

Hygric characterization of wood fiber insulation under uncertainty with dynamic measurements and Markov Chain Monte-Carlo algorithm

Simon Rouchier^a **, Thomas Busser^{a,b}, Mickaël Pailha^a, Amandine Piot^b and Monika Woloszyn^a

^aLOCIE, CNRS-UMR5271, Université de Savoie; Campus Scientifique, Savoie Technolac, 73376 Le Bourget-du-Lac Cedex, France

^bUniv. Grenoble Alpes, INES, F-73375 Le Bourget du Lac, France
CEA, LITEN, Department of Solar Technologies, F-73375 Le Bourget du Lac, France

Postprint: Rouchier S, Busser T, Pailha M, Piot A, Woloszyn M 2017. Hygric characterization of wood fiber insulation under uncertainty with dynamic measurements and Markov Chain Monte-Carlo algorithm, *Building and Environment* vol. 114, p. 129-139

Abstract

The present work is the hygric characterization of wood fibre insulation boards, using dynamic measurements of relative humidity and sample weight, analyzed in the frame of Bayesian inference for parameter identification under uncertainty. It is an attempt at identifying detailed profiles of moisture-dependent properties, and thus a relatively high number of parameters. Because of this ambition, some caution should be exercised once the outcome of the inversion algorithm is available: in addition to confidence intervals of parameters provided by the Bayesian framework, a simplified form of identifiability analysis is performed by analysing a posteriori parameter correlations and likelihood-based confidence intervals.

The characterization methodology does not require for the model structure to have a differentiable analytical formulation, or for material samples to reach mass equilibrium between each RH step of the experimental process. Two separate experimental designs were used for material characterization and for validation, respectively. Results show a clear relation between available information (experimental data) and inference (confidence intervals of parameters). A single relative humidity step is not informative enough for a precise inference of moisture-dependent properties such as vapour permeability and moisture capacity. A two-step experiment however holds enough information to significantly reduce parameter uncertainty.

Keywords HAM ; characterization ; MCMC ; identifiability

1 Introduction

Heat, air and moisture (HAM) transfer phenomena in porous building materials may be simulated with various levels of detail, from simplified models (Effective Moisture Penetration Depth, Moisture Buffer Value)

**Corresponding author. Email: simon.rouchier@univ-savoie.fr

Table 1: List of abbreviations

HAM	Heat, air and moisture
RH	Relative humidity
MCMC	Markov Chain Monte-Carlo
ACF	Auto-correlation function
MLE	Maximum likelihood estimator
FEM	Finite-element method

to detailed FEM calculations of coupled transfer. In the latter case, the extensive knowledge of material properties such as liquid conductivity, water vapour permeability and retention curve may be required for an accurate prediction of these phenomena. This is problematic because of constraints on the measurement of each of these properties: the estimation of a sorption isotherm is time consuming (ISO 12571, 2013); the standard method for measuring water vapour permeability is also a time-consuming steady-state method (ISO 12572, 2013) and its accuracy concerning highly hygroscopic materials has been questioned [14]; the full profile of liquid conductivity requires specific equipment in addition to a simple moisture uptake test. Such issues motivate the application of inverse methods for characterization, relying on dynamic measurements interpreted by advanced algorithms.

The purpose of inverse problems is to learn the structure or the parameters of a model from the physical observation of its outputs. Inverse methods are now well anchored in the field of heat transfer [4] and there is a large, well-documented knowledge base for addressing their challenges [24, 25, 2, 21, 34]. In the general field of building physics, inverse methods are applied in many particular cases, though without a systematic reference for proper use. They are not quite as widespread as in other fields but are progressively gaining interest because of their increasing accessibility to non-mathematicians. Possible applications may be, among others: inverse modelling of multi-zone buildings [11, 22]; in-situ diagnostics of thermal properties [8, 7, 6]; model selection and calibration for thermal retrofit investigation [20].

Hygrothermal material property characterization is also a recent application of optimization-like algorithms, because of its constraints mentioned above. They allow the interpretation of dynamic temperature and concentration measurements, which contain more information than steady-state observations traditionally used by standard methods. [5] address coupled heat and mass transfer and identify non-dimensional transport and storage properties of several materials through a Levenberg-Marquardt optimization scheme. [33] add regularisation to a Covariance Matrix Adaptation algorithm to identify the parameters of a coupled HAM model. [13] show how the Moisture Buffer Value test can be improved by using its dynamic measurements as observations in a Bayesian inference scheme. [28] also analyse dynamic RH measurements to infer a mass diffusion coefficient of several hygroscopic materials. A common questioning in many applications of inverse methods is: what level of detail can be reached in the description of each property (such as moisture permeability, sorption isotherm), by harnessing a given set of observations. This question can be inversely formulated: which experimental design, possibly with a moderate time cost, will yield the most informative characterization of a given parameterization. This is the purpose of optimal experimental design [15].

Another important questioning, and a crucial step in the discussions on the results of an inverse problem, is how reliable these results are. The main difficulty in solving an inverse problem arises from the combination of its ill-posed nature and the amplification effect of the measurement noise [35]. In the case of linear inverse problems, this is equivalent to inverting a badly-conditioned matrix [24]. The consequence is that small measurement errors may result in large error on parameter estimations. Thus, the identifiability of each parameter is not always guaranteed: the solution of an inverse problem must therefore be presented along with a measure of its reliability. We will argue that the application of an inverse technique should be assorted to a form of sensitivity or identifiability analysis [9, 31]. Other difficulties may be added: the search space dimension, or number of unknown parameters, quickly increases the computational time required to solve a problem; a proper parameterisation of the model, though essential for its identifiability, is not always trivial; etc.

The present work is the hygric characterization of wood fibre insulation boards. Material samples were

placed in a controlled environment and submitted to stepwise RH variations. Measurements of their weight and internal humidity were then interpreted by a Markov Chain Monte-Carlo (MCMC) algorithm to estimate their moisture permeability and sorption isotherm. The results of this parameter estimation phase were then validated against observations from a second experiment, in an entirely different setup.

It is an attempt at identifying a somewhat higher number of parameters than recent similar studies [13, 28]. Because of this ambition, some caution should be exercised once the outcome of the inversion algorithm is available. The reliability of identified parameters relate to their identifiability in the outline of the problem: model structure and available data. Bayesian inference, implemented by the MCMC algorithm, was applied for the interpretation of dynamic RH and sample weight measurements, after stepwise RH variations in a controlled environment. This allowed estimating a set of seven parameters along with their confidence intervals. Although we do not demonstrate a rigorous identifiability analysis, we wish to provide some guidelines for a cautious interpretation of the results of HAM characterization. This was done by analysing posterior correlations between parameters and pairwise likelihood profiles [31] to detect eventual non-identifiability. As a last step, characterization results were validated on the basis of a second experimental design, in which samples were submitted to outside weather boundary conditions.

The paper is made of three main sections.

- **Sec. 2** describes the scope of the study and provides general considerations on HAM modelling, characterization, Bayesian inference and identifiability analysis;
- **Sec. 3** describes the hygric characterisation of wood fibre: experimental setup, parameterization and discussion on results;
- **Sec. 4** is the validation phase: a separate experimental design is presented and its measurements are compared to predictions arising from the characterization phase.

2 Background

2.1 HAM modelling

The physical phenomenon under investigation is the one-dimensional water vapour transfer inside a single material sample. The modelling of heat and moisture transport processes in open porous material was dealt with in the main theory of [29] and adapted to building components [23]. Assuming that liquid transfer is overlooked, the mathematical model for moisture transport Eq. 1 and boundary condition Eq. 2 are written here with the vapour pressure p_v as driving coefficient.

$$\frac{\partial w}{\partial t} - \nabla \cdot [\delta_p \nabla p_v] = 0 \quad (1)$$

$$-\delta_p \nabla p_v = h_m (p_{v, \text{amb}} - p_v) \quad (2)$$

where w (kg/m³) is the volumetric moisture content and δ_p (kg/(m.s.Pa)) is the water vapour permeability. The boundary condition Eq. 2 involves a surface transfer coefficient h_m (kg/(m².s.Pa)) and the ambient vapour pressure $p_{v, \text{amb}}$.

Due to the high dependency of moisture transfer to temperature gradients [29], the temperature distribution within a medium is generally calculated by solving the heat conservation equation along with moisture conservation equation:

$$(c_p \rho + c_{p,l} w) \frac{\partial T}{\partial t} + (c_{p,l} T) \frac{\partial w}{\partial t} - \nabla \cdot [\lambda \nabla T + L_v \delta_p \nabla p_v] = 0 \quad (3)$$

where $c_p \rho$ [J/(m³.K)] its volumetric thermal capacity and λ [W/(m.K)] its thermal conductivity. $c_{p,l}$ and L_v are the specific heat and latent heat of evaporation of water.

Although the present work only focuses on characterising moisture transport and storage properties, Eq. 1 and 3 will be solved in a fully coupled fashion. The HAM modelling tool used for this purpose is the Hamopy Python package [32].

2.2 HAM characterization

The evolution of relative humidity and moisture content within a material is predicted by Eq. 1. This equation involves transport and storage coefficients, which must be estimated by experimental characterization methods.

The most common measurement procedure for vapour permeability is the steady-state cup method, following the standard (ISO 12572, 2013). Its principle is to measure a steady mass variation under a controlled vapour flow driven by a difference of RH between both sides of a sample. The mass uptake or decrease profile of the whole cup is then linked to the vapour permeability of the sample. Since the vapour permeability is moisture dependent, a dry and wet cup experiment can be carried with different RH levels. This test is standard and widely used, but its ability to characterise highly hygroscopic materials has been questioned [14]. Because the method does not account for the variation of total pressure within the cup, advective vapour transfer is overlooked. In the case of a very permeable material, this implies that the permeability may be underestimated.

The storage function is traditionally measured by gravimetric methods (ISO 12571, 2013). Material samples are successively placed in different relative humidity conditions controlled by salt solutions. The equilibrium mass uptake is then measured at each point. These methods are highly time consuming, as several points are required to draw a sorption isotherm curve, each of which may only reach equilibrium after several weeks.

Some recent works have attempted to overcome these long characterization periods, by using dynamic measurement data instead of steady-state methods. [33] use the Covariance Matrix Adaptation evolutionary algorithm to solve the optimization problem of fitting a HAM model on dynamic temperature and RH observations. [28] apply another gradient-free optimization algorithm (Simplex) to minimize the residuals between RH measurements and predictions, and estimate the moisture diffusivity of spruce wood and fiber board without reaching mass equilibrium. [13] use the Moisture Buffer Value experimental protocol to derive more material properties than just the moisture buffering capacity. These papers show that HAM characterization can be solved as an inverse problem, where the accuracy of the outcome depends on the available data. The present work pursues the same target and applies the Bayesian inference technique described below.

2.3 Bayesian inference

2.3.1 Principle

In the Bayesian framework, all quantities are stochastic variables and carry a certain amount of uncertainty. As opposed to deterministic methods seeking to find only a single solution for an inverse problem, Bayesian inference describes sought parameters as probability density functions and returns not only point and spread estimates of the likely solutions, but also a complete description of their uncertainty, conditioned by potential measurement noise and inaccuracy [34]. Solving inverse heat transfer problems in the Bayesian framework is fairly recent [21, 38], but already has several interesting examples in the building physics field. These applications mainly fall within two categories: the calibration of building energy models [40, 20], and the characterisation of thermal properties of materials and components [7, 6].

The principle of Bayesian inference is to update an initial probability distribution describing a prior knowledge on the unknown θ into a posterior distribution accounting to knowledge gained by observation. The basis for it is the Bayes theorem:

$$p(\theta|y) \propto p(y|\theta) p(\theta) \quad (4)$$

- $p(\theta)$ are the prior distributions of the parameter vector θ , describing all initial knowledge the expert may have. The prior is generally a result of subjective information.
- $p(y|\theta)$ is the likelihood of observations given model parameters. It gives the relative probability density of measurement outcomes y with a fixed θ .

- $p(\theta|y)$ is the posterior distribution of θ , i.e. its probability density given observations.

Bayesian inference inherently introduces a bias in the solution of the inverse problem. Indeed, the solution will always be a compromise between the prior knowledge on θ and the information brought by observations. The respective weight of these two (sometimes contradictory) sources of information is determined on one hand by the quantity of available data, and on the other hand by the refinement of the prior distribution. Whether or not the expert has a strong initial opinion on the supposed value of the parameters, will influence the solution. In terms of inverse problem theory, the role of the prior is equivalent to a form of regularisation: it helps overcoming the ill-posedness of the problem, and to increase the identifiability of unknown parameters.

Note that the characteristics of noise are not necessarily known. If they are not, this information can be captured by incorporating the parameters controlling noise distributions into the problem unknowns [38]. Similarly, the exact position of temperature and humidity sensors may be uncertain: the Bayesian framework allows stating a prior knowledge on this position as a probability density function instead of a single value. This uncertainty then propagates towards the confidence intervals of the estimated parameters.

2.3.2 Metropolis-Hastings algorithm

In the generic case where the formulation of the likelihood function is the outcome of a grey-box or black-box model, the posterior distribution $p(\theta|y)$ cannot be expressed analytically. A numerical sampling method is then required to provide a series of samples which approximate the posterior distribution. The Metropolis-Hastings (MH) algorithm [19] is one of the Markov Chain Monte Carlo (MCMC) methods and is used here for this purpose.

A Markov Chain is a sequence of random variables $\{x_n, n = 0, 1, 2, \dots\}$ which satisfy the following property: $p(x_{n+1}|x_n, x_{n-1}, \dots, x_0) = p(x_{n+1}|x_n)$, i.e. the transition from the state n to the state $n + 1$ depends solely on the current position, and not on previous ones. MCMC methods, including the MH algorithm, are designed to generate a sequence that will approximate a specific probability distribution at $n \rightarrow \infty$ [17].

The chain starts with an initial guess θ_0 . Each iteration of the chain from the state n to $n + 1$ then runs the following steps [21]:

1. Propose a new state θ' from a proposal distribution $q(\theta'|\theta_n)$
2. Calculate the acceptance ratio α of the new candidate:

$$\alpha(\theta'|\theta_n) = \min \left[1, \frac{p(\theta'|y) q(\theta_n|\theta')}{p(\theta_n|y) q(\theta'|\theta_n)} \right] = \min \left[1, \frac{p(y|\theta') p(\theta') q(\theta_n|\theta')}{p(y|\theta_n) p(\theta_n) q(\theta'|\theta_n)} \right] \quad (5)$$

This step requires calculating the likelihood $p(y|\theta')$: this counts as one evaluation of the model.

3. Accept the candidate θ' with probability α ($\theta_{n+1} = \theta'$) or reject it with probability $1 - \alpha$ ($\theta_{n+1} = \theta_n$)

This algorithm generates a sequence $\{\theta_n, n = 0, 1, 2, \dots\}$ which accurately represents the posterior distribution if the chain is long enough. An explanation on the meaning of “long enough” is given below. At each iteration, α is the probability for a new candidate to be accepted as the new state of the chain. Step 3 means that candidates which offer a likelihood improvement are systematically accepted, while candidates which imply a likelihood decrease have some chance of being accepted. This allows the posterior space to be highly explored

The only available lever for the user is the choice of proposal distribution $q(\theta'|\theta_n)$. A good choice for q is crucial in order to improve the performance of a chain (i.e. how fast it will converge to a stationary distribution). The most common choice is a random walk proposal: $\theta' = \theta_n + \gamma$, where γ is a vector of random variables with a distribution centered around zero, either uniform or Gaussian, with a specific width. This width will directly influence the acceptance rate of the chain, as is illustrated by [21]: too high, it results in a high rejection rate of candidates and thus generates a stationary chain; too low, the chain

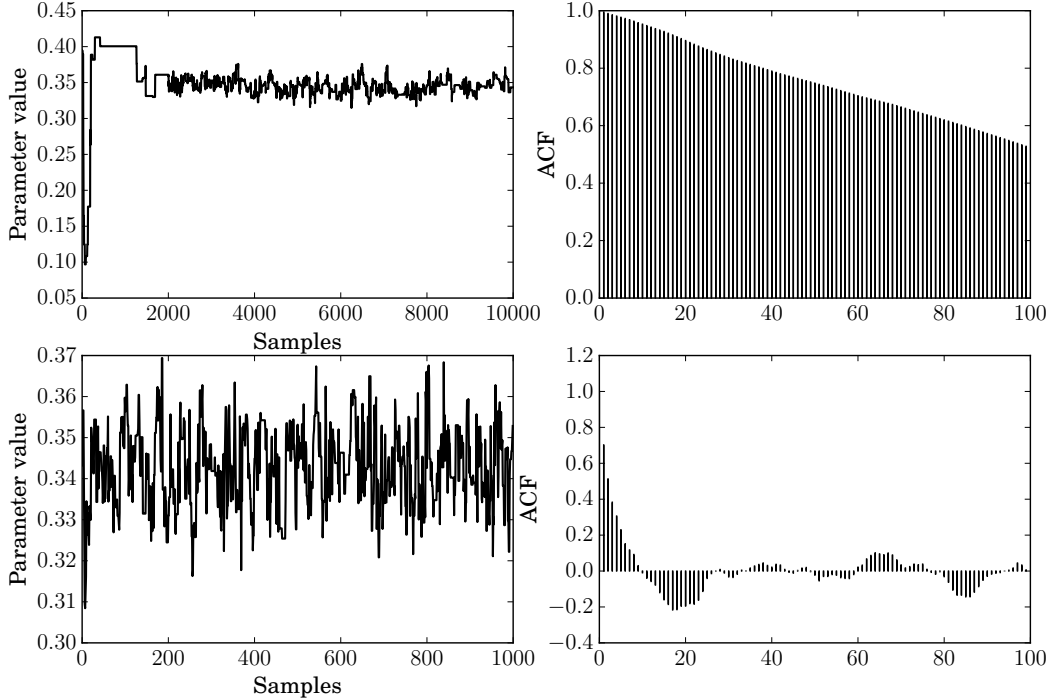


Figure 1: Up: trace and autocorrelation function (ACF) of a parameter before burn-in and thinning. The beginning of the chain must be discarded until the tuning of the adaptive proposal distribution is over. The chain is also strongly autocorrelated and thus requires thinning. Down: trace and ACF after sufficient burn-in and thinning. The ACF has decreased sufficiently and the remaining samples can be considered representative of the posterior distribution.

takes a very long time to explore the state space and generate independent samples. Since a good initial guess for the proposal distribution $q(\theta'|\theta_n)$ is not elementary, we use an adaptive scheme [18]: the width of the distribution is regularly tuned according to the evolution of the acceptance rate. It is reasonable to aim for an acceptance ratio between 20% and 50%.

The first convergence diagnostics of an interrupted chain is a visual examination of its trace in each dimension of the search space. This matter is illustrated on Fig. 1. In the sense of MCMC, convergence means stationary traces with low auto-correlation. First, all samples during which the chain is exploring the state space must be discarded since they are not representative of the posterior distributions: this is denoted as the burn-in period and can typically cover the first half of the chain. Such a period is visible on the 2,000 first samples of the chain shown on the upper left part of Fig. 1. Then, the chain still produces correlated samples because of the nature of the stepping procedure described above. The number of samples kept after the burn-in period must then be reduced by a thinning factor of up to 10. The criterion for an appropriate thinning factor is the autocorrelation function of the trace (see right part of Fig. 1), which must quickly fall to values below 0.1.

The stopping criterion of the chain is when it correctly represents independent samples from the posterior distribution. Convergence diagnostics are essential for assessing whether the chain can be stopped or if it needs further burn-in or thinning. Many methods exist for convergence diagnostics in a less empirical way [16, 27]. Some trial and error may be necessary until convergence can be judged satisfactory, but should not be much of a practical issue since a chain can be paused, examined, and resumed if necessary.

2.4 Parameter identifiability

The notion of identifiability can be defined as the possibility for a parameter to be correctly estimated, given a specific set of measurements and model structure. A parameter is identifiable if the confidence interval of its estimate is finite [31]. It relies on the sensitivity of the parameter, i.e. its importance in the model, but also on the type of available measurements. A typical cause for a low identifiability is either an inconvenient parametrization of the problem, or insufficient observations.

Sensitivity is not a sufficient condition for identifiability: a parameter may have a strong influence on a model output without being easily identifiable. This is typically the case when two or more parameters are strongly coupled by the way the problem is parameterized. A model calibration problem may be unsolvable as two parameters are pairwise dependent and exhibit any number of combinations as possible solutions. In the case of inverse heat transfer problems, a good example for this is when attempting to estimate both the thermal conductivity λ and capacity c_p from a single series of temperature measurements: only their ratio (thermal diffusivity) may be identified, leading to an infinite number of solutions for λ and c_p . The matter of identifiability becomes a crucial issue when large models are involved [9], impaired by an overparametrization with respect to the available observations, or when these observations are partial [31].

Structural identifiability, or theoretical identifiability [3], is satisfied if the following condition is met [37]:

$$\mathcal{M}(\theta, u) = \mathcal{M}(\theta^*, u) \implies \theta_i = \theta_i^* \quad (6)$$

In other words, an identifiable model structure yields distinct outcomes with distinct parameters: structural identifiability relates the identification accuracy to the parameterisation of the problem. In [36], a quantitative measure of this identifiability is given by the formulation of a data-independent information matrix, assuming a linear model structure. This allows the authors to reduce the parameterization dimensionality towards the best parameter identifiability.

Practical identifiability [9] relates the parameter estimation to the experimental design (type and amount of measurements), independently from the model structure. To test for data-dependent identifiability, [1, 12] consider the Fisher information matrix which bounds the covariance of the parameter estimation error. [12] calculate this matrix based on a pre-estimation data set in order to filter out non-influential parameters.

In the present work, the model structure is not linear and there is no explicit formulation of the sensitivity matrix. Nonidentifiability of parameters will therefore be detected *a posteriori* in the form of a strong correlation in the posterior density [30], and related to the size of likelihood-based confidence intervals [31]. Adding a rough identifiability analysis to a HAM characterization problem is one of the novelties of the present work.

3 Identification

A first experimental setup is used for the identification of moisture transport and storage properties of wood fibre insulation boards. Validation of these results will then follow in Sec. 4.

3.1 Experimental setup

The test facility, called RH-Box and shown on Fig. 2(a), is made of two connected climatic chambers. The temperature is controlled with a thermostatically-controlled water bath allowing water to circulate in a heat exchanger inside the chamber. In each chamber, RH is independently controlled by saline solutions. Two locks, at each side, allow the operator to insert or remove samples while minimizing system disturbances. An airlock links the two chambers and facilitates an instantaneous change in humidity for samples while passing from one box to another. Two fans, operating continuously, allow to homogenize air inside box when saline solutions are renewed. A more thorough description of the experiment was given by [10].

The sensors for measuring the RH and the temperature inside the samples and RH-Box are HygroPuce from Waranet, with an accuracy of $\pm 2\%$ RH and $\pm 0.5^\circ\text{C}$. These are wireless sensors, 0.6 cm thick and 1.6

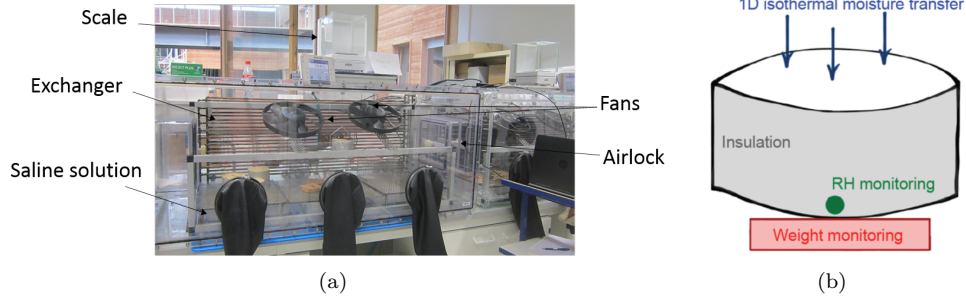


Figure 2: (a) RH-Box: double climatic chamber regulated in temperature and relative humidity (b) Setup of material samples and sensors inside the box

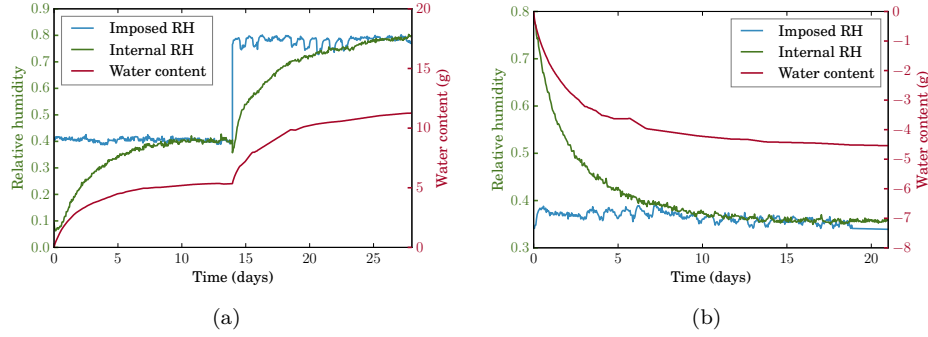


Figure 3: Stepwise imposed RH within the box, and records of sample weight and internal RH: (a) Increasing RH, (b) Decreasing RH

cm in diameter. The sample weight evolution is recorded by a precision scale (± 0.2 mg) placed outside the chamber, linked to a suspended plate within the box and connected to an automatic recording system.

The material investigated here is wood fibre insulation, the same that was used in the French collaborative project Hygrobat [39]. Samples are cylindrical, with a 10 cm diameter and 8 cm thickness. Their setup is schematically shown on Fig. 2(b): to ensure one-dimensional moisture transfer, the side and bottom surfaces of samples were covered with aluminium tape, glued on a white acrylic seal. In addition to weight monitoring, RH sensors were inserted within the samples: either one, two or three sensors. The present study only focuses on one sample, in which a single RH sensor was placed at the bottom (see Figure 2(b)).

The experimental protocol is based on the study reported by [28], where mass diffusion coefficients were determined by inverse methods as well. The experiment consists in submitting one face of the samples to variations of RH, while measuring RH on its back face. In our case, weight monitoring is an addition to the initial protocol of [28].

The samples were first pre-conditioned and dried in an oven at 70°C and 10% RH. They were then placed in the RH-Box and submitted to three successive steps: 33% with a MgCl_2 solution, 75% with a NaCl solution, and 33% again. Measurements obtained from this protocol are displayed on Fig. 3, where increasing and decreasing RH phases are separated.

3.2 Parameterization

For the purpose of parameter identification, all unknown properties of the mathematical problem (Eq. 1 and 2) must be summarised into a finite parameter vector. A proper choice of parameterisation is the key to a successful and informative characterisation procedure, as previously explained in Sec. 2.4.

- The vapour permeability δ_p is a function of the relative humidity. It is modelled by a piecewise linear interpolation of $\log(\delta_p)$ between three unknown values $\delta_{p,25\%}$, $\delta_{p,50\%}$ and $\delta_{p,75\%}$.
- The surface transfer coefficient h_m is considered unknown.
- The sorption isotherm is modelled by a third-degree polynomial. This assumption is generally correct within the hygroscopic range, if the moisture content does not exceed the capillary threshold. The moisture retention curve is represented by three unknown scalar values $\xi_{25\%}$, $\xi_{50\%}$ and $\xi_{75\%}$, which are the values of the moisture capacity ξ at RH = 25%, 50% and 75% respectively:

$$\xi = \frac{\partial w}{\partial \text{RH}} \quad (7)$$

These choices for the formulation of material properties result in an unknown parameter vector θ of dimension $n = 7$:

$$\theta = \{h_m, \delta_{p,25\%}, \delta_{p,50\%}, \delta_{p,75\%}, \xi_{25\%}, \xi_{50\%}, \xi_{75\%}\} \quad (8)$$

Since the heat conservation equation 3 is being solved, the modelling relies on the values of thermal properties. These properties have been measured separately: $\lambda = 0.038$ W/m.K; $\rho = 146$ kg/m³; $c_p = 1103$ J/kg.K.

This parameterization is purposely more descriptive than recent similar studies [13, 28]. The reason is that we wish to investigate whether the available data is sufficient to learn generic shapes of the sorption isotherm and permeability profiles, and to demonstrate identifiability issues arising from parameter interactions.

Measurements gathered in the increasing and decreasing RH phases were used separately in two runs of the MCMC algorithm, in order to obtain two sets of values for θ . The target is to identify the sorption and desorption curves separately, and to ensure that the permeability has roughly the same value in both directions.

3.3 Results

Two runs of the MCMC algorithm were performed, labelled Sorption and Desorption below, according to the data set used as observations (Fig. 3(a) and 3(b), respectively).

3.3.1 Fitting and residuals

In the following, W denotes the mass variation of a sample since the beginning of a measurement period, i.e. the total water content uptake or loss over its volume V at time t :

$$W(t) = \int_V (w_t - w_{t=0}) dV \quad (9)$$

Results of the MCMC search are first displayed in terms of fitting between observations y^D and predictions y . Fig. 4(a) and 4(b) respectively show the fitting of RH and W in the increasing and decreasing RH phases, respectively. The set of parameters used for this prediction is the mean of the posterior distribution in each of its dimensions $\bar{\theta}$. Fig. 4(c) and 4(d) respectively show the mean residuals $r(\bar{\theta}) = y(\bar{\theta}) - y^D$.

An accurate estimation of the real material properties should yield uncorrelated residuals [25]. The graph of $r(\bar{\theta})$ should resemble a profile of white noise, for the model structure to accurately recreate the physical phenomenon and the estimator θ to be correct.

In the present results, the mean residuals $r(\bar{\theta})$ are not entirely uncorrelated: in addition to noise, average variations of roughly $\pm 2\%$ RH and ± 0.2 g can be seen on 4(c) and 4(d). In addition, the RH residuals of the sorption phase occasionally show higher deviations of up to 7 % RH. This suggests two possible, non mutually exclusive explanations:

- An inadequacy of the model due to some simplification in its formulation;
- RH or W measurement error that varies in time, possibly caused by an uncontrolled part of the experimental setup.

This uncorrelated error should however be put in perspective as its order of magnitude is low.

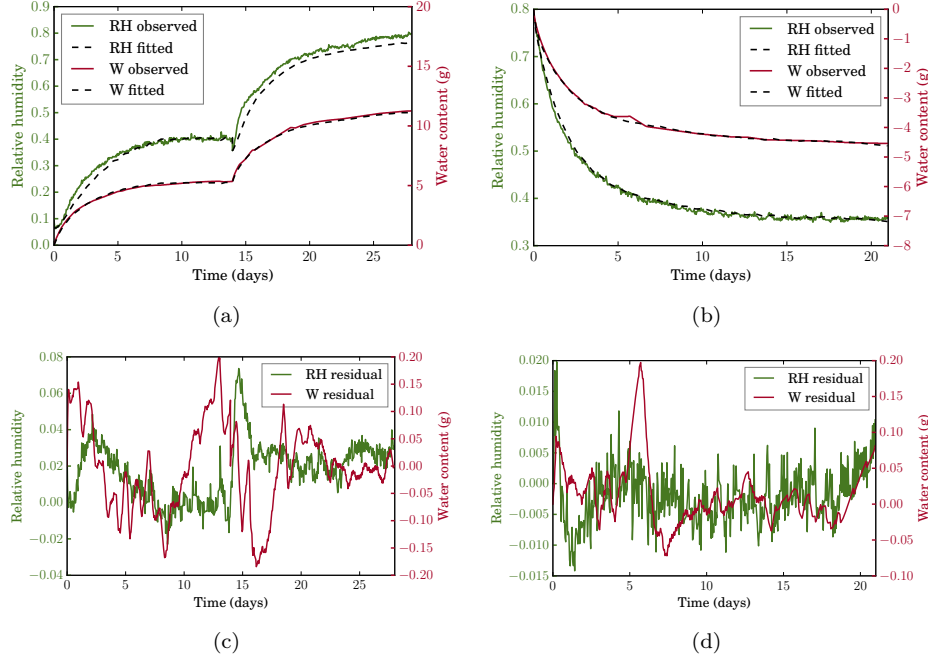


Figure 4: Average fitting of RH and mass observations: (a) increasing and (b) decreasing steps. Below: residuals between observations and maximum likelihood estimator

3.3.2 Parameter values

Prior to the present study, the wood fibre boards under investigation have had their hygrothermal properties characterized within the Hygrobat project [39]. Specifically, its water vapour resistivity has been measured by the standard wet cup and dry cup technique (ISO 12572, 2013), while its sorption and desorption isotherm were measured by gravimetric techniques (ISO 12571, 2013). A reference value for these parameters is therefore available for comparison.

The first identified parameter is the surface transfer coefficient h_m influencing the boundary condition Eq. 2. As already mentioned, it has been identified twice: using sorption and desorption separately (Fig. 3(a) and 3(b), respectively). Fig. 5 displays its posterior distributions in both cases.

The first observation concerning this result is how broad the posterior distribution of h_m is in the desorption experiment, compared to the sorption one. A trivial explanation would be because the desorption experiment is shorter and only consists in one relative humidity step, whereas the sorption experiment has two. More observations imply more information and more knowledge gained on the parameter. The second observation is that both distributions hardly overlap, although this parameter should roughly yield the same value in both cases, because of a similar experimental design. However, according to [28], this parameter only has a mild influence on the solution of the forward problem. This could explain some difficulty in its identification.

The apparent moisture permeability profiles are then shown on Fig. 6, for both the sorption and the desorption experiment. For each of them, the coloured area is the 95% confidence interval, i.e. the area delimited by the limits of the 95% confidence intervals on the posterior distributions of $\delta_{p,25\%}$, $\delta_{p,50\%}$ and $\delta_{p,75\%}$. The uncertainty arising from the desorption experiment is again much higher than from the sorption one, as was seen on the posterior distributions of h_m .

Another important observation is how high the identified values of moisture permeability are. In order for model predictions to fit with measurements, the search algorithm returns values for δ_p which are up to ten times higher than from the wet cup and dry cup experiments. This result correlates previous observations

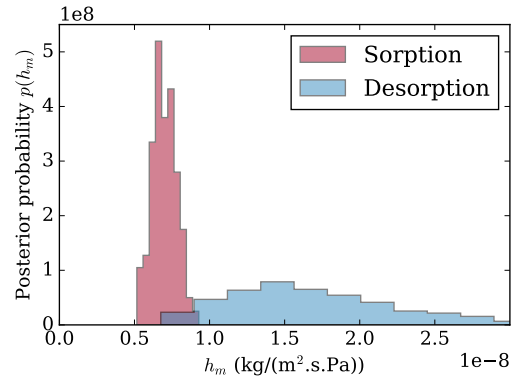


Figure 5: Posterior distributions of surface transfer coefficient

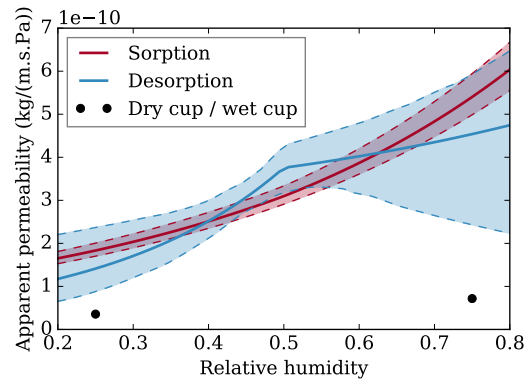


Figure 6: Profile of apparent moisture permeability with 95% confidence intervals

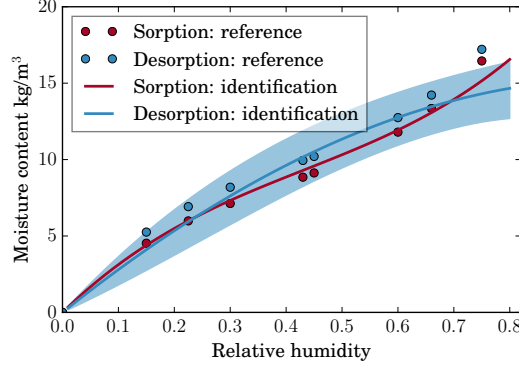


Figure 7: Sorption and desorption isotherms with 95% confidence intervals, compared to reference measurements

from validation studies [39]. If this were only explained by purely diffusive water vapour transfer, this would amount to resistance factors lower than 1, i.e. vapour diffusion is faster than in air. In this highly hygroscopic and permeable material, a possible explanation is that this *apparent* permeability includes effects which are not explicitly formulated in Eq. 1: a significant air permeability of the material driving moisture by convection.

The identified sorption and desorption isotherm profiles are displayed on Fig. 7. of $\xi_{25\%}$, $\xi_{50\%}$ and $\xi_{75\%}$, with 95% confidence intervals. The curves are only displayed up to RH 80%, which is the highest humidity reached by the experiment.

The sorption isotherm can be seen to fit very well with the reference measurements from the original material characterization. Its confidence intervals are very narrow, suggesting that the measurement period (four weeks, see Fig. 3(a)) can be shortened without compromising the identification uncertainty. This can be seen as a very satisfactory improvement from the traditional method. The desorption curve however has broader confidence intervals, since only one stepwise change, from 75% to 33% RH, was imposed on the samples for this phase of characterization. Its position with respect to the sorption curve is coherent, but does not extrapolate well outside of the RH range of the experiment. Around 75% RH, the curvature of both the sorption and the desorption curve are underestimated as a result of the lack of observations above this humidity range.

3.3.3 Identifiability

A posterior distribution may be well refined yet inaccurate: this may happen in case of an inappropriate model structure, or if the measurement error is not an independent and identically distributed (iid) random variable [25]. We therefore need a measure for the quality of identification results besides posterior refinement.

As previously mentioned, non-identifiability of parameters can be highlighted as a strong correlation among parameters in the posterior density [30]. This can be a direct consequence of an overparameterized model, or of non-informative data. For explicit models structures, the origin of non-identifiability may be distinguished a priori by calculating structural information and Fisher information matrices [1]. [30] however points out that it may be impossible for complex models to recognize overparameterization a priori.

Our focus lies on a posteriori detection of parameter correlations on the basis of two-dimensional likelihood regions. The weighted sum of squared residuals measures the agreement of experimental data with model predictions, given a parameter set θ :

$$r^2(\theta) = \sum_{k=1}^m \sum_{l=1}^d \left(\frac{y_{kl}^D - y_k(\theta, t_l)}{\sigma_{kl}^D} \right)^2 \quad (10)$$

where y_{kl}^D denotes d data points for each observable k (relative humidity and weight), measured at time-points

t_l (we followed symbols chosen by [31]). σ_{kl}^D are the corresponding measurement errors. The r^2 indicator is related to the likelihood L by:

$$L(\theta) \propto \exp \left[-\frac{1}{2} r^2(\theta) \right] \quad (11)$$

Therefore, the maximum likelihood estimate (MLE) $\hat{\theta}$ minimizes r^2 over the search space, under the assumption of iid measurement noise of standard deviation σ^D . Following the definition of [26], [31] translate the r indicator into two-dimensional likelihood-based confidence regions using a threshold in the likelihood. An approximate $100(1 - \alpha)\%$ likelihood-based confidence region for θ is the set of all values such that

$$-2 \log \left[\frac{L(\theta)}{L(\hat{\theta})} \right] = r^2(\theta) - r^2(\hat{\theta}) < \chi_{1-\alpha, k}^2 \quad (12)$$

where $\chi_{1-\alpha, k}^2$ is the $1 - \alpha$ quantile of the chi-square distribution with k degrees of freedom, k being the number of parameters [26]. As a consequence, there is a direct link between the value of r^2 of each sample of the posterior distribution, and confidence intervals.

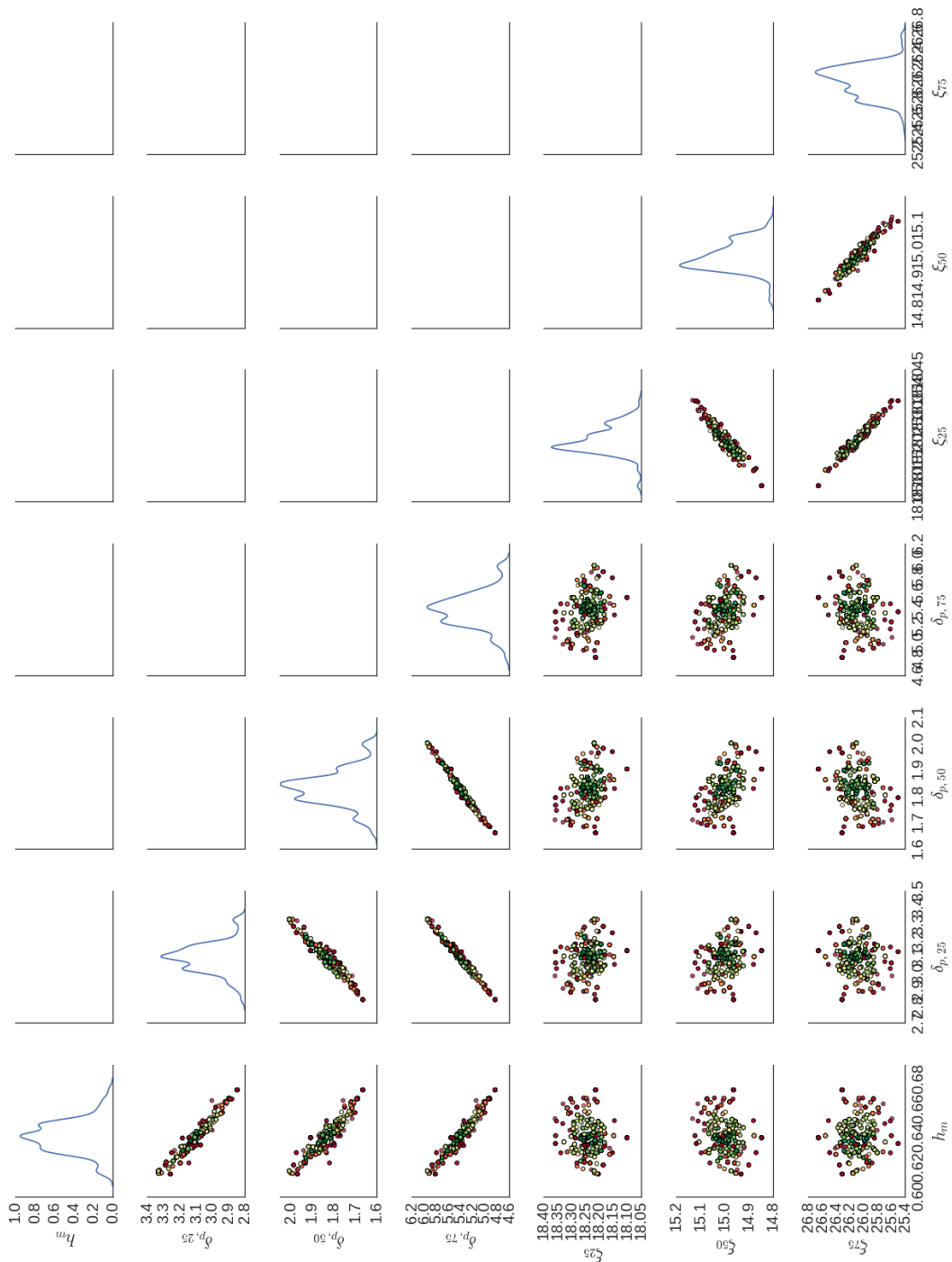


Figure 8: Pairwise posterior distributions. Color scale : $r^2(\theta) - r^2(\hat{\theta})$. Green to red means low to high deviation from MLE

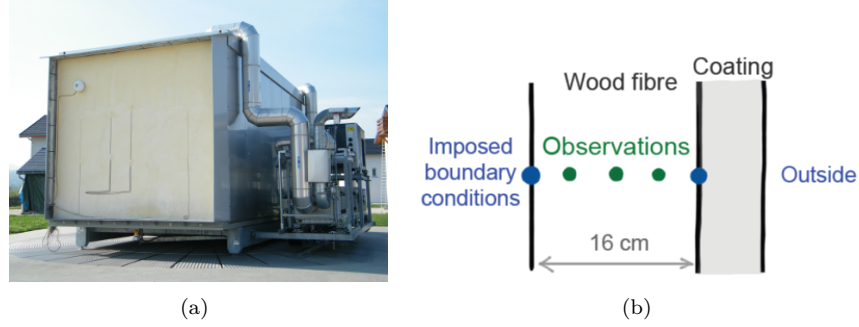


Figure 9: Experimental setup for the validation exercise: (a) PASSYS test cell and (b) sensor locations within the wall

Fig. 8 is the matrix of pairwise posterior distributions between each of the 7 parameters obtained from the sorption experiment. Each pairwise scatter plot accounts for a visual representation of the correlation between two parameters, with an additional information: the color scale indicates the value of $r^2(\theta) - r^2(\hat{\theta})$, i.e. log-likelihood of each sample relatively to the MLE. The first observation from Fig. 8 is the presence of two groups of correlated parameters: first, h_m is negatively correlated with each component of the permeability $\delta_{p,25\%}$, $\delta_{p,50\%}$ and $\delta_{p,75\%}$, which are positively correlated to each other; then, a strong correlation between each component of the moisture capacity $\xi_{25\%}$, $\xi_{50\%}$ and $\xi_{75\%}$ can be seen as well.

Our interpretation of these results is the following:

- Although h_m interacts with the parameters of moisture permeability, it is most likely not an important constraint on the identification procedure. Some authors [28] found that the external mass transfer coefficient has a negligible effect on the identified value of moisture diffusivity. Moreover, likelihood-based confidence intervals between h_m and δ_p seem finite.
- The apparent moisture permeability δ_p seems overparameterized, or the available data does not seem sufficient to estimate a profile of permeability described by three variables. The results shown on Fig. 6 should therefore be seen with caution.
- The interaction between parameters of the moisture capacity ξ is not problematic, because of how well the sorption isotherm fits reference data (see Fig. 7)

4 Validation

A choice of material description was made in Eq. 8 and these parameters have been identified from the experiments and the MCMC algorithm, along with their confidence intervals. We now wish to validate these results with a second experimental setup, accounting for parameter value uncertainty.

4.1 Experimental setup

The experimental validation of the identification procedure is attempted on the basis of measurements previously gathered in the frame of the HYGROBAT project [39]. This project aimed at establishing tools to precisely quantify the impact of mass transfer on heat transfer in highly hygroscopic materials. One of the implemented experimental facilities were PASSYS test cells shown on Fig. 9(a).

An instrumented wall is exposed on one side to controlled boundary conditions, and to the exterior weather on the other side. The wall is made of a 16 cm layer of wood fibre board, which is protected from solar radiation and wind-driven rain by a permeable coating. Temperature and RH sensors were evenly distributed within the wood fibre layer, every 4 cm (Fig. 9(b)). These sensors provide observation data which

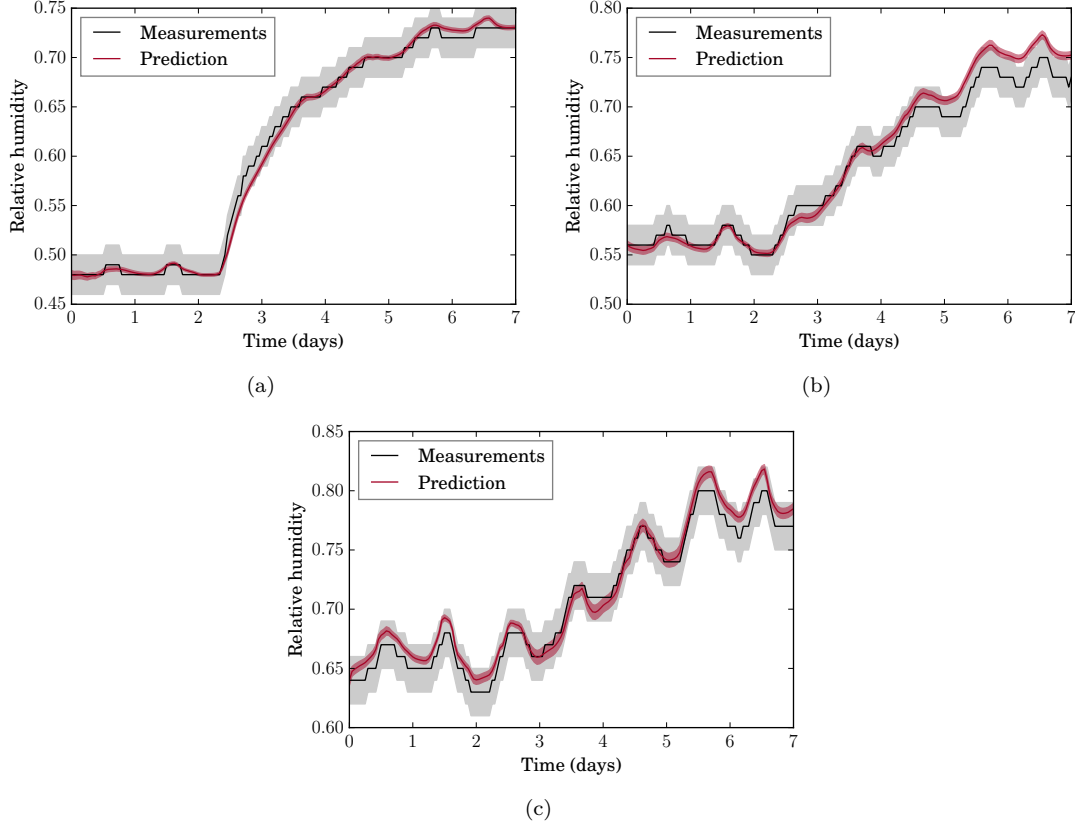


Figure 10: Validation results: (a) $x = 4\text{cm}$, (b) $x = 8\text{cm}$, (c) $x = 12\text{cm}$

will be used for comparison with model outputs. More details on the experimental design are available in [33]. The conditions on each side of the instrumented layer are recorded by sensors placed at the material surface: the boundary conditions of the simulation are prescribed temperature and vapour pressure. Eq. 2 is thus replaced by a Dirichlet condition in the following simulations.

A measurement sequence of 7 days is used in this validation phase. The interior temperature is constant at 20°C and the relative humidity follows a 50%-75% step after two days of equilibrium. The exterior boundary conditions used in the simulations are given by the temperature and relative humidity measured at the interface between the wood fibre board and the external coating. Only the wood fibre layer is simulated.

4.2 Results

Simulations are a FEM implementation of Eq. 1 and Eq. 3, where coupling effects of heat and moisture transfer are significant. The material is described by the identified set of parameters and their uncertainty: instead of a single simulation, it is repeated 100 times by drawing parameter samples from the posterior distribution. The deviation between each simulation defines a range of uncertainty for the relative humidity predictions at the locations of the sensors within the wall: $x = 4\text{cm}$, $x = 8\text{cm}$ and $x = 12\text{cm}$. Parameter values and uncertainties were averaged between the outcome of the sorption and desorption identification experiments. This does not account for a real hysteresis model for the sorption isotherm, but was considered adequate in the present case. Results are shown on Fig. 10

The grey area around RH observations depicts a measurement uncertainty of 2%RH, and the area around the mean prediction profiles are the 95% confidence intervals for RH calculations. The agreement between

measurements and predictions can be seen to be very satisfactory. It is noteworthy that this validation exercise has a radically different setup from the first experiment used for parameter characterization. A better agreement could likely still be reached by including a sorption isotherm hysteresis model, which would better capture desorption phases.

5 Conclusion

The present work showed the hygric characterization of wood fibre insulation boards. Material samples were placed in a controlled environment and submitted to stepwise RH variations. Dynamic measurements of their weight and internal humidity were then interpreted by a MCMC algorithm to estimate their moisture permeability and sorption isotherm. The reliability of identification results was then evaluated by a posteriori parameter correlations, which may indicate the relevance of the chosen parameterization. The results of this parameter estimation phase were then validated against observations from a second experiment, in an entirely different setup. This validation phase confirmed that characterization results could be successfully extrapolated to another experiment.

- It is not necessary for the model structure to have a differentiable analytical formulation. The MCMC algorithm is derivative-free and returns a description of the unknown parameters in the form of posterior probability distributions.
- It is not necessary for material samples to reach mass equilibrium between each RH step of the experimental process. A dynamic set of boundary conditions with relatively fast fluctuations holds enough information for the inference of the sorption isotherm, if both sample weight and internal RH are recorded dynamically. This allows for a significant time reduction of the experiments compared to standard steady-state methods.
- A single RH step is not informative enough for a precise inference of moisture-dependent properties such as vapour permeability and moisture capacity: the two-step sorption experiment yields results with a noticeably lower uncertainty than the one-step desorption experiment.
- The examination of interactions between parameters can easily be done a posteriori, in order to qualitatively estimate the reliability of identification results. This is an important step if the model is likely to be overparameterized.

The first possible application of these results is the development of a new protocol for fast and exhaustive numerically assisted characterisation, improving traditional methods at the laboratory scale. In this particular case, the experimental observation period could be reduced to a month for a joint identification of moisture permeability and sorption isotherm. This time can most likely be reduced further, either by only harnessing the beginning of each observation step, or by designing the experiment (RH values and duration) to maximize the information gained by inference: this is the purpose of Optimal Experiment Design, which aims at improving identification accuracy within a shortened time frame.

The second possible application is at the building scale: a local estimation of the hygrothermal features of a building envelope may allow a better diagnosis in the prospects of building retrofitting. An example of this was done by [6] in a problem of one-dimensional heat transfer. However, a necessary condition for a successful parameter estimation is that the model used for inference accurately represents the complexity of the physical system, and sensor information is sufficient to capture it. The in-situ characterisation of real building components featuring thermal bridges and effects of 2D or 3D transfer is therefore a challenge.

6 Acknowledgements

This work uses experimental records that were gathered in the scope of the HYGROBAT project, and data processing algorithms developed within the BAYREB project, both funded by the French National Research Agency (ANR). The authors would like to thank the ANR for supporting these projects.

References

- [1] C. Agbi, Z. Song, and B. Krogh. Parameter identifiability for multi-zone building models. In *2012 IEEE 51st Annual Conference on Decision and Control (CDC)*, pages 6951–6956, 2012.
- [2] Oleg M. Alifanov. *Inverse Heat Transfer Problems*. Springer Science & Business Media, December 2012.
- [3] Thierry Bastogne. Experimental Modeling of Dynamical Systems - Applications in Systems Biology. Habilitation à diriger des recherches, Université Henri Poincaré - Nancy I, 2008.
- [4] James V. Beck. *Inverse Heat Conduction: Ill-Posed Problems*. James Beck, October 1985.
- [5] Abdelmalek Bensefia, Mohammed Boussaid, and Tahar Loulou. Optimal Heat Input for Estimating Luikov’s Parameters in a Heat and Mass Transfer Problem. *Numerical Heat Transfer, Part B: Fundamentals*, 60(5):399–423, November 2011.
- [6] Julien Berger, Helcio R. B. Orlande, Nathan Mendes, and Sihem Guernouti. Bayesian inference for estimating thermal properties of a historic building wall. *Building and Environment*, 2016.
- [7] Phillip Biddulph, Virginia Gori, Clifford A. Elwell, Cameron Scott, Caroline Rye, Robert Lowe, and Tadj Oreszczyn. Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements. *Energy and Buildings*, 78:10–16, 2014.
- [8] Jordan Brouns, Alexandre Nassiopoulou, Frédéric Bourquin, and Karim Limam. State-parameter identification for accurate building energy audits. In *Building Simulation 2013 conference*, page 7, France, January 2013.
- [9] Roland Brun, Peter Reichert, and Hans R. Künsch. Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37(4):1015–1030, 2001.
- [10] Thomas Busser, Amandine Piot, Mickael Pailha, Timea Bejat, and Monika Woloszyn. From material properties to modelling hygrothermal transfers of highly hygroscopic walls. In *Central European Symposium on Building Physics*, Dresden.
- [11] Jie Cai and James Braun. An inverse hygrothermal model for multi-zone buildings. *Journal of Building Performance Simulation*, 0(0):1–19, November 2015.
- [12] Jie Cai, Donghun Kim, James E. Braun, and Jianghai Hu. Optimizing Zone Temperature Setpoint Excitation to Minimize Training Data for Data-driven Dynamic Building Models. In *American control conference*, pages 6–8, Boston, July 2016.
- [13] Samuel Dubois, Fionn McGregor, Arnaud Evrard, Andrew Heath, and Frédéric Lebeau. An inverse modelling approach to estimate the hygric parameters of clay-based masonry during a Moisture Buffer Value test. *Building and Environment*, 81:192–203, November 2014.
- [14] Thierry Duforestel. *Des transferts couplés de masse et de chaleur à la conception bioclimatique: recherches sur l’efficacité énergétique des bâtiments*. HDR Thesis, Université Claude Bernard Lyon 1, 2015.
- [15] Valerii Fedorov. Optimal experimental design. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(5):581–589, September 2010.
- [16] John Geweke. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Research Department Staff Report 148, Federal Reserve Bank of Minneapolis, 1991.
- [17] Peter J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, December 1995.

- [18] Heikki Haario, Eero Saksman, and Johanna Tamminen. An Adaptive Metropolis Algorithm. *Bernoulli*, 7(2):223–242, 2001.
- [19] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, April 1970.
- [20] Y. Heo, R. Choudhary, and G. A. Augenbroe. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, 47:550–560, 2012.
- [21] Jari P. Kaipio and Colin Fox. The Bayesian Framework for Inverse Problems in Heat Transfer. *Heat Transfer Engineering*, 32(9):718–753, 2011.
- [22] Rick Kramer, Jos van Schijndel, and Henk Schellen. Inverse modeling of simplified hygrothermal building models to predict and characterize indoor climates. *Building and Environment*, 68:87–99, 2013.
- [23] Hartwig M. Künzel and Kurt Kiessl. Calculation of heat and moisture transfer in exposed building components. *International Journal of Heat and Mass Transfer*, 40(1):159–167, October 1996.
- [24] Denis Maillet, Yvon Jarny, and Daniel Petit. Problèmes inverses en diffusion thermique: Formulation et résolution du problème des moindres carrés. *Techniques de l’ingénieur Transferts thermiques*, base documentaire : TIB214DUO(ref. article : be8266), 2014.
- [25] Denis Maillet, Yvon Jarny, and Daniel Petit. Problèmes inverses en diffusion thermique: Outils spécifiques de conduction inverse et de régularisation. *Techniques de l’ingénieur Transferts thermiques*, base documentaire : TIB214DUO(ref. article : be8267), 2014.
- [26] William Q. Meeker and Luis A. Escobar. Teaching about Approximate Confidence Regions Based on Maximum Likelihood Estimation. *The American Statistician*, 49(1):48–53, 1995.
- [27] Johan A. A. Nylander, James C. Wilgenbusch, Dan L. Warren, and David L. Swofford. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, 24(4):581–583, February 2008.
- [28] Patrick Perré, Floran Pierre, JoÃl Casalinho, and Mehdi Ayouz. Determination of the Mass Diffusion Coefficient Based on the Relative Humidity Measured at the Back Face of the Sample During Unsteady Regimes. *Drying Technology*, 33(9):1068–1075, 2015.
- [29] J. R. Philip and D. A. De Vries. Moisture movement in porous materials under temperature gradients. *Eos, Transactions American Geophysical Union*, 38(2):222–232, 1957.
- [30] Bruce Rannala. Identifiability of Parameters in MCMC Bayesian Inference of Phylogeny. *Systematic Biology*, 51(5):754–760, September 2002.
- [31] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. KlingmÃijller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, August 2009.
- [32] Simon Rouchier. hamopy: Heat, Air and Moisture transfer in Python, June 2015. DOI: 10.5281/zenodo.18399.
- [33] Simon Rouchier, Monika Woloszyn, Yannick Kedowide, and Timea Béjat. Identification of the hygrothermal properties of a building envelope material by the covariance matrix adaptation evolution strategy. *Journal of Building Performance Simulation*, 0(0):1–14, 2015.
- [34] Andrew M. Stuart. The Bayesian Approach To Inverse Problems. *arXiv:1302.6989 [math]*, February 2013.
- [35] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.

- [36] Jorn F. M. Van Doren, Paul M. J. Van den Hof, Jan Dirk Jansen, and Okko H. Bosgra. Determining Identifiable Parameterizations for Large-scale Physical Models in Reservoir Engineering. *IFAC Proceedings Volumes*, 41(2):11421–11426, 2008.
- [37] Eric Walter and Luc Pronzato. Identification of parametric models. *Communications and Control Engineering*, 8, 1997.
- [38] Jingbo Wang and Nicholas Zabaras. A Bayesian inference approach to the inverse heat conduction problem. *International Journal of Heat and Mass Transfer*, 47(17-18):3927–3941, 2004.
- [39] Monika Woloszyn. Vers une méthode de conception HYGRO-thermique des BATiments performants: démarche du projet HYGRO-BAT. In *Conférence IBPSA France*, Arras, 2014.
- [40] Chadia Zayane. *Identification d’un modèle de comportement thermique de bâtiment à partir de sa courbe de charge*. PhD thesis, École Nationale Supérieure des Mines de Paris, January 2011.