Parameter estimation and model selection for water sorption in a wood fibre material

Julien Berger1 · Thibaut Colinart2 · Bruna R. Loiola3 · Helcio R. B. Orlande4

Received: 21 February 2020 / Published online: 19 September 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
The sorption curve is an essential feature for the modelling of heat and mass transfer in porous building materials. Several models have been proposed in the literature to represent the amount of moisture content in the material according to the water activity (or capillary pressure) level. These models are based on analytical expressions and few parameters that need to be estimated by inverse analysis. This article investigates the reliability of eight models through the accuracy of the estimated parameters. For this, experimental data for a wood fibre material are generated with special attention to the stop criterion to capture long time kinetic constants. Among five sets of measurements, the best estimate is computed. The reliability of the models is then discussed. After proving the theoretical identifiability of the unknown parameters for each model, the primary identifiability is analysed. It evaluates whether the parameters influence on the model output is sufficient to proceed the parameter estimation with accuracy. For this, a continuous derivative-based approach is adopted. Seven models have a low primary identifiability for at least one parameter. Indeed, when estimating the unknown parameters using the experimental observations, the parameters with low primary identifiability exhibit large uncertainties. Finally, an Approximation Bayesian Computation algorithm is used to simultaneously select the best model and estimate the parameters that best represent the experimental data. The GAB and Fredlund-Xing models, together with a proposed model in this work, were the best ones selected by this algorithm.

List of symbols

Physical parameters: Latin letters
\( a, \hat{a} \) Water activity (–)
\( K \) Slope of the sorption model (–)

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00226-020-01206-0) contains supplementary material, which is available to authorized users.

✉ Julien Berger
julien.berger@univ-lr.fr

Extended author information available on the last page of the article
\( m, m_0 \) Mass (kg)

\( p \) Parameter of a model (–)

\( R_i \) Water vapour gas constant (J kg\(^{-1}\) K\(^{-1}\))

\( T \) Temperature (K)

\( u, u_0, \hat{u}, u_m \) Moisture content dry basis (–)

**Physical parameters: Greek letters**

\( \delta \) Measurement uncertainty (–)

\( \delta_c \) Random component of measurement uncertainty (–)

\( \delta_s \) Systematic component of measurement uncertainty (–)

\( \Psi \) Capillary pressure (Pa)

\( \rho \) Material dry density (kg m\(^{-3}\))

\( \rho_2 \) Liquid water specific mass (kg m\(^{-3}\))

**Mathematical notations: Latin letters**

\( d \) Distance function \( i \)

\( \mathfrak{D}_i \) Polynomial of order \( i \)

\( f \) Sorption model function

\( F \) Fisher matrix

\( K \) Kernel perturbation

\( N \) Number of parameters of a model

\( N_a \) Number of measurement points for a population

\( N_e \) Number of carried out measurements

\( N_i \) Number of populations

\( N_v \) Number of particles in each population

\( \mathfrak{R}_i \) Polynomial of order \( i \)

\( \mathfrak{S}_N \) Polynomial of order \( N \)

\( \mathcal{U} \) Uniform distribution

\( w \) Weight

**Mathematical notations: Greek letters**

\( \alpha_i \) Polynomial coefficients

\( \beta_i \) Polynomial coefficients

\( \varepsilon, \varepsilon_i \) Tolerance

\( \eta_n \) Relative error estimator

\( \gamma_n, \gamma^T \) Derivative-based sensitivity metric related to parameter \( p_n \)

\( \Omega \) Set of elements

\( \pi \) Probability distribution

\( \nu_n, \nu^T \) Local derivative-based sensitivity metric related to parameter \( p_n \)

\( \kappa, \kappa_0 \) Kernel parameter

\( \tau \) Acceptation rate

\( \theta_n \) Sensitivity function related to parameter \( p_n \)

**Subscripts and superscripts**

\( a \) Related to water activity

\( \text{apr} \) A priori parameter

\( \copyright \) Springer
Introduction

Within the environmental context, bio-based materials such as wood fibre have been increasingly used in building constructions due to their reduced ecological footprint and their thermal performance. The properties of such wood-based materials have a good reproducibility due to their industrial production. Several recent studies pointed out the importance of modelling accurately the phenomena of adsorption in such materials to predict the phenomena of heat and mass transfer and assess the moisture disorder risks (Berger et al. 2015). In the Conclusion section of Patera et al. (2016), the error in the prediction can reach 20–30% of moisture content in the wood without modelling accurately the adsorption phenomena combined with hysteresis effects. Those results are confirmed by Zhang et al. (2016a) for a wood fibre material where the authors highlight the importance of modelling accurately adsorption phenomena to evaluate the risk of mould growth. Furthermore, the moisture content is a crucial parameter in the modelling framework since several other properties such as thermal conductivity depend on it (Willems 2014).

As presented in Skaar (1988), several phenomenological models have been proposed in the literature to reproduce the moisture sorption curve. Among them, one can mention the Guggenheim–Anderson–de Boer (GAB) model with examples of applications in Singh and Singh (1996) and Iglesias and Chirife (1995). The Brunauer, Emmett and Teller model is employed in Colinart and Glouannec (2017) for a hygroscopic material. In Carmeliet and Roels (2002), the van Genuchten model is used for several building materials. All these models are based on a few important parameters that can be estimated using experimental data of moisture content according to the water activity or capillary pressure. These experimental observations can be obtained using static gravimetric methods. Then, the unknown parameters can be inferred by solving an inverse problem. For instance, the parameters of several models are retrieved in Iglesias and Chirife (1995), Ouertani et al. (2014) or Stolarska and Garbalinska (2017). Those studies use least square estimator algorithm, and the discussion among the models is based only on the residual between numerical predictions and experimental data. In Furmaniak (2012) and Karoglou et al. (2005), investigations are carried out considering one model. The
parameters are determined for several materials without information on the accuracy of the estimation.

It is essential to have reliable models, which accurately represent the physical phenomena when compared to experimental observations. Such evaluation has already been done for wood materials in Zhang et al. (2015). Nevertheless, the reliability is also based on the capacity of estimating the important parameters with accuracy. This accurate estimation is also required to verify the theoretical models based on first principles of physics. This paper proposes to investigate the reliability of eight models, seven being mainly used for heat and mass transfer modelling in building materials. The last one is proposed according to the general curve of moisture sorption. To discuss the reliability, experimental measurements in a wood fibre material are taken using a DVS equipment. Then, using the experimental observations, the parameter estimation problem can be solved for each model. The accuracy of the retrieved parameters and the resultant robustness of the models are discussed using two important approaches. First, the primary identifiability of the parameters is evaluated. It enables to evaluate the sensitivity of the output models to each unknown parameter, using a continuous derivative-based approach. If a model is not sensitive to a parameter, it indicates that the latter cannot be retrieved with accuracy. Then, an efficient Approximate Bayesian Computation (ABC) algorithm is employed to conduct a selection over the eight competing models. The selection is achieved sequentially with decreasing tolerances. The selected model is the one having the highest probability to minimize the distance between predictions and experimental observations for the smallest tolerance.

The article is organized as follows: “Physical model” section presents the eight physical models to predict the moisture content according to the water activity. Moreover, the methodology is described to evaluate the robustness of the models in the framework of a parameter estimation problem. “Experiments” section introduces the experimental measurements obtained for a wood fibre material. Then, “Reliability of the models” section discusses the reliability of the models and “Conclusion” section gives some general remarks on the results of the study.

**Physical model**

**Models for water adsorption**

The sorption model of water in porous material describes the water content $u(-)$ contained in the porous matrix for a defined water activity $a(-)$. The moisture content $u$ is defined as

$$u \overset{\text{def}}{=} \frac{m - m_0}{m_0},$$

where $m_0$ (kg) is the dry mass of the material and $m$ (kg) the mass of the material. In the literature, several models are proposed to represent the dependency of $u$ on
the water activity $a$. From a mathematical point of view, the sorption curve can be formulated as:

$$
  u : \bigcup_{n=1}^{N} \Omega_{p_n} \times \Omega_a \longrightarrow \mathbb{R}_{>0},
  \quad (p_1, \ldots, p_N, a) \mapsto f(p_1, \ldots, p_N, a),
$$

where $a$ is the water activity, $f(-)$ is the sorption model and $p_n$ is an unknown parameter involved in the model definition. Depending on the model, the parameter $p_n$ may have a physical meaning and specific unit. As presented below, the total number of parameter $N$ varies from two to four depending on the sorption model. The set of the parameter $p_n$ verifies $\mathcal{P}_{p_n} \subset \mathbb{R}$. The water activity $a$ belongs to the interval $[0, 1]$. However, $a$ is restricted to $[0.05, 0.95]$ for several reasons. First, some models proposed in the literature are mathematically not defined for $a = 0$. Then, in this study, it is stated that the moisture content at saturation $a = 1$ is unknown. Indeed, for this value, moisture content may cover several definitions like hygroscopic, capillary or complete saturation moisture content (Nilsson 2018). Moreover, the discussion is carried out with the perspective of using sorption models for the simulation of heat and mass transfer in porous building materials under normal conditions. Thus, the fully dry and saturated states are never reached in practice.

A total of eight models are investigated, namely the Brunauer, Emmett and Teller (BET), the normalized Guggenheim–Anderson–de Boer (GAB), the thermodynamic (TRM), the empirical Oswin (OSW), the Fredlund-Xing (FX), the van Genuchten (VG), the Smith (SM) and an additional moisture adsorption (MADS) model which is proposed based on the general shape of the sorption curve. A detailed presentation of each model is given in Appendix A in Electronic Supplementary Material. The indicator $m \in \{1, \ldots, 8\}$ is linked to each model $u_m \equiv f_m$.

**Parameter estimation problem**

The eight models depend on several parameters that can be determined using experimental observations of moisture content $u$ according to the water activity $a$. The procedure to solve the parameter estimation problem is now detailed. First, some important notations are clarified. It is assumed that measurement, denoted as $\hat{u}$, of moisture content in a material is obtained for different levels of water activity $\hat{a}$. The set of model indicator is:

$$
  \Omega_m \overset{\text{def}}{=} \{m|m \in \{1, \ldots, 8\}\},
$$

where $m$ is an indicator representing one of the eight investigated models. We denote as $u_m := f_m(p, \hat{a})$ the computed moisture content with the model $m$ for the water activity $\hat{a}$ of the experiments (with its respective domain $\Omega_{\hat{a}}$). In addition, for model $m$, the set of unknown parameters is defined as:
Several distinctions are made among the unknown parameters. First, the solution of the parameter estimation problem is denoted by $\mathbf{p}^\circ$. Then, the so-called a priori parameters, which are used in the preliminary identifiability investigations, are written as $\mathbf{p}^{apr}$.

**Primary identifiability**

The primary identifiability of the unknown parameters is discussed according to Jumabekova et al. (2019). It aims at carrying out a sensitivity analysis of the unknown parameters on the output model using a continuous derivative-based approach. For this, the sensitivity function $\theta_n$ of the model $f_m$ relatively to the parameter $p_n$ is defined as (Saltelli et al. 2004):

$$\theta_n : \bigcup_{n=1}^{N} \Omega_{p_n} \times \Omega_a \rightarrow \mathbb{R},$$

$$(p_1, \ldots, p_N, a) \mapsto \frac{\partial f_m}{\partial p_n}, \quad \forall n \in \{1, \ldots, N\}.$$

The differentiation of sorption model relatively to the parameters is performed analytically. Then, the following sensitivity metrics are defined (Sobol 1990; Dickinson and Gelinas 1976; Sobol and Kucherenko 2009; Kucherenko and Song 2016):

$$\nu_n : \bigcup_{n=1}^{N} \Omega_{p_n} \times \Omega_a \rightarrow \mathbb{R},$$

$$(p_1, \ldots, p_N, a) \mapsto \int_{\Omega_{p_n}} \theta_n^2 \, dp_n, \quad \forall n \in \{1, \ldots, N\}.$$

and

$$\nu_n^1 : \bigcup_{n=1}^{N} \Omega_{p_n} \times \Omega_a \rightarrow \mathbb{R},$$

$$(p_1, \ldots, p_N, a) \mapsto \int_{\Omega_a} \int_{\Omega_p} \theta_n^2 \, da \, dp, \quad \forall n \in \{1, \ldots, N\}.$$

The quantities $\nu_n$ and $\nu_n^1$ translate how changes of parameter $p_n$ impact the sorption model $f_m$. The first one is local and depends on the value of the water activity $a$, while the second is global over the whole range of $a$. A large value of those metrics reveals an important influence of the parameter. For the analysis, it is transformed into a dimensionless metric to get the derivative-based sensitivity indexes $\gamma_n$ and $\gamma_n^1$:
Both metrics \( v_n \) and \( \gamma_n \) assess the sensitivity of the parameter over its whole domain of variation \( \Omega_{p_n} \). The metrics \( \gamma_n \) and \( \gamma_n^\top \) are local and global according to the water activity, respectively.

With the computation of the sensitivity functions, the so-called Fisher information matrix \( F \) (Karalashvili et al. 2015; Ucinski 2004) is defined as:

\[
F : \bigcup_{n=1}^{N} \Omega_{p_n} \times \Omega_a \longrightarrow \mathcal{M}(\mathbb{R}^{N \times N}),
\]

\[
(p_1, \ldots, p_N, a) \mapsto \left[ F_{n_1n_2} \right], \quad \forall (n_1, n_2) \in \{1, \ldots, N\}^2,
\]

\[
F_{n_1n_2} \overset{\text{def}}{=} \int_{\Omega_a} \theta_{n_1} \cdot \theta_{n_2} \, da.
\]

The matrix relates the total sensitivity of the system. From its computation using the estimated parameters \( p_\circ \), a relative error estimator \( \eta_n \) can be obtained for the parameter retrieved \( p_\circ_n \) (Walter and Lecourtier 1982; Walter and Pronzato 1990):

\[
\eta_n \overset{\text{def}}{=} \sqrt{(F^{-1})_{nn} / p_\circ_n^2}, \quad \forall n \in \{1, \ldots, N\}.
\]

High values of \( \eta_n \) mean a high error during the estimation process.

**Solving the inverse problem combined with model selection**

The Approximate Bayesian Computation (ABC) algorithm is an efficient tool to infer the posterior distributions when the likelihood function is computationally too expensive to evaluate. It has been successfully implemented in various fields of research (Toni et al. 2009; Liepe et al. 2014; da Costa et al. 2018; Loiola et al. 2020). With the ABC technique, the prior information about the parameters is taken into account. In this way, it limits the variability of the estimated parameters in the inverse analysis. Within the Bayesian framework, the objective is to approximate the posterior distribution \( \pi(p | \hat{u}) \) using Bayes’ theorem:
\[
\pi(p|\hat{u}) = \frac{\pi(\hat{u}|p) \cdot \pi(p)}{\pi(\hat{u})},
\]

where \(\pi(p)\) is the a priori density of the parameters \(p^{\text{apr}}\), \(\pi(\hat{u}|p)\) is the likelihood function and \(\pi(\hat{u})\) is the marginal probability density of measurements. The ABC algorithm can also be employed for model selection. An indicator \(m\) is defined for each model of interest. Each model has a prior density \(\pi(m)\). The marginal posterior distribution is approximated among all models and parameter subspaces, such as \(\pi(m,p|d(\hat{u},u_m) \leq \varepsilon)\). So, it is useful to rank the models. Here, to increase the acceptance rate, the ABC method based on sequential Monte Carlo (SMC) sampling is used. The complete description of the algorithm is provided in Appendix B in Electronic Supplementary Material.

Figure 1 summarizes the ABC algorithm used in this work. It proceeds as follows. At the beginning, the different models have uniform a priori distributions, so that they have exactly the same probability of being selected. Uniform priors are also assigned to the parameters of each model. The algorithm runs for successive \(N_{\varepsilon}\) populations with decreasing tolerances \(\varepsilon\). For the population \(i\), i.e. the tolerance \(\varepsilon_i\), the algorithm is divided into three main steps. (1) First, sample a model \(m^{**}\) from the prior distribution among the eight competing models. If it is the first
population, uniform a priori distributions are defined for the competing models and the unknown parameters. It enables to not favour any of the competing models. For the other populations, the priors for the parameters are obtained from the weights at the previous population. (2) Then, for the candidate model sampled from its prior, the algorithm samples a candidate parameter $p^{**}$, using a sequential scheme and a small perturbation to ensure that the whole parameter space is explored. (3) For each candidate parameter, the distance $d_m$, i.e. the error, between the direct model and the experiment is computed. Two cases are distinguished. If the distance is lower than the tolerance $\epsilon_s$, then both the parameter and model are selected. It is stated that the particle validated the distance test. If the test is not valid, then the algorithm comes back to the first step in order to sample a new model from its prior. This operation is repeated until $N$ particles have been accepted, denoting by $\tau$ the acceptance rate. The tolerance of the last population corresponds to the desired agreement between the model and measurement data. Among the successive populations, less models validate the distance test since the tolerances are decreasing. At the end, only the best model(s) remain(s) and the samples for the parameters approximate their posterior distribution. In other words, the algorithm provides the best model and corresponding parameters values that minimize the error between measured and estimated quantities.

**Experiments**

**Materials and methods**

The tested material is wood fibre insulation (Soprema 2019) with density of $\rho = 50$ kg m$^{-3}$. Due to its low thermal conductivity and vapour permeability, this bio-based material is increasingly used in building envelope. Furthermore, it presents a strong hygroscopic behaviour as highlighted by the previous sorption isotherm measurement (Vololonirina et al. 2014).

Sorption isotherms are measured with five samples with the DVS equipment IGASorp-HT system (Hiden Isochema, Warrington, UK). The instrument has a microbalance with a resolution of $0.1 \mu g$ on which a stainless-steel mesh basket containing the sample is suspended. The sample is then placed inside a separate chamber with controlled temperature and water activity, and sample mass is continuously recorded. Prior to the start of the adsorption measurement, the sample is dried under flow of dry nitrogen at $65^\circ$C for 24 h (with a flow of 250 mL min$^{-1}$) until stabilization (ISO12570 2000). Sample dry mass is recorded after setting the temperature to $23^\circ$C under flow of dry nitrogen. Here, sample dry mass varies between 10.8 and 28.2 mg. Then, the sample is exposed to increasing humidity from 0.05 to 0.9, with a 0.05 step, the testing temperature being $23^\circ$C. The humidity is controlled by mixing dry and water vapour-saturated nitrogen streams at a total flow of 250 mL min$^{-1}$ using electronic mass flow controllers. It is measured by a sensor placed in the chamber near the sample. The experiment is run at given temperature and water activity until a user-defined stop criterion is reached. Usual stop criteria are (1) hold time, (2) final rate of derivative of mass with respect to time $\frac{dm}{dt}$ or (3) accuracy of asymptotic moisture content from a kinetic model.
fit to the moisture content versus time data. While several stop criteria may be found in the literature for cellulosic materials, Glass et al. (2017) underlined that the commonly used stop criteria may mischaracterize equilibrium moisture content up to 1% of the moisture content. Therefore, the recommendation is to increase hold times to catch long time constants of sorption kinetic (the order of 500 min or longer), even if it increases the isotherm measurement time to several weeks for a single replicate of a single material. Later, the same authors Glass et al. (2018) proposed a new methodology to improve measurement accuracy and to reduce measurement times. Based on these previous works, a slope $\frac{dm}{dt} = 10 \mu g g^{-1} min^{-1}$ calculated over a 15 min window combined with a maximum hold time of 24 h was used as stop criterion. When this condition is met, the apparent equilibrium moisture content is taken as the last measured moisture content.

**Experimental data**

The experimental results for the five samples are presented in Fig. 2a. Except for three points with water activity higher than 0.9, all equilibrium moisture contents are obtained by meeting the stop criterion. The measurement time increases with water activity, ranging between 2 and 12 h. In the hygroscopic range (i.e. $a < 0.8$), acquisition could be stopped because of high signal-to-noise ratio. Nevertheless, the calculated slope over a 60 min window did not exceed $15 \mu g g^{-1} min^{-1}$. Therefore, we have good confidence in the results in the hygroscopic range and the discrepancy is limited. For water activity higher than 0.8, the signal-to-noise ratio is better because of the larger mass change. Nevertheless, even if each measurement lasts for at least 4 h, it might not be sufficient for the identification of long time constants of sorption kinetic. For instance, the calculated slope over a 120 min window did not drop below $45 \mu g g^{-1} min^{-1}$, which is much higher than the value of $3 \mu g g^{-1} min^{-1}$ suggested by Glass et al. (2018). Therefore, we expect measuring moisture content with an accuracy of at least 0.36% (Glass et al. 2018). Nevertheless, the sorption at high water activity involves a complex phenomenon (like polymer softening) that may vary from one sample to another. This variability is highlighted by the higher discrepancy between the sets for water activity higher than 0.8.

According to Taylor (1997), the best estimates of the moisture content are:

$$\hat{u} \overset{\text{def}}{=} \frac{1}{N_e} \sum_{i=1}^{N_e} u_i,$$

where $N_e$ is the number of carried out measurements. In the present case, $N_e = 5$ since measurement for five samples has been taken. To evaluate the total measurement uncertainty $\delta$, both the random and the systematic components of the uncertainty are considered:

$$\delta \overset{\text{def}}{=} \sqrt{\delta_{\infty}^2 + \delta_{\Sigma}^2}.$$

The random part $\delta_{\infty}$ is computed through the standard deviation of the mean:
The systematic component $\delta_\Sigma$ is due to the experimental DVS device and given as follows:

$$\delta_\Sigma \overset{\text{def}}{=} \frac{1}{\sqrt{N_e}} \sqrt{\frac{1}{N_e} \sum_{i=1}^{N_e} (u_i - \hat{u})^2}.$$ 

The systematic component $\delta_\Sigma$ is due to the experimental DVS device and given as follows:

$$\delta_\Sigma = 10^{-7}.$$ 

Here, it is given by the balance resolution divided by the dry mass. It is assumed that there is no other systematic error in the measurement design. The best estimate for the moisture content is given in Fig. 2c as well as in Table 1. The different uncertainty components are also indicated. The variation of the relative uncertainty with the water activity is shown in Fig. 2b. It can be noticed that the random uncertainty
component is significant compared to the systemic one. In Fig. 2c, it seems that the uncertainty increases at high water activity. However, from a relative point of view, the uncertainty is high for $a \geq 0.8$, of the order of 5%. As samples have different dry masses and shapes, it may explain the observed higher uncertainty. In Fig. 2c, the measured sorption isotherm is compared to previous results (Vololonirina et al. 2014) from the literature. The shape in the hygroscopic domain is similar. For higher water activity, large difference may be observed, probably due to differences in stop criteria. Indeed, if the measurement is stopped too early, it may lead to an underestimation of equilibrium moisture content during adsorption (Glass et al. 2017).

**A priori distribution of unknown parameters**

The slope of the sorption model MADS in supplementary material (Eq. (9)) is $K = 0.2416$. Uniform distributions are considered for the prior density of the unknown parameters. The interval of variation of each parameter is given in Table 2 for each model. To challenge each of the competing models, the interval of variation of the prior density is defined to represent a large range of sorption curves. The interval of variation of the a priori parameters is chosen so that the image of each

| Water activity | Moisture content $a$ | Total uncertainty $\delta$ | Random uncertainty $\delta_u$ |
|----------------|---------------------|--------------------------|-----------------------------|
| 0              | 0                   | 0.01                      | 0                           |
| 0.05           | 0.0121              | 0.0006                    | 0.0005                      |
| 0.1            | 0.0208              | 0.0009                    | 0.0008                      |
| 0.15           | 0.0276              | 0.0012                    | 0.0011                      |
| 0.2            | 0.0344              | 0.0014                    | 0.0014                      |
| 0.25           | 0.0402              | 0.0017                    | 0.0016                      |
| 0.3            | 0.0468              | 0.0019                    | 0.0019                      |
| 0.35           | 0.0527              | 0.0022                    | 0.0021                      |
| 0.4            | 0.059               | 0.0024                    | 0.0024                      |
| 0.45           | 0.0652              | 0.0027                    | 0.0026                      |
| 0.5            | 0.072               | 0.0029                    | 0.0029                      |
| 0.55           | 0.0795              | 0.0032                    | 0.0032                      |
| 0.6            | 0.0895              | 0.0036                    | 0.0036                      |
| 0.65           | 0.1013              | 0.0042                    | 0.0041                      |
| 0.7            | 0.1203              | 0.005                     | 0.0048                      |
| 0.75           | 0.1531              | 0.0064                    | 0.0061                      |
| 0.8            | 0.1921              | 0.008                     | 0.0077                      |
| 0.85           | 0.2589              | 0.0119                    | 0.0143                      |
| 0.881          | 0.3568              | 0.0186                    | 0.012                       |
| 0.9            | 0.4647              | 0.025                     | 0.0186                      |
| 0.916          | 0.8113              | 0.0428                    | 0.0325                      |
model is included in the range of sorption curves illustrated in Fig. 2c. It should be noted that the magnitude of variation of the parameters is very different among the models. It will be verified that those differences do not influence the model selection through the choice of the kernel parameter.

### Reliability of the models

The purpose is to demonstrate that the unknown parameters of the sorption models are identifiable. From a theoretical point of view, the Structural Global Identifiability (SGI) property is evaluated for each model in Appendix C in Electronic Supplementary Material. As a synthesis, it is demonstrated that all eight models have parameters theoretically identifiable if a set of observations is obtained. The next section investigates the primary identifiability, i.e. if the parameters sufficiently influence the output of the model to be estimated with accuracy.

### Primary identifiability

The sensitivity function is computed for the eight models. An illustration is shown for the GAB model in Fig. 3a, b, for two values of parameter $p$. The latter corresponds to the lower and upper bounds of each parameter in $\Omega_p$ for the GAB model, according to Table 2. It is noticed that over the domain $\Omega_p$, the sensitivity functions of parameters $p_1$ and $p_2$ have the highest values. The sensitivity function of the parameter $p_3$ is at least two orders of magnitude lower than others. These results are consistent with the sensitivity metric $\gamma$ computed for each of the three parameters and presented in Fig. 4a. The parameter $p_3$ has a very negligible influence on the sorption model. The sensitivity of parameters $p_1$ and $p_2$ is higher, with $\gamma_1 > \gamma_2$. Thus, the parameter $p_1$ is the most sensitive of the model. Therefore, it is the easiest to identify from a practical point of view. As the water activity $a$ increases, the sensitivity of the parameter $p_3$ increases. Thus, one could imagine to use some observations for $a \in [0.05, 0.8]$ to estimate the parameter $p_1$ and, then, some observations

### Table 2

| Models | Parameter $p_1$ | Parameter $p_2$ | Parameter $p_3$ | Parameter $p_4$ |
|--------|----------------|----------------|----------------|----------------|
| GAB    | $\mathcal{U}(1.06 \times 10^{-2}, 5.31 \times 10^{-2})$ | $\mathcal{U}(0.95, 1.05)$ | $\mathcal{U}(5.171)$ | – |
| TRM    | $\mathcal{U}(1.0, 3.0)$ | $\mathcal{U}(1.0, 1.55)$ | $\mathcal{U}(2.0, 3.0)$ | – |
| OSW    | $\mathcal{U}(0.013, 0.14)$ | $\mathcal{U}(0.75, 1.28)$ | – | – |
| FX     | $\mathcal{U}(4.9, 40)$ | $\mathcal{U}(18.0, 18.69)$ | $\mathcal{U}(7, 18.5)$ | $\mathcal{U}(1.46, 1.7)$ |
| BET    | $\mathcal{U}(0.034, 0.09)$ | $\mathcal{U}(0.2, 10)$ | – | – |
| VG     | $\mathcal{U}(0.99, 179)$ | $\mathcal{U}(26.6, 3.73 \times 10^3)$ | $\mathcal{U}(1.7, 2.4)$ | – |
| SM     | $\mathcal{U}(0.0026, 0.013)$ | $\mathcal{U}(0.04, 0.20)$ | – | – |
| MADS   | $\mathcal{U}(-1.1, -0.4)$ | $\mathcal{U}(2.1, 2.9)$ | – | – |
for $a \in [0.8, 0.95]$ to retrieve $p_2$. Even with this procedure, the parameter $p_3$ cannot be estimated with accuracy.

The variation of the sensitivity metric $\gamma$ with $a$ is given for the other models in Fig. 4b–f. The results of the primary identifiability are reported in Table 3. A general observation is that five models out of eight have one parameter with a very high sensitivity, other parameters having a very low influence on the models output. In other words, these models have parameters with sensitivity metrics of the same orders of magnitude. One can conclude that their primary identifiability is very low for all the domain of $a$. The accuracy of the results of the parameter estimation problem might be very low. Three models, namely SM, TRM and MADS, have parameters with medium sensitivity for two parameters. For the SM and MADS model, both parameters have a similar influence on the model output. It shows a good primary identifiability for these models. For the TRM, parameters $p_2$ and $p_3$ have good primary identifiability. However, the accuracy of estimation for parameter $p_1$ might be poor since it has a very small influence on the model.

**Parameter estimation and model selection**

Previous investigations demonstrated that all models have parameters identifiable in theory. From a practical point of view, it has been shown that not all parameters of the models have sufficient influence on the model predictions to be retrieved with accuracy. Here, the parameter estimation problem is first solved to examine the consequence of a bad primary identifiability on the model reliability. Then, a selection is operated over the competing models to distinguish the most reliable ones.

**Parameter estimation**

First, GAuß algorithm is employed in the least squares sense to solve the parameter estimation problem. The estimated parameters and their estimated uncertainties are reported in Table 4. Figure 5a, c, e, g compares the prediction of the
Fig. 4 Variation of the global sensitivity metrics for the models according to the water activity
models computed with the estimated parameters and the experimental observations. Figure 5b, d, f, h presents the residuals between computations and observations. Globally, the discrepancies between the model and measurements are relatively low for all models in the so-called hygroscopic state $a \leq 0.80$. At high water activity $a \geq 0.80$, the discrepancies increase for almost all models, except the MADS one. This can be clearly noticed from the analysis of the residuals. The SM and MADS model residuals have a particular pattern which does not vary around zero. As indicated in Fig. 5h (and Fig. 2b), the pattern is similar to the standard deviation $\sigma$ of the measurements. Looking at the distance presented in Table 4, the models closest to the experimental observations are the MADS and FX ones. However, for the latter, as shown in Fig. 5d, the discrepancy is relatively high for $a \leq 0.2$. In addition, the model has four parameters to be estimated, which increases the complexity of the estimation problem.

Thus, in general, we note that the parameters could be estimated and the measurements could be accurately predicted with all models. Nevertheless, the reliability of the models may be discussed looking at the relative estimator error $\eta$ in Table 4. Indeed, some parameters are estimated with a very high error estimator. Namely, the parameters $p_3$ for the GAB model, $p_1$ for the TRM model, $(p_1, p_2, p_3, p_4)$ for the FX model, $p_2$ for the BET model, $(p_1, p_2)$ for the VG model and $p_1$ for the SM model are estimated with a very low accuracy. For the models FX and VG, the error estimator is higher or equal to half of the standard deviation of the a priori uniform distribution. These results are consistent with the ones of the primary identifiability. In other words, parameters with a very low influence on the model prediction correspond to the ones with a high error estimator. For instance, for the TRM model, a bad primary identifiability is observed for parameter $p_1$. Indeed, the estimation of parameter $p_1$ is very inaccurate with almost 70% of relative error. On the contrary, parameters $p_2$ and $p_3$ have good primary identifiability. Consequently, their error estimator is better, around 10%.

### Table 3: Results of the primary identifiability for each model

| Models | Global sensitivity indexes $\gamma^T$ |
|--------|--------------------------------------|
|        | $p_1$ | $p_2$ | $p_3$ | $p_4$ |
| GAB    | 0.89  | 0.10  | $O(10^{-6})$ | – |
| TRM    | 0.06  | 0.55  | 0.38  | – |
| OSW    | 0.975 | 0.025 | –      | – |
| FX     | $3 \times 10^{-4}$ | $3.7 \times 10^{-3}$ | $6.2 \times 10^{-3}$ | 0.98 |
| BET    | 0.99  | $3.2 \times 10^{-4}$ | –      | – |
| VG     | $7.7 \times 10^{-2}$ | $9 \times 10^{-4}$ | 0.99  | – |
| SM     | 0.48  | 0.52  | –      | – |
| MADS   | 0.54  | 0.46  | –      | – |
Fig. 5 Comparison between the model predictions computed with the estimated parameters and the experimental observations (a, c, e, g) and residual between both (b, d, f, h)
Model selection

Now, the ABC algorithm is used for model selection and model calibration (estimation of the model parameters), among the eight competing ones described above, for a kernel $\kappa_0 = 0.01$. A number of 22 populations are chosen as illustrated in Fig. 6a. The tolerance is decreasing with the number of population respecting the Morozov’s discrepancy principle at the final population. The last tolerance scales with the square roots of the sum of the uncertainties for each measurement $\varepsilon_{22} = 1.4 \cdot \sqrt{\sum_{i=1}^{N_a} \delta_i^2}$, $N_a$ being the total number of measurements. A total number $N_v = 4000$ particles is chosen. Figure 6b shows the variation of the acceptance rate according to each population. It decreases with the population. For the last population, the acceptance rate is 0.4%.

Table 4  Results of the parameter estimation problem for each model

| Models | Estimated parameter | Relative error estimator $\eta$ | Distance $d_m(p, \hat{u})$ |
|--------|---------------------|--------------------------------|--------------------------|
|        | $p_1^c$            | $p_2^c$            | $p_3^c$            | $p_4^c$            | $p_1^s$            | $p_2^s$            | $p_3^s$            | $p_4^s$            |
| GAB    | 0.035              | 1.027              | 15.33             | –                 | 0.12              | 0.04              | 0.64              | –                 | 0.05              |
| TRM    | 1.156              | 1.383              | 2.165             | –                 | 0.67              | 0.15              | 0.10              | –                 | 0.11              |
| OSW    | 0.069              | 0.75               | –                 | –                 | 0.05              | 0.09              | –                 | –                 | 0.15              |
| FX     | 7.98               | 18.39              | 9.586             | 1.471             | $>1$              | $>1$              | $>1$              | 0.48              | 0.007             |
| BET    | 0.04               | 9.18               | –                 | –                 | 0.07              | 0.44              | –                 | –                 | 0.12              |
| VG     | 51.33              | 2299               | 1.895             | –                 | $>1$              | $>1$              | 0.12              | –                 | 0.14              |
| SM     | 0.0078             | 0.099              | –                 | –                 | 0.53              | 0.09              | –                 | –                 | 0.38              |
| MADS   | $-0.76$            | 2.47               | –                 | –                 | 0.04              | 0.04              | –                 | –                 | 0.005             |

Figure 6  Variation of the tolerance (a) and the acceptance rate (b) according to the population
does not succeed in representing the phenomena for high water activity, as reported in Fig. 5g. At population 5, no particles validate the distance test for this model with the candidate parameters. Thus, at population 5, the model SM is no longer selected. At population 9, it can be noticed that only three models are still competing, namely the MADS, the FX and the GAB. At the final population, the MADS model is the only one satisfying the distance test for the lowest tolerance.

The model selection can also be discussed by analysing the evolution of estimated parameters according to the population presented in Fig. 7a, b. Figure 7a shows the dissemination of the estimated parameters in the plan $\Omega_{p_1} \times \Omega_{p_2}$ according to the population, for both models OSW and MADS. For the first population, the estimated parameters are scattered. At population 8, for the OSW model, the estimated parameters still exhibit a large variability. For the MADS model, the parameters are

![Parameter Dissemination](https://via.placeholder.com/150)

**Fig. 7** Variation of the estimated parameters (a) and their standard deviation (b) according to the population
already concentrated in a narrow region, corresponding to the final estimated parameters of the model. This analysis is consistent with the evolution of the standard deviation of the estimated parameters according to the population. Figure 7b shows that the GAB model is not selected anymore after population 11. Indeed, the standard deviation of the parameters estimated for this model is large. The algorithm did not find new candidate parameters that validate the distance test of the GAB model.

At the final population, the parameters are estimated for the selected model. Figure 8 gives the posterior distribution of the parameters $p_1$ and $p_2$ of the MADS model. It can be noticed that it corresponds to the one estimated in previous section, using the least estimator algorithm, in Table 4. In addition, the standard deviation of posterior distributions is very low indicating an accurate estimation.

To confirm the results, the model selection is performed for another kernel $\kappa_0 = 0.1$, while keeping 22 populations. As noted in Fig. 6b, the acceptance rate $\tau$ decreases when $\kappa_0$ increases. Thus, the computational time of the algorithm rises significantly (from 15 min for $\kappa_0 = 0.01$ to 30 min for $\kappa_0 = 0.1$ in the MATLAB™ environment with a computer equipped with Intel i7 CPU and 32 GB of RAM). The model selection among the 22 populations is shown in Fig. 9b. The results are unchanged, while the MADS model is selected. It is noted that additional simulations have been carried out for higher kernel $\kappa_0 = 1$ and 400 particles, also resulting in the selection of the MADS model. However, the acceptance rate was too low indicating a bad choice for the kernel parameter.

**Conclusion**

The sorption curve is an essential property for wood materials with the objective of modelling the interaction of wood with ambient moist air. This paper proposes to investigate the reliability of eight models through their robustness to the identified parameters. Seven have been proposed in the literature with various examples of applications. The last one is proposed based on the general shape of moisture
Experimental measurements are taken for a wood fibre material. Using a DVS equipment, the moisture content is obtained according to several levels of water activity.

Using the experimental observations, the reliability of the models is discussed in “Reliability of the models” section through the accuracy of the parameter estimation. The so-called primary identifiability of each parameter is discussed. It investigates whether the parameters sufficiently influence the output of the model to be estimated with accuracy. For this, a continuous derivative-based approach is adopted using the sensitivity function of the model. A global sensitivity metric is computed for each parameter. Seven models have a low primary identifiability for at least one parameter. In other words, one parameter is not influencing the model sufficiently to be estimated with accuracy. These results are confirmed when solving the parameter estimation problem in “Parameter estimation and model selection” section. For all models, a set of parameters can be identified to represent accurately the

Fig. 9 Variation of the model selection according to the population for different kernels
moisture sorption curve. However, the parameter with low primary identifiability is retrieved with a large error estimator. Last, an ABC algorithm is used for simultaneously model selection and model calibration, among eight competing models. The proposed model appears to have the best reliability based on the distance between measurements and estimation. The so-calleds GAB and Fredlund-Xing (Fredlund and Xing 1994) models are also reliable candidates. The first has a high parametric complexity since it is composed of four parameters.

Future works should focus on the hysteresis effects. It should be taken into account since it has a significant influence on the precision of the numerical predictions of heat and mass transfer (Zhang et al. 2016b).

Acknowledgements The authors acknowledge the Junior Chair Research program “Building performance assessment, evaluation and enhancement” from the University of Savoie Mont Blanc in collaboration with The French Atomic and Alternative Energy Center (CEA) and Scientific and Technical Center for Buildings (CSTB). The authors acknowledge the French and Brazilian agencies for their financial support through the project CAPES–COFECUB, as well as the CNPQ of the Brazilian Ministry of Education and of the Ministry of Science, Technology and Innovation, respectively, for co-funding. The support provided by FAPERJ, agency of the Rio de Janeiro state government, is gratefully appreciated. The authors also thank A. Moreau and A. Jumabekova for the productive discussions on the MADS and GAB models, respectively.

Compliance with ethical standards

Conflict of interests Authors have no conflict of interest relevant to this article.

References

Berger J, Guernouti S, Woloszyn M, Buhe C (2015) Factors governing the development of moisture disorders for integration into building performance simulation. J Build Eng 3:1–15. https://doi.org/10.1016/j.jobe.2015.04.008

Carmeliet J, Roels S (2002) Determination of the moisture capacity of porous building materials. J Therm Envel Build Sci 25(3):209–237. https://doi.org/10.1106/109719602022835

Colinart T, Glouannec P (2017) Temperature dependence of sorption isotherm of hygroscopic building materials. Part 1: experimental evidence and modeling. Energy Build 139:360–370. https://doi.org/10.1016/j.enbuild.2016.12.082

da Costa J, Orlande H, da Silva W (2018) Model selection and parameter estimation in tumor growth models using approximate Bayesian computation-abc. Comput Appl Math 37(3):2795–2815. https://doi.org/10.1007/s40314-017-0479-0

Dickinson RP, Gelinas RJ (1976) Sensitivity analysis of ordinary differential equation systems—a direct method. J Comput Phys 21(2):123–143. https://doi.org/10.1016/0021-9991(76)90073-6

Fredlund D, Xing A (1994) Equations for the soil–water characteristic curve. Can Geotech J 31(4):521–532. https://doi.org/10.1139/t94-061

Furmaniak S (2012) The alternative model of water vapour sorption in porous building materials. Transp Porous Med 95:21–23. https://doi.org/10.1007/s11242-012-0030-0

Glass S, Boardman C, Zelinka SL (2017) Short hold times in dynamic vapor sorption measurements mischaracterize the equilibrium moisture content of wood. Wood Sci Technol 51(2):243–260. https://doi.org/10.1007/s00226-016-0883-4

Glass S, Boardman C, Thybring E, Zelinka S (2018) Quantifying and reducing errors in equilibrium moisture content measurements with dynamic vapor sorption (DVS) experiments. Wood Sci Technol 52(4):909–927. https://doi.org/10.1007/s00226-018-1007-0
Iglesias H, Chirife J (1995) An alternative to the guggenheim, anderson and de boer model for the mathematical description of moisture sorption isotherms of foods. Food Res Int 28(3):317–321. https://doi.org/10.1016/0963-9969(94)00002-P

ISO12570 (2000) Hygrothermal performance of building materials and products—determination of moisture content by drying at elevated temperature

Jumabekova A, Berger J, Foucquier A (2019) An efficient sensitivity analysis for energy performance of building envelope: a continuous derivative based approach. Build Simul: Int J. https://doi.org/10.1007/s12273-020-0712-4

Karalashvili M, Marquardt W, Mhamdi A (2015) Optimal experimental design for identification of transport coefficient models in convection–diffusion equations. Comput Chem Eng 80:101–113. https://doi.org/10.1016/j.compchemeng.2015.04.036

Karolou M, Moropoulou A, Maroulis Z, Kroikida M (2005) Water sorption isotherms of some building materials. Drying Technol 23(1–2):289–303

Kucherenko S, Song S (2016) Derivative-based global sensitivity measures and their link with sobol’ sensitivity indices. In: Cools R, Nuyens D (eds) Monte Carlo and quasi-Monte Carlo methods. Springer, New York, pp 455–469. https://doi.org/10.1007/978-3-319-33507-0_23

Liepe J, Kirk P, Filippi S, Toni T, Barnes C, Stumpf M (2014) A framework for parameter estimation and model selection from experimental data in systems biology using approximate bayesian computation. Nat Protoc 9:439–456. https://doi.org/10.1038/nprot.2014.025

Lioiola BR, Orlande HR, Dulikravich GS (2020) Approximate bayesian computation applied to the identification of thermal damage of biological tissues due to laser irradiation. Int J Therm Sci 151:106243. https://doi.org/10.1016/j.ijthermalsci.2019.106243

Nilsson L (2018) Methods of measuring moisture in building materials and structures. Springer, Berlin

Ouertani S, Azzouz S, Hassini L, Koubaa A, Belghith A (2014) Moisture sorption isotherms and thermodynamic properties of jack pine and palm wood: comparative study. Ind Crops Prod 56:200–210. https://doi.org/10.1016/j.indcrop.2014.03.004

Patera A, Derluyn H, Derome D, Carmeliet J (2016) Influence of sorption hysteresis on moisture transport in wood. Wood Sci Technol 50:259–283. https://doi.org/10.1007/s00226-015-0786-9

Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity analysis in practice: a guide to assessing scientific models. Halsted Press, New York

Singh PC, Singh RK (1996) Application of gab model for water sorption isotherms of food products1. J Food Process Preserv 20(3):203–220. https://doi.org/10.1111/j.1745-4549.1996.tb00743.x

Skaar C (1988) Wood-water relations. Springer, Berlin. https://doi.org/10.1007/978-3-642-73683-4

Sobol IM (1990) Sensitivity estimates for nonlinear mathematical models. Mat Model 2:112–118

Sobol IM, Kucherenko SS (2009) Derivative based global sensitivity measures and their link with global sensitivity indices. Math Comput Simul 79(10):3009–3017

Soprema (2019) https://www.soprema.fr/fr/product/isolation/toiture/toiture-en-pente/rampants/pavaflex-confort, consulted the 23 Oct 2019

Stolarska A, Garbalinska H (2017) Assessment of suitability of some chosen functions for describing of sorption isotherms in building materials. Heat Mass Transf 53:1603–1617. https://doi.org/10.1007/s00231-016-1926-y

Taylor JR (1997) An introduction to error analysis: the study of uncertainties in physical measurements. University Science Books, Mill Valley

Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf M (2009) Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface 6(31):187–202. https://doi.org/10.1098/rsif.2008.0172

Ucinski D (2004) Optimal measurement methods for distributed parameter system identification. CRC Press, New York

Vololonirina O, Coutand M, Perrin B (2014) Characterization of hygrothermal properties of wood-based products—impact of moisture content and temperature. Constr Build Mater 63:223–233. https://doi.org/10.1016/j.conbuildmat.2014.04.014

Walter E, Lecourtier Y (1982) Global approaches to identifiability testing for linear and nonlinear state space models. Math Comput Simul 24(6):472–482. https://doi.org/10.1016/0378-4754(82)90645-0

Walter E, Pronzato L (1990) Qualitative and quantitative experiment design for phenomenological models: a survey. Automatica 26(2):195–213. https://doi.org/10.1016/0005-1098(90)90116-Y

Willems W (2014) Hydrostatic pressure and temperature dependence of wood moisture sorption isotherms. Wood Sci Technol 48(3):483–498. https://doi.org/10.1007/s00226-014-0616-5
Zhang X, Zillig W, Kunzel H, Zhang X, Mitterer C (2015) Evaluation of moisture sorption models and modified Mualem model for prediction of desorption isotherm for wood materials. Build Environ 92:387–395. https://doi.org/10.1016/j.buildenv.2015.05.021

Zhang X, Zillig W, Künzel H, Mitterer C, Zhang X (2016a) Combined effects of sorption hysteresis and its temperature dependency on wood materials and building enclosures—part II: hygrothermal modeling. Build Environ 106:181–195. https://doi.org/10.1016/j.buildenv.2016.06.033

Zhang X, Zillig W, Künzel H, Mitterer C, Zhang X (2016b) Combined effects of sorption hysteresis and its temperature dependency on wood materials and building enclosures—part I: measurements for model validation. Build Environ 106:143–154. https://doi.org/10.1016/j.buildenv.2016.06.025

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Julien Berger1 · Thibaut Colinart2 · Bruna R. Loiola3 · Helcio R. B. Orlande4

1 Laboratoire des Sciences de l’Ingénieur pour l’Environnement (LaSIE), UMR 7356 CNRS, La Rochelle Université, CNRS, 17000 La Rochelle, France
2 Univ. Bretagne Sud, UMR CNRS 6027, IRDL, 56100 Lorient, France
3 Mechanical Engineering Department, Military Institute of Engineering, Rio de Janeiro 22290-270, Brazil
4 POLI/COPPE, Mechanical Engineering Graduate Program, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil