i, x_j)$ is the estimated covariance associated to x_i and x_j . To quantify the uncertainty of a measurement using the GUM approach, one does not need to compute any complicated calculation. The methodology is clearly stated in the GUM and simple to apply.

2.2 Monte Carlo method for uncertainty estimation

An updated supplement to the GUM, GUM S1 [13], specifies uncertainty evaluation in terms of a probability density function. It describes the numerical propagation of probability density function through the model mentioned before. The Monte Carlo methodology [14], as presented by the GUM S1, induces the propagation of the input quantities distributions through the model in order to provide the output probability distribution [15].

When the inputs density functions are defined, a number of Monte Carlo trials is selected. Generally, the greater the number of simulation is, the better are the results. The evaluation of the measurement uncertainty from Monte-Carlo simulation is done through the following steps.

After setting up the relationship between the analytical result y and the input quantities x_i , which is somewhat the model equation for the measurement process, the probability density function $p(x_i)$ of each input quantity x_i , have to be identified according to all the uncertainty sources. A number M of Monte Carlo trials is selected. Then, one generates M samples $x_{i1}, x_{i2}, \dots, x_{iM}$ of each x_i . From the estimated set of result, $y_{i1}, y_{i2}, \dots, y_{iM}$ the combined uncertainty $u(y)$ can be easily calculated by the propagation of the generated data through the functional relationship f as stated in equation (3). The Monte Carlo Method (MCM) is then said to be a method of propagation of distributions. However, when the inputs quantities are correlated, their probability density function are suitably estimated by a joint probability density function. Then, a confidence interval of the output y is estimated after the choice of a coverage probability p .

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = f \left(\begin{bmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1M} & x_{2M} & \dots & x_{NM} \end{bmatrix} \right). \quad (3)$$

The Monte Carlo Simulation is easy to apply and returns the information about the model distribution. However, it also has some limitations: the model simulation runtime could be long in some complex cases. Furthermore, the selection of the proper probability distribution functions for the model parameters may be difficult due to inaccurate data. In addition, the accuracy of the numerical simulation depends on the number of simulations.

Regarding the Monte Carlo Method, it is sometimes difficult to reliably specify the probability density function of each input variable. If some inputs are known to be correlated, it will be difficult to specify their joint probability distribution. Most of time, an approximation is made from the available data and information. Furthermore, related to the ISO-GUM approach, the Monte Carlo Simulation does not make the distinction between the uncertainties estimated by the statistical approach and those arising from other source such as sensors specifications, previous tests results or from the operator judgment. The difference should be done to estimate the probability density function of the input quantities.

2.3 Bayesian uncertainty quantification

For Bayesian Uncertainty Quantification, all the information available on a quantity is updated by the measurement in order to assign a value to this quantity. Bayesian statistics is based on Bayes Theorem [10]. This theorem comes from the axiom which states that given two propositions X and Y , the probability of both of them being true is equal to the product of the probability of X being true by the probability of Y being true given that X is true (Eq. (4)).

$$p(X/Y) = \frac{p(X)p(Y/X)}{p(Y)}. \quad (4)$$

Let us consider a sample of independent and identically distributed (i.i.d), x_1, x_2, \dots, x_N random variables, which depends on an unknown parameter θ , with a probability density function $p(x/\theta)$ $\pi(\theta)$ is the probability density function of θ and $L(x/\theta)$ the measurement likelihood function (Eq. (5)). The posterior distribution of θ denoted by $\pi(\theta/x)$, is the conditional distribution of θ given the observations of x , with $x = (x_1, x_2, \dots, x_N)$, the i.i.d variables considered above.

$$L(x/\theta) = \prod_{i=1}^n p(x_i/\theta). \quad (5)$$

The posterior distribution $\pi(\theta/x)$, gives the information about θ after measurement is done (Eq. (6)). This posterior distribution is found through the Bayesian inference.

$$\pi(\theta/x) = \frac{L(x/\theta) \pi(\theta)}{\int_{\theta} L(x/\theta) \pi(\theta) d\theta}$$

$$\pi(\theta/x) = \frac{\prod_{i=1}^n p(x_i/\theta) \pi(\theta)}{\int_{\theta} \prod_{i=1}^n p(x_i/\theta) \pi(\theta) d\theta}. \quad (6)$$



Fig. 1. Data-driven method flow chart.

Equation (6) can then be written as:

$$\pi(\theta/x) \propto \prod_{i=1}^n p(x_i/\theta) \pi(\theta). \quad (7)$$

Bayesian inference about a quantity [11], is made through the probability density function that describes the information acquired from measurement and the knowledge about the quantity before the measurement is performed [16, 17]. Bayes formula is a mechanism that combines the prior information on the parameters and the information provided by the measured data. The Bayesian approach applied with probability density functions is given in equation (8) by [18].

$$p(y/x_1, x_2, \dots, x_N) \propto p(x_1, x_2, \dots, x_N/y) p(y). \quad (8)$$

When estimating the uncertainty of a quantity by the Bayesian method, it is often difficult to assign prior information to the quantity being measured. To construct the prior density function, one should make use of all available prior information such as calibration data, data from other related experiments, personal judgment and other available evidences about the quantity. Still, since that information comes from a personal opinion and are unreliable, it can result to significant uncertainties [19].

2.4 A simplified building energy model: the BEECHAM model

One of the major concerns for the building sector nowadays is the evaluation and the prediction of a building energy performance. Many factors influence the energy behaviour of a building. Weather conditions, the occupants and their activities, the components such as lighting, the building thermal properties, the HVAC systems performance are the relevant parameters. These elements induce several variables, which are involved in the building energy model and can be split into two groups: the dynamic and the static parameters. The dynamic factors are known to change with time like the weather, the occupancy and the static parameters are the non-variable factors such as the building construction and thermal property. The energy model of buildings is then said to be complex and when predicting its energy consumption one might take into account all these factors.

Several methods rather complex or simplified are used to model the thermal performance of a building and its heat gains. Different techniques, going from statistical methods such as simple regression model to engineering methods based on physical principles are used. Artificial intelligence methods for example Artificial Neural Networks (ANN) [20, 21] or Support Vector Machines

(SVMs) [22] are also used to forecast the building energy consumption. In this paper, we will focus on an engineering and statistical based method especially a data driven model (Fig. 1) which uses physical and thermal principles to estimate the energy behaviour of a building.

According to Coakley et al. [23], data-driven models can be used to model a system behaviour with a minimum set of adjustable inputs. The statistical methods are used to construct physical model from measured data as illustrated in Figure 1. This kind of methods mainly focuses on the thermal phenomena occurring between a building and its surroundings to calculate precisely, at each time step the building heat load. The most influential and the most important factors to determine a building energy demand are the weather conditions. The outdoor temperature, the solar radiations and the wind speed have a non-negligible impact on the conduction through the building envelop and on the infiltrations through the building walls.

Caucheteux et al. [12] developed a model, the BEECHAM (Building Energy Efficiency Characterization Method) with which the heating requirement of a given building can be estimated based on its thermal balance.

The energy balance of a given building is assessed as previously said, by summing the heat losses and the heat gains under operational conditions. Then, the heat loss due to conduction through the building envelop Q_C , the heat loss due to ventilation Q_V , the heat loss due to infiltrations by wind and the stack pressure effect Q_{inf} , the heat gain due to solar radiations A_S , the heat load due to occupancy A_o (occupation, the use of electrical equipment's, lighting, ...), and the heat gain or loss (depending on the case) of energy in the building construction materials, which is called inertia Q_I , are summed in order to find the required amount of energy so that the building is balanced on a thermal basis. For a given period, the heating requirement of the building can be estimated by equation (9). Each element in these equations is denoted in the nomenclature part.

$$Q_{building} = Q_{cv} + Q_{inf} + Q_I - A_s - A_o \quad (9)$$

$$\Delta\Psi = \sum_{24} (T_{out} - T_{in}). \quad (10)$$

At a specified time step, the conduction and ventilation heat loss through a building envelop Q_{Cv} , depends on the outdoor and indoor temperature difference $\Delta\Psi(K)$, like specified in equation (10), on the ventilation average flow rate and on the conduction coefficient β (kW/(K)). β represents the irradiative and convective exchange of the walls in the overall zone and also the ventilation flow rate. Then, the conduction and ventilation heat loss through the building envelop can be expressed by equation (11).

$$Q_{cv} = \beta\Psi. \quad (11)$$

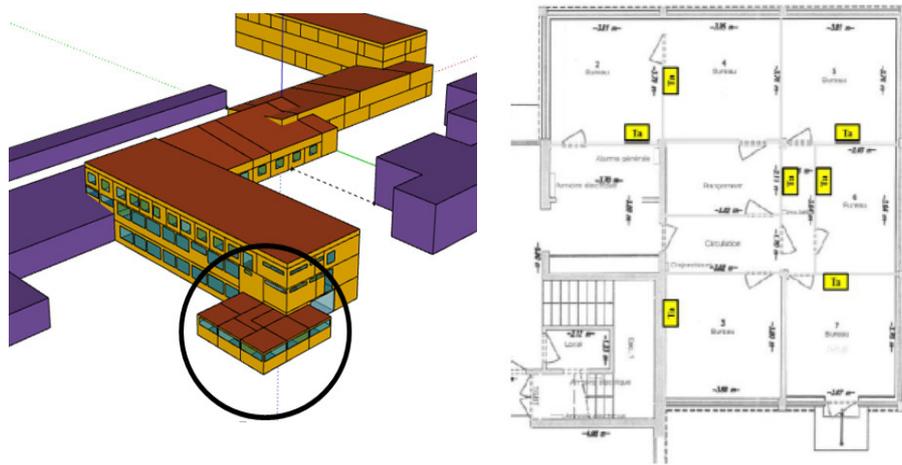


Fig. 2. (a) A 3-dimension view of the monitored building. (b) The monitored zone.

Two effects are mainly responsible of infiltrations: the stack effect and the wind pressure. Due to the simplicity of the model, the stack effect was not considered (the model is supposed to be applied to small-rise buildings). The infiltrations are then expressed by equation (12) where γ represents the infiltrations coefficient due to wind.

$$Q_{inf} = \gamma \left((\vartheta^2)^\alpha \right) \Psi, \quad \text{with } \alpha = 2/3. \quad (12)$$

The airflow through the building envelop is related to the difference of pressure created in the building within its environment. The parameter α is usually called the flow exponent and depends on the nature of the airflow (turbulent or laminar). By default, one retains usually a value of 2/3 [24]. An exponent close to 1 characterize laminar flow while a value of 1/2 represents a turbulent airflow.

The capability of a given building to charge and discharge the intake energy is called inertia. It is assumed that it depends essentially on the difference in temperature from one day to the next one. For example, if the considered day temperature is lower than the previous one, the building will give in the energy from the day before. This heat transfer can be expressed as:

$$Q_I = \xi \frac{\partial \Psi}{\partial t}, \quad (13)$$

where ξ is the building inertia coefficient and t is the considered time step. Usually, the heat gain due to solar radiations is estimated with an accurate description of the building (surfaces and windows being exposed to the sun, their orientations and titles angles ...). The BEECHAM model aims to avoid the description of the overall building which is time consuming and which parameters are often unknown. With these assumptions, the heat gain due to solar radiations is supposed to be constant during a time scale of one week and is stated in equation (14) where Φ represents the building geometry and R_{ay} the average solar radiations on horizontal.

$$A_s = \Phi R_{ay}. \quad (14)$$

Unlike the other coefficients β, γ, C_m which are stationary, Φ can be represented over the year, by a polynomial of degree equal or greater than 2 [12]. The degree of the polynomial depends on the presence or not of solar masks adjacent to the building. The time scale used to represent this polynomial is the week. A polynomial regression is then performed to estimate the coefficients a, b and c in equation (15). ω represents the week of the year and the resulting model for the estimation of the building energy demand is given by equation (16).

$$\Phi = a\omega^2 + b\omega + c \quad (15)$$

$$Q_{building} = \beta\Psi + \gamma \left(\vartheta^{4/3} \right) \Psi - (a\omega^2 + b\omega + c) R_{ay} + \xi \frac{\partial \Psi}{\partial t} - A_o. \quad (16)$$

The BEECHAM model, in equation (16), is somewhat simplified and can be qualified as a data-driven approach among the building energy models [23]. It uses the building monitored data through the model to predict the system thermal behaviour. Thus, with a minimal set of dynamic inputs (the outdoor and indoor temperature, the solar radiations, the wind speed), the user is able to estimate the heating load of a building. It also simplifies the measurements, as there is no analysis of the user behaviour or the solar gain evaluations. In this way, it allows to reduce the errors coming from the operator judgement.

3 Building description

The purpose of this work is to evaluate the heating requirement of a building zone and its associated uncertainty with the three methods described before: the GUM, the Monte Carlo Method and the Bayesian Method for uncertainty quantification. The monitored zone is a part of the ground floor of an administrative building located in Angers, France (Fig. 2). It is composed of six offices with a circulation area and a total surface of 102 m² (Fig. 2a).

Table 1. Estimation of the static parameters distributions.

Static parameters	Distributions
Envelop Conduction Coefficient β (kWh/°K)	$N(0.4, 0.057)$
Infiltration coefficient due to wind γ (kW/°K(km/h))	$N(9.57 \times 10^{-7}, 6.28 \times 10^{-7})$
Building inertia Coefficient ξ (kW/°K)	$N(-0.139, 0.1559)$
Solar radiation Coefficient $a\omega^2 + b\omega + c$ (m ²)	$0.07\omega^2 - 0.29\omega + 1.39$

Most of the building occupancy occurs between 8:00 and 18:00. The sensors have been fixed in all the offices to get the temperature data of the overall zone. The outdoor temperature, the solar radiations and the wind speed data for the monitored period are provided by a meteorological station implanted near the building being experimented. The temperature data are collected during one year from January to December 2013 in each office (Fig. 2b) with a time step of one hour. The average wind speed, horizontal solar radiation, and outdoor temperature are also measured each hour. For the occupancy rate, we assumed that there is one person in each office during work time in order that the occupancy rate is proportional to the electrical consumption being measured. Then, with the static parameters of the building estimated above, the heating requirement of the zone was calculated.

Regarding the test-bed, each sensor can be independent and works with a battery, which allows them to have their own memory to record the data. The sensors have been put on walls, at half height and away from openings (doors and windows).

4 Results and discussion

4.1 Inputs quantities uncertainty quantification

As the BEECHAM model is known to be a data-driven method, it requires prior data in order to model the building energy behaviour. For the static parameters (the envelop conduction coefficient, the infiltration coefficient due to wind, the building inertia coefficient and the solar radiation coefficient), as they are not changing in time, their daily uncertainty cannot be measured. A suitable methodology is required to assign them a variation range, in which they could be included, in order to do the uncertainty analysis. For this purpose, we used three months heating consumption measurements results, going from January to mid-March 2013 to get very wide upper and lower bonds. We realised 10 000 random samples from an uniform distribution and we calculated the daily energy consumption with these values.

To select the parameters ranges, we assumed that the model is calibrated. The ASHRAE guidelines [3] suggested the use of standardised statistical indices to sense the performance of a model. Among these criterion, the Coefficient of Variation of Root Mean Square Error, *CVRMSE* (Eq. (17)), allows to determine how well the energy demand model being used fits the data. For a model to be calibrated, ASHRAE sets the *CVRMSE* on a monthly basis to 15% and on an hourly basis to 30%. Thus, only

the set of static parameters with a *CVRMSE* < 20% are picked up (Tab. 1 and Fig. 3).

Equation (17) gives the *CVRMSE* of a model, with m_i and s_i , the measured and simulated data points, N_i the number of data points at regular interval (hourly, daily or monthly) and \bar{m} the mean of the measured data points.

$$CVRMSE = \frac{\sqrt{\left(\sum_{i=1}^{N_i} \frac{(m_i - s_i)^2}{N_i}\right)}}{\bar{m}}. \quad (17)$$

The probability density plots of the static model parameters β, γ, ξ, a, b and c are determined according to the previous results (Fig. 3). These probability density results give us information on the variation range of each parameters and the probability to obtain each value, which will be useful for the computation of the GUM, MCM and BST approaches. For the estimation of the building solar radiation equation please refer to the experiments in reference [12].

On the other hand, for the dynamical parameters involved such as the weath24 hourser, the occupancy and the indoor temperature, measurements data are available during the selected period. Still, the user needs to find the adequate probability distribution of each parameter and the uncertainty associated to the measured data.

After having performed several distributions fittings to the measured data, we assumed that each dynamic variable X , follows a normal distribution with its associated distribution coefficients. The related uncertainty $u(X)_k$, is obtained by adding the uncertainty of the measurement sensors $u(X)_s$ and those arising from the data itself [8]. Then, in order to take into account the variability of the data, an uncertainty $u(X)_k$ is assigned to the variable at each time step k (Eq. (18)).

$$u(X)_k^2 = u(X_k)^2 + u(X)_s^2. \quad (18)$$

For example, in Figure 4 is represented the temperature histogram in the six offices of the monitored zone and the circulation area. The represented temperature is the sum on 24 hours of the difference between the outdoor and indoor temperature (Eq. (10)). Then, we assumed that the temperature data follows a normal distribution in all the offices and the same procedure is used for the solar radiations and the wind speed data.

4.2 Propagation of uncertainty in the model

The purpose of this research work is to estimate the uncertainty associated to the daily energy consumption of a

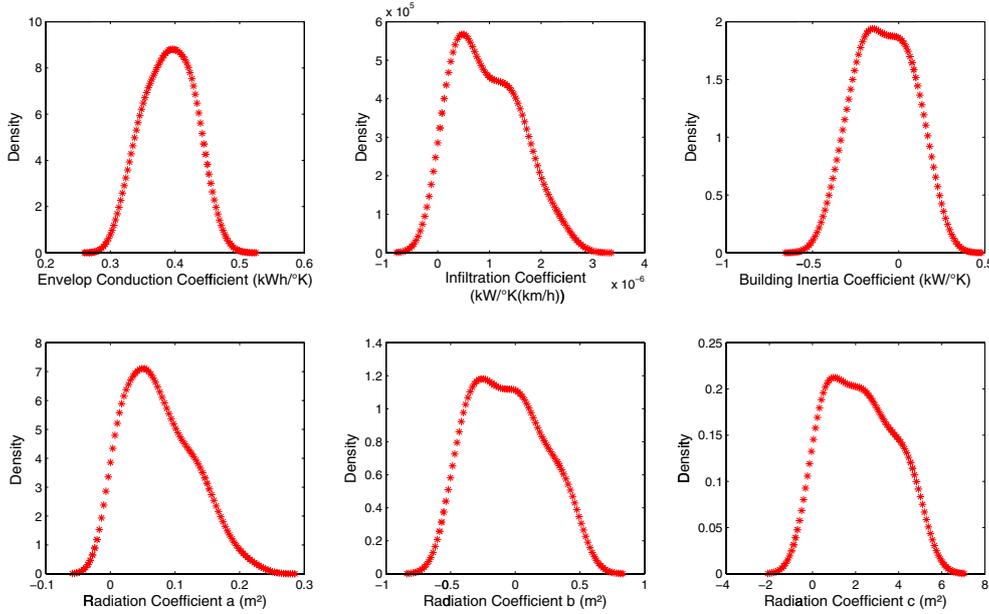


Fig. 3. Static parameters density plots obtained from the 10 000 samples realised.

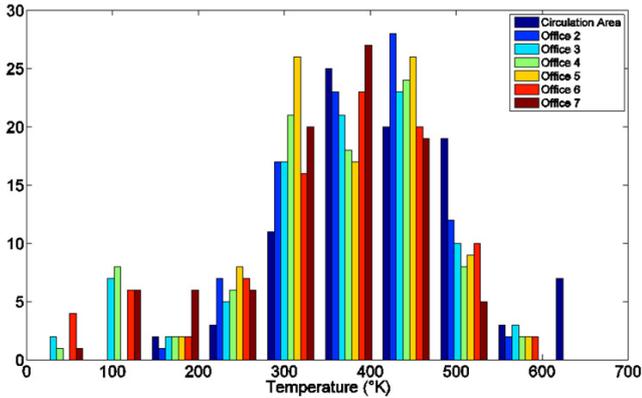


Fig. 4. Sum on a day basis of the difference between the outdoor and indoor temperature distribution ($^{\circ}\text{K}$) in each part of the monitored zone.

building estimated by a given model. The uncertainty will be estimated by the three different methods cited above in order to find the better or the more suitable one for the estimation of a daily energy consumption.

As stated above, the BEECHAM model dynamic parameters are:

- the difference between the indoor and outdoor temperature over a day Ψ ;
- the sum of the solar radiations over a day Φ ;
- the wind speed v found by taking the average on a day basis also.

Knowing that the weather conditions and the occupancy within the building energy demand model are changing at each iteration, all the uncertainty analysis have been done with the assumption that their daily variability is equal to zero which means that the weather conditions and the

occupancy rate for example are the same from one day to another.

Then, to get the uncertainty associated to these three dynamic parameters, we combined the uncertainty of the sensor being used (depending on the measured parameter) and the uncertainty coming from the data itself (depending on the selected method) like stated in equation (16). If it is the GUM method, the user will have an amount of uncertainty associated to its data; if it is the Monte Carlo method, the uncertainty will be on the form of a probability density function and if it is the Bayesian Analysis, then according to the prior information available on the parameter under consideration and the data recorded the user will have a posterior distribution to perform the uncertainty analysis.

Then, for the GUM method and referring to equation (2), we assumed that the input quantities are not correlated and we estimated thanks to the BEECHAM model, the value $\frac{\partial Q_{building}}{\partial X}$ of each input quantity X and we obtained the uncertainty associated to each variable $u(X)$ thanks to equation (18). The results are presented in Tables 1 and 2.

To give a better interpretation to the uncertainty obtained, the GUM [8] proposed the calculation of a confidence interval in which the best estimate of the measurand will fall in. A lower and upper bounds are estimated with a coverage factor α like stated in the following equation.

$$X - \varphi * u(X)_k \leq X \leq X + \varphi * u(X)_k. \quad (19)$$

With the 60 days data sets, we estimated the probability to obtain the energy consumption of the monitored zone with its upper and lower bounds within a confidence level of 95% ($\varphi = 2$).

The estimation of the uncertainty with the Monte Carlo Method was quite easier after knowing the distribution functions of all the input quantities. We performed

Table 2. Estimation of the dynamic parameters distributions with measured data.

Dynamic parameters	Distributions
Outdoor and indoor temperature Difference Ψ ($^{\circ}\text{K}$)	$N(386.76, 59.04)$
Solar Radiations over a day R_{ay} (W/m^2)	$N(1.19 \times 10^{+03}, 695.49)$
Wind Speed v (km/h)	$N(10.84, 6.18)$
Occupancy rate A_o (Wh)	$N(94.4, 47.7)$

Table 3. The daily energy consumptions and its uncertainty estimated by the three approaches and compared to the measured energy consumption in the monitored zone.

Method	Estimated Value (kWh)	Associated Uncertainty (kWh)	Proportion of Uncertainty (%)
GUM Principle	1.493×10^2	29.56	19.79
Monte Carlo Method	1.491×10^2	29.91	20.06
Bayesian Uncertainty Quantification	1.567×10^2	24.99	15.95
Measured Energy Consumption	1.363×10^2	32.97	24.18

a number of simulations of 10^6 in order to have a large dataset for the output estimation. We generated 10^6 data for each input quantity according their distribution functions and the data obtained from the generation are spread through the BEECHAM model like stated in equation (3). The energy consumption result obtained on the 60 days basis is then compared to the GUM results in Figures 6 and 7.

However, for the Bayesian analysis the process was quite different from the last two methods. We have to define the prior knowledge for each parameter and the likelihood functions. Having performed measurements on the dynamic parameters, the likelihood functions was known as well as their probability density functions and the prior was made of the weather data used when performing a building simulation. In TRNsys 17 [25] for example, to perform a building energy simulation, TMY (Typical Meteorological Year) files are opened for users and represent a typical weather data. The files are built from the weather data measured during the last 20 years and are available in several cities. Then, we chose two months measurement files to make the weather parameters (temperature, solar radiations and wind speed) prior functions (Fig. 5).

Nevertheless, for the static parameters, we decided to use normal distributions to build the prior knowledge of each parameter and the dataset built previously for the GUM and Monte Carlo Method for the likelihood functions. The uncertainty associated to the building zone estimated by the Bayesian method is obtained by the propagation of each posterior distribution through the BEECHAM model and the results are compared to the previous methods (Figs. 7 and Tab. 3).

5 Discussion and conclusion

In order to assess the BEECHAM model consistency, we decided to compare the two months real energy consumption to the energy consumption estimated by the model for the monitored zone presented before and during the same period. The heat load of the building zone was measured during January and February providing a 60 days

dataset. Table 3 shows the results comparison. One can notice that the measured consumption is much closed to the consumption estimated by the BEECHAM model but the real daily energy consumption is overestimated by an average of 27 kWh in the whole period. This difference can be explained by the fact that a building energy consumption is badly estimated and sometimes lead to a gap between the estimated energy performance (by simulation or by calculation), of a building and the real one. De Wilde [4] pointed out this issue and stated that it appears mainly at three stages of the building life: the design stage, the construction stage and the operational stage. The real energy consumption can go from two to three times the predicted energy, which is simply huge to be acceptable.

However, the amount of uncertainties arising from the data analysis can be explained in different ways. The first issue is the presence of uncertainties in the model itself. A building model, either a simulation or an analytical model, aims to represent the building in real conditions and as true as possible. The use of improper methods or incorrect input data can lead to increase the amount of uncertainty. However, even if the model is supposed to be correct, the user needs to have the proper qualifications and knowledge to apply it.

Uncertainties could also come from the measurement procedures and the devices used for the data acquisition. Sometimes, these instruments are the root causes of data deviation and the operator judgement and assumptions can have a huge influence on the results in different situations. For example, in the Bayesian analysis case, the user sometimes has to make hypotheses in order to have a consistent prior knowledge on a quantity, which will influence the results on the posterior distributions.

Finally, the last cause could be the variability in the dynamical data. It is known that the outdoor temperature at a moment t for example, depends on the temperature at $t - 1$ and will be different from the temperature at time $t + 1$. This variability in time of the dynamical parameters is fairly unpredictable and cannot be quantify on a regular basis. This dynamical feature is a non-negligible source of uncertainty and should be taken into account

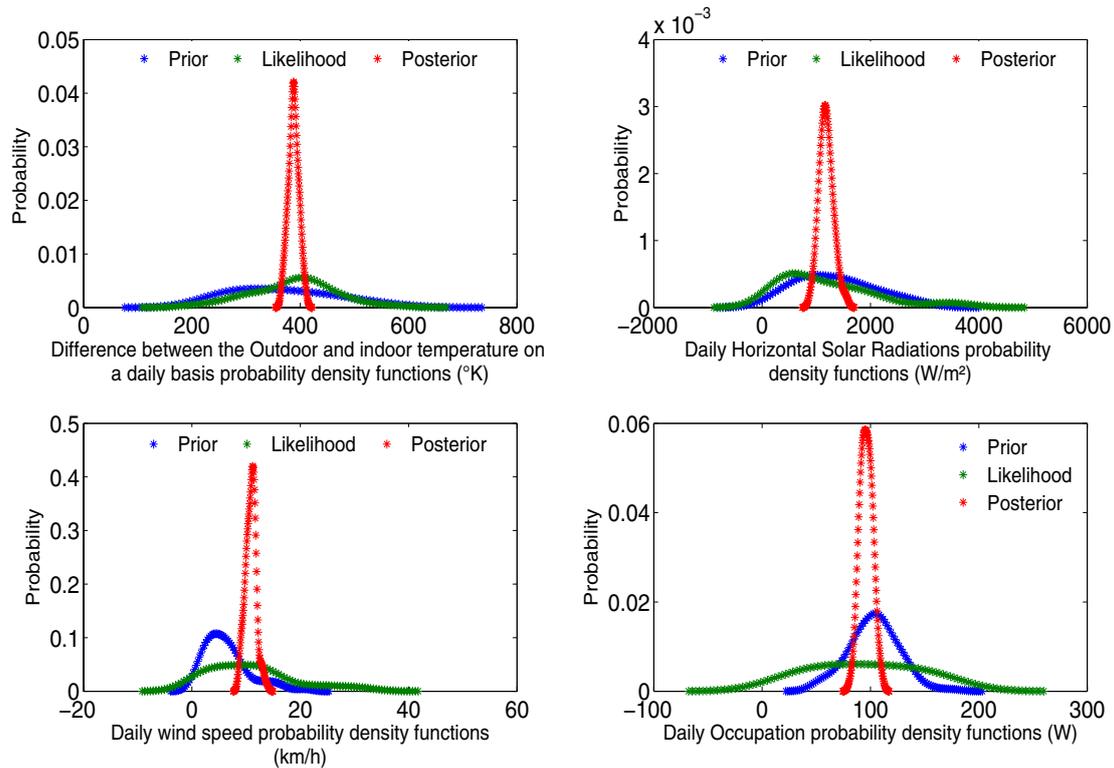


Fig. 5. Daily probability density plots (prior, likelihood and posterior) for the dynamic input parameters.

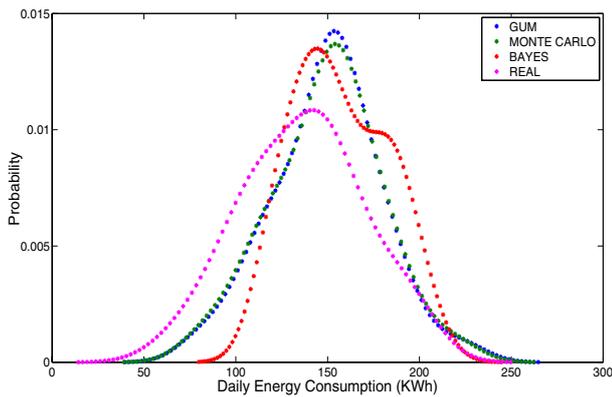


Fig. 6. Daily energy consumption probability plots obtained from the gum, the Monte Carlo method, the bayesian principle and the measured consumption.

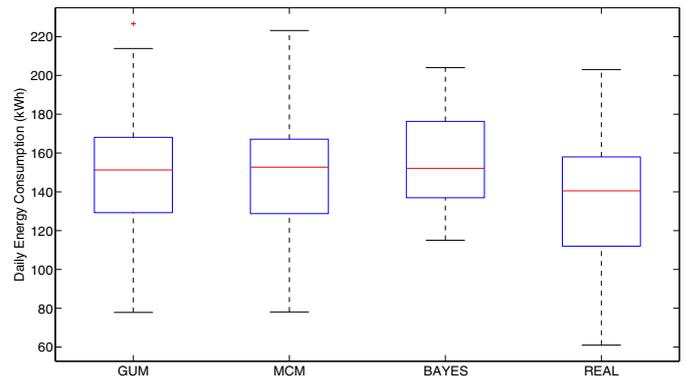


Fig. 7. Dispersion of the daily energy consumption data estimated by the three approaches and compared to the measured data.

when quantifying the uncertainty related to the dynamic parameters.

The main goal of this research work was to compare three approaches of uncertainty quantification: the Guide to the expression of Uncertainty in Measurement, the Monte Carlo Method and the Bayesian Uncertainty Quantification with a simplified model. These three approaches allow quantifying the uncertainty associated to the energy consumption of a given building with low variation among them. They give almost the same results and have affirmed the research which have already been made in metrology [19,26]. Each method has certainly its own limitations

and its advantages, thus, the user will have to choose according to the data and the information he already has, the convenient one. This paper has also pointed out by its results, the existence of a gap between the predicted and the actual energy consumption in a given building and has highlighted the causes of such discrepancies. Moreover, this analysis was done with the assumption that all the input parameters are uncorrelated. A logical continuation of this research work will be to model the relationship between them and study its influence on the uncertainty analysis. Then, the analysis can be extended to the optimization of the measurement plan.

Acknowledgements. This study was funded by the grants from the National Association of Research (ANR) in France within the OMEGA project and the data was provided by the Cerema.

References

1. IEA, "Key World Energy Statistics 2014 – Key World 2014.pdf," 2014. [Online]. Available: <http://www.iea.org/publications/freepublications/publication/KeyWorld2014.pdf>. [Accessed: 10-Dec-2014].
2. L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, *Energy Build.* **40**, 394–398 (2008)
3. ASHRAE, ASHRAE Guideline 14, Measurement of energy and demand savings (2002)
4. P. de Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, *Autom. Constr.* **41**, 40–49 (2014)
5. A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, *Appl. Energy* **97**, 355–364 (2012)
6. S. De Wit, G. Augenbroe, Analysis of uncertainty in building design evaluations and its implications, *Energy Build.* **34**, 951–958 (2002)
7. S. Attia, M. Hamdy, W. O'Brien, S. Carlucci, Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design, *Energy Build.* **60**, 110–124 (2013)
8. GUM, Guide to the Expression of Uncertainty in Measurement (1995)
9. O. Bodnar, G. Wübbeler, C. Elster, On the application of Supplement 1 to the GUM to non-linear problems, *Metrologia* **48**, 333 (2011)
10. K. Weise, W. Woger, A Bayesian theory of measurement uncertainty, *Meas. Sci. Technol.* **4**, 1 (1993)
11. I. Lira, G. Kyriazis, Bayesian inference from measurement information, *Metrologia* **36**, 163 (1999)
12. A. Caucheteux, A Building Energy Efficiency Characterisation Method (BEECHAM) to Assess Existing Buildings Performance (2012)
13. C. Elster, W. Wöger, M.G. Cox, Draft GUM Supplement 1 and bayesian analysis, *Metrologia* **44**, L31 (2007)
14. M. Ángeles Herrador, A.G. González, Evaluation of measurement uncertainty in analytical assays by means of Monte-Carlo simulation, *Talanta* **64**, 415–422 (2004)
15. M.G. Cox, B.R. Siebert, The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty, *Metrologia* **43**, S178 (2006)
16. C. Elster, B. Toman, Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the Supplement 1 to the Guide: a comparison, *Metrologia* **46**, 261 (2009)
17. I. Lira, D. Grientschnig, Bayesian assessment of uncertainty in metrology: a tutorial, *Metrologia* **47**, R1 (2010)
18. I. Lira, W. Wöger, Bayesian evaluation of the standard uncertainty and coverage probability in a simple measurement model, *Meas. Sci. Technol.* **12**, 1172 (2001)
19. R. Kacker, A. Jones, On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent, *Metrologia* **40**, 235 (2003)
20. S.A. Kalogirou, C.C. Neocleous, C.N. Schizas, Building heating load estimation using artificial neural networks, in *Proceedings of the 17th international conference on Parallel architectures and compilation techniques*, (1997), Vol. 8, p. 14
21. T. Olofsson, S. Andersson, Long-term energy demand predictions based on short-term measured data, *Energy Build.* **33**, 85–91 (2001)
22. Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, *Appl. Energy* **86**, 2249–2256 (2009)
23. D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, *Renew. Sustain. Energy Rev.* **37**, 123–141 (2014)
24. AFNOR, Norme Européenne, Norme Française, Performance Energetique des batiments, Calcul des besoins d'énergie pour le chauffage, Batiments résidentiels. Norme AFNOR, Aout-1999.
25. A. Trnsys, Transient System Simulation Program, University. Wis. (2000)
26. C. Elster, B. Toman, Bayesian uncertainty analysis for a regression model versus application of GUM Supplement 1 to the least-squares estimate, *Metrologia* **48**, 233 (2011)