



**HAL**  
open science

## Bayesian calibration and sensitivity analysis of heat transfer models for fire insulation panels

Paul-Remo Wagner, P.-R Wagner, Reto Fahrni, Michael Klippel, Andrea Frangi, B. Sudret

► **To cite this version:**

Paul-Remo Wagner, P.-R Wagner, Reto Fahrni, Michael Klippel, Andrea Frangi, et al.. Bayesian calibration and sensitivity analysis of heat transfer models for fire insulation panels. *Engineering Structures*, 2020, 205, 10.1016/j.engstruct.2019.110063 . hal-02290027v2

**HAL Id: hal-02290027**

**<https://hal.science/hal-02290027v2>**

Submitted on 13 Nov 2019 (v2), last revised 7 Jan 2020 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BAYESIAN CALIBRATION AND SENSITIVITY ANALYSIS OF HEAT TRANSFER MODELS FOR FIRE INSULATION PANELS

P.-R. Wagner, R. Fahrni, M. Klippel, A. Frangi, B. Sudret



## Data Sheet

---

**Journal:**

**Report Ref.:** RSUQ-2019-002

**Arxiv Ref.:** <http://arxiv.org/abs/1909.07060> - [stat] [stat.CO]

**DOI:**

**Date submitted:** November 07, 2019

**Date accepted:** -

---

# Bayesian calibration and sensitivity analysis of heat transfer models for fire insulation panels

P.-R. Wagner, R. Fahrni, M. Klippel, A. Frangi, B. Sudret

07.11.2019

## Abstract

A common approach to assess the performance of fire insulation panels is the component additive method (CAM). The parameters of the CAM are based on the temperature-dependent thermal material properties of the panels. These material properties can be derived by calibrating finite element heat transfer models using experimentally measured temperature records. In the past, the calibration of the material properties was done manually by trial and error approaches, which was inefficient and prone to error. In this contribution, the calibration problem is reformulated in a probabilistic setting and solved using the Bayesian model calibration framework. This not only gives a set of best-fit parameters but also confidence bounds on the latter. To make this framework feasible, the procedure is accelerated through the use of advanced surrogate modelling techniques: polynomial chaos expansions combined with principal component analysis. This surrogate modelling technique additionally allows one to conduct a variance-based sensitivity analysis at no additional cost by giving access to the Sobol' indices. The calibration is finally validated by using the calibrated material properties to predict the temperature development in different experimental setups.

**Keywords:** Bayesian model calibration, sensitivity analysis, surrogate modelling, component additive method, polynomial chaos expansions.

## 1 Introduction

Knowledge about the basic behaviour of materials exposed to fire is extremely important to successfully develop fire safety strategies. Depending on the type and height of buildings, certain fire requirements need to be fulfilled, e.g. requirements w.r.t. the *load-bearing function* (R) and *separating function* (EI). In case of timber buildings, the performance not only of the timber members, but also of protective materials such as gypsum plasterboards and



insulations is of high importance for the fire design of the building structure. These different materials are usually combined to build floor and wall elements with different layups, so-called *timber frame assemblies*. The separating function of timber frame assemblies is usually verified using the component additive method (CAM) (Frangi et al., 2010; Mäger et al., 2017; Just and Schmid, 2018). This method is rather flexible for calculating the separating function because it handles arbitrary layups made of various materials and thickness. Producers of fire protection products (e.g. gypsum plasterboard and insulation) need to determine input factors for the individual materials so that the separating function of a timber frame assembly with these materials can be verified using the CAM.

Indeed, the same protective material (e.g. a gypsum plasterboard with a given thickness) contributes differently to the fire resistance of a timber frame assembly in different setups. The CAM therefore considers (1) the material and thickness of a layer and (2) modification factors taking into account the neighbouring layers. This leads to a high number of possible combinations for timber frame assemblies, which cannot all be tested in fire resistance tests. Therefore, the factors of the CAM are usually derived based on finite element (FE) models and accompanying fire resistance tests.

Fire resistance tests using the standard EN/ISO temperature time curve according to ISO 834-1:1999 (1999) and EN 1363-1:2012 (2012) constitute the basis for these simulations. In the fire tests, the temperature is recorded over time inside the specimen at specific distances to the fire exposed surface. These recordings are used as a reference for FE simulations of the same setup. Heat transfer models using *effective thermal material properties* that depend on the temperature  $T$  (specific heat capacity  $c(T)$ , thermal conductivity  $\lambda(T)$  and material density  $\rho(T)$ ) are usually employed to simulate the temperature development inside timber frame assemblies exposed to fire. These material properties are then calibrated such that the output matches the recorded temperatures. Since these properties are not strictly physical quantities, they are called *effective* material properties. They account for not explicitly modeled effects such as fissures, cracks and moisture flow inside the specimen (Frangi et al., 2010). Despite these simplifications, using temperature-dependent effective material properties together with a common heat transfer analysis is appropriate and state-of-the-art for the calibration of parameters in the CAM. This is especially true since the introduced simplifications are negligible compared to the simplifications made within the CAM.

The conventional process of determining these effective material properties is slow and inaccurate as the calibration is usually done manually. All temperature measurements are averaged, thus eliminating the variability in the material behaviour and not accounting for it in the calibration. The derivation of thermal material properties is conventionally done as follows (Mäger et al., 2017):

**Step 1** Assume effective thermal material properties and simulate a specific layup with FE heat transfer models;

**Step 2** Compare the resulting temperatures with the averaged measurements;

**Step 3** Iterate **Step 1** and **Step 2** until the simulation results are similar to the measured temperatures.

A more rigorous calibration of these effective material properties can be achieved by parameterizing the thermal time-dependent material properties with a set of model parameters. Through this parametrization the problem of determining the time-dependent effective material properties is recast as a problem of determining the real-valued parameterizing model parameters. Then the calibration problem can be posed in a probabilistic setting. This allows a proper treatment of uncertainties arising from material fluctuations, measurement errors and model insufficiencies. One general way to do this is the so-called *Bayesian inversion* framework (Beck and Katafygiotis, 1998; Gelman et al., 2014; Yu et al., 2019). In this framework, the model parameters are seen as random variables. Instead of trying to determine one particular value for these parameters, this probabilistic framework determines the full probability distribution of the model parameters conditioned on the observed measurements. This distribution contains much more information about the calibrated properties than the single point estimate from the conventional approach. For example, it allows computing expected values, maximum a posteriori estimates, confidence intervals on the calibrated values and the full correlation structure. Furthermore, the calibration can be verified easily by computing the *posterior predictive distribution* (Gelman et al., 1996).

To determine the probability distribution of the material properties, it is necessary to repeatedly evaluate the FE heat transfer model. To reduce the computational burden associated with repeated model evaluations, it has become customary to replace the computational forward model with a cheap-to-evaluate surrogate. Therefore, the Bayesian inversion framework is here combined with the polynomial chaos expansions (PCE) surrogate modelling technique (Sudret, 2007; Blatman, 2009; Guo et al., 2018).

When working with models with multiple input parameters, the question of the *relative importance* of individual parameters with respect to the output arises naturally. Quantifying this influence is called sensitivity analysis. One family of approaches are the so-called *variance decomposition techniques* (Saltelli et al., 2000; Arwade et al., 2010). These methods attempt to apportion the variance of the probabilistic model output to the individual input parameters. The *Sobol' indices* are one such variance decomposition technique (Sobol', 1993). Determining the Sobol' indices is typically computationally expensive, but it has been shown by Sudret (2006) that they can be computed easily for a PCE surrogate model. Their computation allows valuable insights into the heat-transfer model's properties.

In this paper, the material properties of four different gypsum insulation boards (Products A-D, E1-E4) are calibrated based on fire resistance tests carried out with these materials. These experimental results are presented and discussed in detail in Section 2. The calibration is carried out with the Bayesian model calibration framework accelerated by constructing a

PCE-based surrogate model as detailed in Section 3. Section 4 outlines how the employed surrogate model can be used to conduct a global sensitivity analysis of the considered computational model. Finally, in Section 5, the calibration is verified using two fire tests that use two of the calibrated materials (Product C and Product D) in different experimental setups (V1 and V2).

## 2 Experiments and modelling

### 2.1 Experiments

Two fire resistance tests with horizontally oriented specimens constitute the experimental basis for the analysis in this paper (Just, 2016; Breu, 2016). The unloaded tests were conducted on the model-scale furnace of SP Wood Building Technology (today's Research Institute of Sweden, RISE) and were exposed to the EN/ISO temperature-time curve (EN 1363-1:2012, 2012; ISO 834-1:1999, 1999). Two (V1 and V2, (Breu, 2016)), respectively four (E1 to E4, (Just, 2016)) different gypsum plasterboard setups with dimensions of  $0.4 \times 0.4$  m were tested in each test (Figure 1). Tables 1 and 2 show the layups of the specimens. The space between and around the specimens was at least 100 mm and was filled with Product D boards to protect the carrying layer, *i.e.* the last layer. The carrying layer was a 19 mm particle board with density  $\rho = 633$  kg/m<sup>3</sup> in Test 1 (specimens E1 to E4) and a Product D 15 mm in Test 2 (specimens V1 and V2). Figure 2 shows the specimens of Test 2 during fabrication.

In specimens E1 to E4 (Test 1), five wire thermocouples and one copper disc thermocouple measured the temperatures at a single *interface* between the layers. The fluctuations of the sensor readings can mainly be attributed to variations in the material properties. The measured temperatures are displayed in Figures 3(a) to 3(d).

In specimens V1 and V2 (Test 2), three wire thermocouples were placed between each layer. The measured temperatures at the two *interfaces 1* and *2* are displayed in Figures 3(e) and 3(f).

Test 1 and Test 2, with specimens E1-E4 and V1,V2 respectively, were exposed to the EN/ISO standard temperature-time curve (EN 1363-1:2012, 2012; ISO 834-1:1999, 1999).

The temperature measurements  $y^{(s)}(t)$  at each sensor  $s$  were collected at  $N$  discrete time steps  $t_i$ :

$$\mathbf{y}^{(s)} = (y_1^{(s)}, \dots, y_N^{(s)})^\top \quad \text{with} \quad y_i^{(s)} \stackrel{\text{def}}{=} y^{(s)}(t_i) \quad \text{for} \quad i = 1, \dots, N, \quad (1)$$

where  $t_i = i\tau$  and  $\tau = 10$  s. For simplicity, the superscript ( $s$ ) is omitted unless required to distinguish between individual measurements. Therefore, in the sequel  $\mathbf{y}$  stands for a vector of measurements captured by a single sensor.

Table 1: Specimens E1-E4 (Test 1)

		Layer 1		Layer 2	
E1	Product A	12.5 mm	particle board	19 mm	
E2	Product B	9.5 mm	particle board	19 mm	
E3	Product C	12.5 mm	particle board	19 mm	
E4	Product D	15 mm	particle board	19 mm	

Table 2: Specimens V1 and V2 (Test 2)

		Layer 1		Layer 2+3	
V1	Product C	12.5 mm	Product D	2×15 mm	
V2	Product D	15 mm	Product D	2×15 mm	

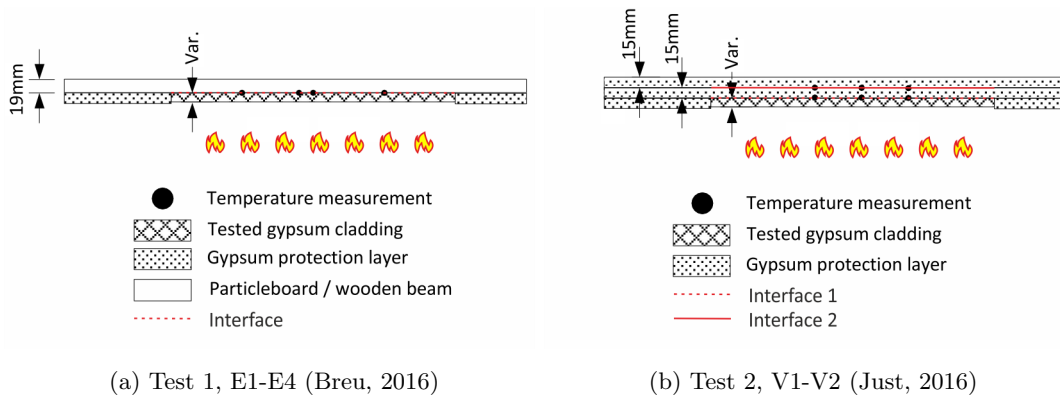


Figure 1: Sketch of the experimental setups from Test 1 and Test 2. For more details refer to the respective publications Breu (2016) and Just (2016).

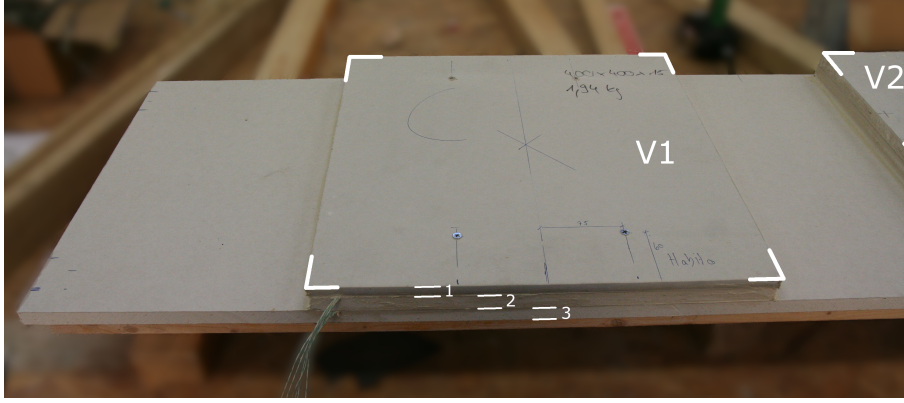


Figure 2: Specimens V1 and V2 (on the right), upside-down, exposed surface on top, before installation of the protection layers around/between the specimens, three wire thermocouples between each layer;

1: exposed protection layer, 12.5 mm Product C; 2: protection layer 15 mm Product D; 3: carrying layer 15 mm Product D; around the specimen other protection layers were applied to protect the carrying layer.

## 2.2 Forward modelling

The experiments described in the previous section were modeled using one-dimensional heat transfer FE-models. The reduction to a one-dimensional setup is justified, because it is known that the experimental heat-flux is mostly perpendicular to the exposed surface. The simulations were conducted using the general purpose finite element software *Abaqus* (Abaqus FEA, 2017). The energy input on the heated surface and the losses on the unexposed side took into account the energy input/loss through convection and radiation.

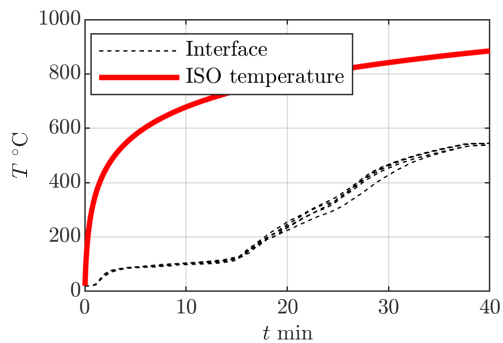
The radiation temperature was assumed to be equal to the gas temperature and followed the EN/ISO temperature-time curve EN 1363-1:2012 (2012); ISO 834-1:1999 (1999) on the exposed side and was constantly  $19.5\text{ }^{\circ}\text{C}$  on the unexposed side (as in the experiments). The emissivity was taken as 0.8 and the convection coefficient as  $25\text{ W/m}^2\text{K}$  according to EN 1991-1-2:2002 (2002). The element size in the FE-mesh was 0.25 mm. The temperature-dependent material properties of the particle board (Test 1) were taken from Schleifer (2009).

The unknown parameters of interest are the *temperature-dependent effective material properties* of the insulation material: the thermal conductivity  $\lambda(T)$ , the heat capacity  $c(T)$  and the material density  $\rho(T)$  of the investigated insulation materials.

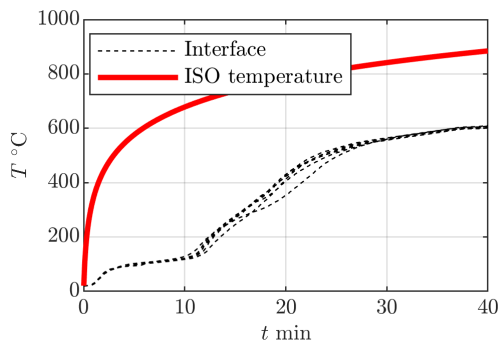
With this, the computational forward model is

$$\mathbf{Y} = (Y_1, \dots, Y_N)^T = \mathcal{M}(\lambda(T), c(T), \rho(T)), \quad (2)$$

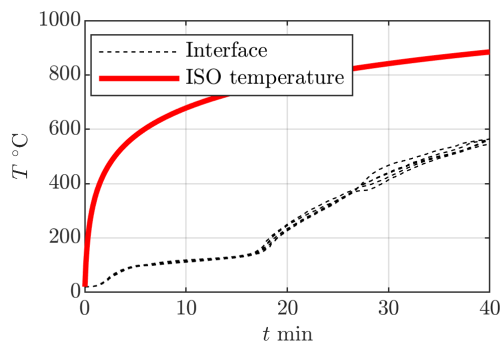
For every set of material properties, this model returns the temperature evolution at locations and at times where measurements are available (see Section 2.1).



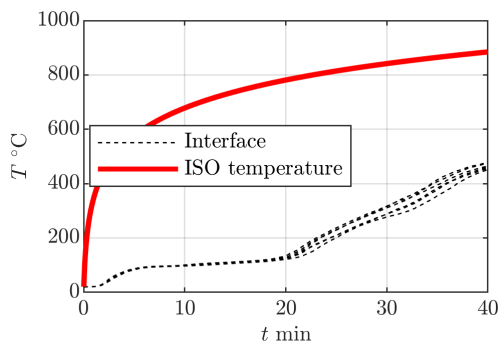
(a) E1 (Just, 2016)



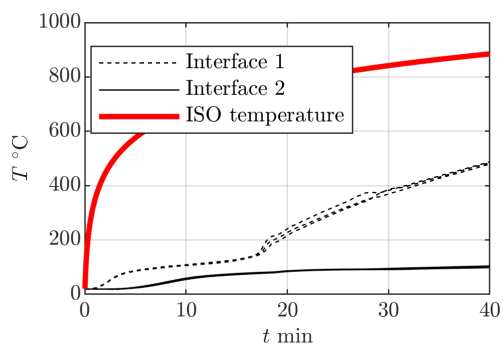
(b) E2 (Just, 2016)



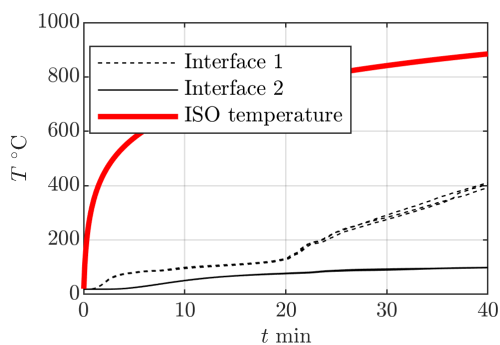
(c) E3 (Just, 2016)



(d) E4 (Just, 2016)



(e) V1 (Breu, 2016)



(f) V2 (Breu, 2016)

Figure 3: Summary of the data used for calibration and the underlying ISO temperature according to EN 1363-1:2012 (2012); ISO 834-1:1999 (1999).

Since each of those parameters is a function of the temperature, they cannot be directly calibrated. Instead, these functions have to be parameterized with a set of scalar parameters, as described next.

### 2.3 Parametrization of material properties

The parametrization of the three temperature-dependent material properties ( $\lambda(T), c(T), \rho(T)$ ) is a crucial step of the calibration procedure. It consists of specifying a set of parameters  $\mathbf{X}_{\mathcal{M}}$  that define the shape of the temperature-dependent function that describes each material property. The choice of these parameters is delicate, as it imposes a certain temperature-dependent behaviour on the material properties. A priori, there are no physical constraints on this thermal behaviour besides positivity, so generally, the properties are defined as  $\lambda : [0, 1200^{\circ}\text{C}] \rightarrow \mathbb{R}^+$ ,  $c : [0, 1200^{\circ}\text{C}] \rightarrow \mathbb{R}^+$  and  $\rho : [0, 1200^{\circ}\text{C}] \rightarrow \mathbb{R}^+$ .

One further complication lies in the fact that these properties are mere *effective* properties and cannot generally be measured. To find constraints on these parameters, it is thus necessary to rely on previous calibration attempts of gypsum insulation boards (Breu, 2016; Schleifer, 2009) in conjunction with measurements of certain properties, where available.

By gathering information from such previous attempts, the thermal properties are parameterized by six parameters  $\mathbf{X}_{\mathcal{M}} = (X_1, \dots, X_6)^{\top}$ . This parametrization is flexible enough to enable inference on  $\mathbf{X}_{\mathcal{M}}$  and follows physical and empirical reasoning as described next.

We propose to distinguish two *key processes* during which the temperature-dependent material properties change significantly:

**First key process** When the free water content in the gypsum insulation boards evaporates, the latent water content of gypsum, which is composed of sulfate dihydrate  $\text{CaSO}_4 \cdot 2 \text{H}_2\text{O}$ , evaporates. In this process, evaporation first forms calcium sulfate hemihydrate  $\text{CaSO}_4 \cdot 1/2 \text{H}_2\text{O}$  (also called bassanite) and then anhydrate III  $\text{CaSO}_4$ . The evaporation consumes heat, which is modelled as a local increase of the specific heat capacity, a reduction in the conductivity and a reduction in the material density.

**Second key process** Thermogravimetric analyses have shown a second peak in the specific heat due to chemical metamorphosis at elevated temperatures of secondary components found in the gypsum insulation boards (Schleifer, 2009). This second key process is modelled as an increase in the material conductivity, a peak in the specific heat and a further reduction of the material density.

The temperatures at which these two key processes occur cannot be equally well prescribed a priori. While the temperature of the water evaporation is well known to occur at approximately  $100^{\circ}\text{C}$  with its main effect taking place at  $140^{\circ}\text{C}$  until it tails off at  $180^{\circ}\text{C}$ , the second key process cannot be characterized this precisely. It is assumed that the second key process starts at the unknown temperature  $X_1$  and ends at  $850^{\circ}\text{C}$ . Additionally, the

relative location of its main effect between  $X_1$  and  $850^\circ\text{C}$  is parameterized with  $X_2$ . These two temperatures heavily influence the evolution of the thermal properties and are thus used as temperatures of change in all effective material properties.

In the present setting of heated gypsum boards, the initial conductivity at ambient temperature is assumed to be  $\lambda(20^\circ\text{C}) = 0.4 \text{ W/mK}$  (Breu, 2016). During the first key process, the conductivity is assumed to linearly decrease to a second value that is parameterized by  $X_3$ . Starting with the second key process the conductivity starts to increase linearly to another value that is parameterized by  $X_4$ , reached at the highest simulation temperature of  $1200^\circ\text{C}$ .

Phase changes require a significant amount of energy. To model the energy requirement associated with the evaporation of water trapped inside the insulation material, the specific heat  $c(T)$  is modelled with two piecewise linear spikes during both key processes, while being constant at  $c = 960 \text{ J/kgK}$  (Schleifer, 2009) for the other temperatures. The specific heat at the peaks is parameterized by  $X_5$  for the first process and  $X_6$  for the second one.

During the first and second key process, gaseous products are emitted (water and carbon dioxide respectively) and thus the density of the gypsum  $\rho(T)$  reduces. This density reduction was studied in Schleifer (2009) and the results are applied here directly. Starting from the density measured at room temperature  $\rho_0$ ,  $\rho(T)$  linearly reduces during the first key process to 82%. It then remains constant and linearly reduces further to 77% from the start of the second key process  $X_1$  to the main effect of the second key process. The parametrization of the material properties is visualized in Figure 4.

To finalize the parametrization, reasonable ranges are defined for all parameters. These ranges correspond to bounds on the parameters that are the results of prior calibration attempts along with expert judgement. These ranges are given along with a summary of the parameters in Table 3 with plots of the resulting temperature-dependent material properties in Figure 5.

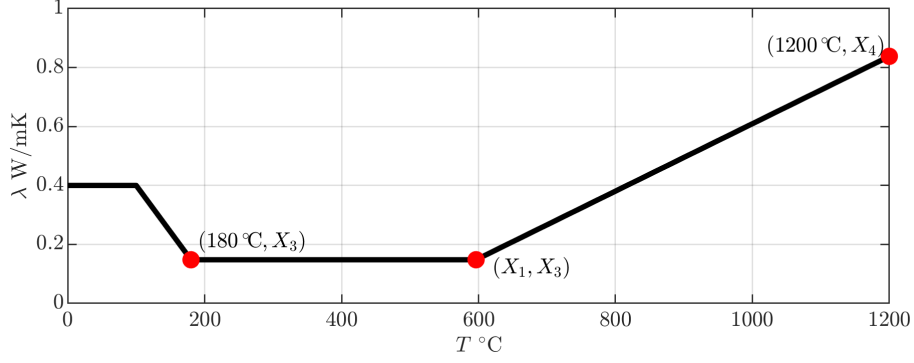
These six parameters are gathered into a vector  $\mathbf{X}_{\mathcal{M}} = (X_1, \dots, X_6)^\top$ , which fully characterizes the temperature-dependent behaviour of the gypsum insulation boards.

## 2.4 Finite element model

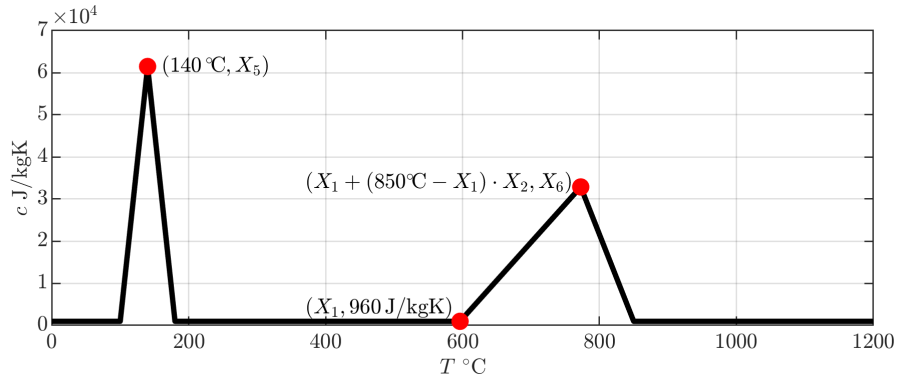
The FE model  $\mathcal{M}(\mathbf{X}_{\mathcal{M}})$  is considered as a verified simulator for the transient heat propagation in gypsum insulation panels under fire exposure. This means that the model is assumed to accurately solve the underlying differential equations posed by the mathematical heat transfer model. For a review of techniques for rigorous model verification see Oberkampf et al. (2004); Oberkampf and Roy (2010).

The FE model yields a discretized time-dependent temperature curve  $\mathbf{X}_{\mathcal{M}} \mapsto \mathbf{Y} = (Y_1, \dots, Y_N)^\top$  for each realization of the parameter vector  $\mathbf{X}_{\mathcal{M}} = (X_1, \dots, X_6)^\top$  that parameterizes the effective thermal properties  $\lambda(T, \mathbf{X}_{\mathcal{M}})$ ,  $c(T, \mathbf{X}_{\mathcal{M}})$  and  $\rho(T, \mathbf{X}_{\mathcal{M}})$ . The dis-

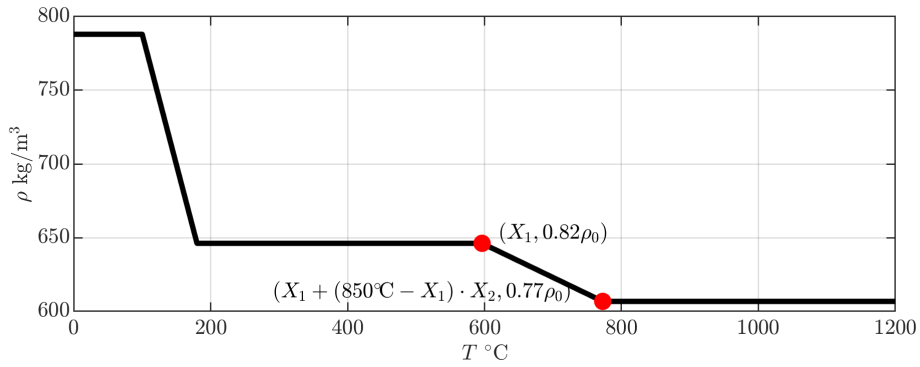




(a)  $\lambda(T, \mathbf{X})$

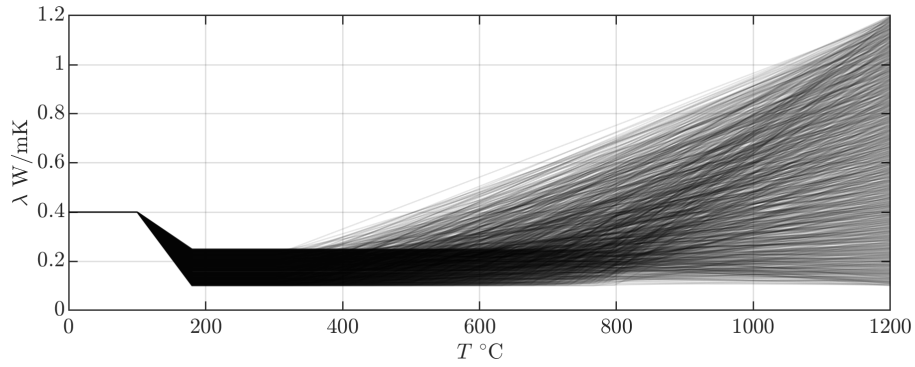


(b)  $c(T, \mathbf{X})$

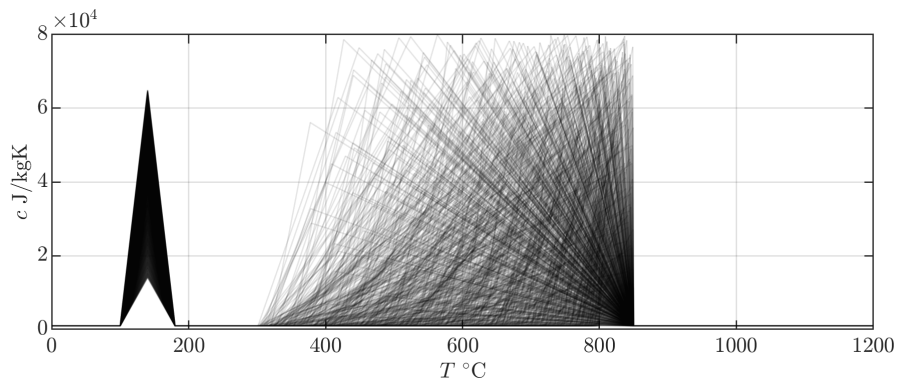


(c)  $\rho(T, \mathbf{X})$

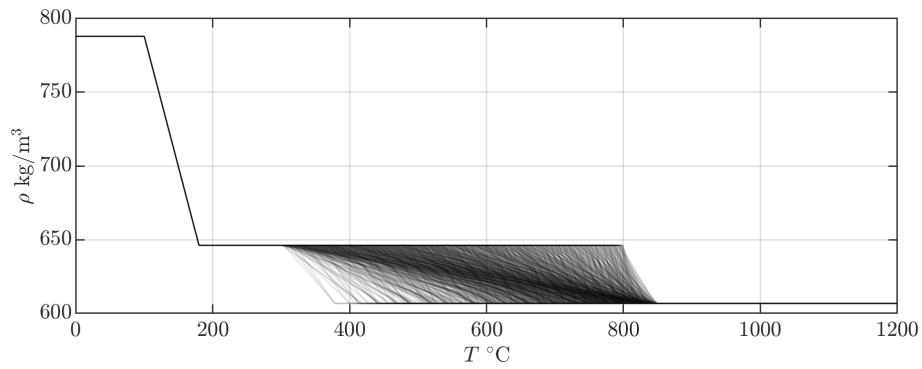
Figure 4: Parametrization of temperature-dependent effective material properties as defined in Table 3.



(a)  $\lambda(T, \mathbf{X})$



(b)  $c(T, \mathbf{X})$



(c)  $\rho(T, \mathbf{X})$

Figure 5: Realizations of temperature-dependent effective material properties in their respective ranges as defined in Table 3.

Table 3: Summary of the parameters  $(X_1, \dots, X_6)^\top$  that describe the material properties with their respective ranges.

Parameter	Physical Meaning	Range	Unit
$X_1$	Start of second key process	[300, 800]	$^\circ\text{C}$
$X_2$	Main effect of second key process (relative between $X_1$ and $850^\circ\text{C}$ )	[0.1, 1]	-
$X_3$	$\lambda(180^\circ\text{C})$ and $\lambda(X_1)$	[0.1, 0.25]	W/mK
$X_4$	$\lambda(1200^\circ\text{C})$	[0.1, 1.2]	W/mK
$X_5$	$c(140^\circ\text{C})$	$[1.4 \cdot 10^4, 6.5 \cdot 10^4]$	J/kgK
$X_6$	$c(X_1 + (850^\circ\text{C} - X_1) \cdot X_2)$	$[1 \cdot 10^3, 8 \cdot 10^4]$	J/kgK

cretization of the time steps is identical to the available measurements, so that  $t_i = i\tau$  with  $\tau = 10$  s.

### 3 Model calibration

The process of finding model parameters so that the model evaluation using this parameter vector agrees with some observations is called calibration. A general probabilistic framework for calibration is presented next. For simplicity, it is assumed that only one measurement series  $\mathbf{y} = \mathbf{y}^{(s)}$  is available for now. This restriction is lifted in Section 5.

#### 3.1 The Bayesian calibration approach

All models are simplifications of reality and all observations made in the real world contain measurement errors. To explicitly account for this combined mismatch between model output and observations, one option is to model the discrepancy as an *additive* mismatch between the model predictions and the observations:

$$\mathbf{y} = \mathcal{M}(\mathbf{X}_{\mathcal{M}}) + \mathbf{E}. \quad (3)$$

One way to address the calibration problem of determining  $\mathbf{X}_{\mathcal{M}}$  is to formulate it in a probabilistic setting. The unknown model discrepancy  $\mathbf{E}$  from Eq. (3) is then seen as a random vector. Commonly and in this work,  $\mathbf{E}$  is assumed to follow a zero mean normal distribution with a covariance matrix parameterized by a set of parameters  $\mathbf{X}_\varepsilon$ :

$$\mathbf{E} \sim \mathcal{N}(\varepsilon | \mathbf{0}, \boldsymbol{\Sigma}(\mathbf{x}_\varepsilon)). \quad (4)$$

Additionally, in the probabilistic setting, the combined parameter vector  $\mathbf{X} = (\mathbf{X}_{\mathcal{M}}, \mathbf{X}_\varepsilon) \in \mathcal{D}_{\mathbf{X}}$  is assumed to be distributed according to a so-called prior distribution (to be further

specified)

$$\mathbf{X} \sim \pi(\mathbf{x}). \quad (5)$$

Then the discrepancy distribution from Eq. (4) can be used together with Eq. (3) to construct a model giving the probability of the observations given a realization of the parameter vector. Denoting by  $\mathbf{x} = (\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon})$  a realization of  $\mathbf{X}$ , this probability as a function of the parameters is the so-called *likelihood function*

$$\mathcal{L}(\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon}; \mathbf{y}) = \mathcal{N}(\mathbf{y} | \mathcal{M}(\mathbf{x}_{\mathcal{M}}), \Sigma(\mathbf{x}_{\varepsilon})), \quad (6)$$

which reads more explicitly:

$$\mathcal{L}(\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon}; \mathbf{y}) = \frac{1}{(2\pi)^{N/2} \det \Sigma(\mathbf{x}_{\varepsilon})} \exp [(\mathcal{M}(\mathbf{x}_{\mathcal{M}}) - \mathbf{y})^{\top} \Sigma(\mathbf{x}_{\varepsilon})^{-1} (\mathcal{M}(\mathbf{x}_{\mathcal{M}}) - \mathbf{y})]. \quad (7)$$

With these definitions, it becomes possible to apply the Bayes' theorem for conditional probabilities (Gelman et al., 2014):

$$\pi(\mathbf{x} | \mathbf{y}) = \frac{\mathcal{L}(\mathbf{x}; \mathbf{y}) \pi(\mathbf{x})}{Z}, \quad \text{with} \quad Z = \int_{\mathcal{D}_{\mathbf{x}}} \mathcal{L}(\mathbf{x}; \mathbf{y}) \pi(\mathbf{x}) \, d\mathbf{x}, \quad (8)$$

where  $\pi(\mathbf{x})$  is the *prior distribution* of the input parameters  $\mathbf{X}$ ,  $\pi(\mathbf{x} | \mathbf{y})$  is the *posterior distribution* and  $Z$  is a normalizing factor called *evidence*.

The probability distributions  $\pi(\cdot)$  in this expression can be interpreted as degrees of belief about the parameters  $\mathbf{X}$  (Beck and Katafygiotis, 1998). Low values of the distribution at a realization  $\mathbf{X}$  indicate low confidence in this particular value, whereas high values indicate high confidence. With this interpretation of probabilities, Eq. (8) encodes the shift of belief about the parameter vector from *before*  $\mathbf{X} \sim \pi(\mathbf{x})$  to *after* the observation of experiments  $\mathbf{X} | \mathbf{y} \sim \pi(\mathbf{x} | \mathbf{y})$ . This process is called *Bayesian updating*, *Bayesian inference* or *Bayesian inversion*.

As mentioned above, the probability distributions in Bayes' theorem are named according to their information content about the parameters  $\mathbf{X}$  in the setting of the updating procedure:

**Prior distribution  $\pi(\mathbf{x})$ :** this distribution captures the belief about the parameters before (*i.e. prior* to) observing data. In the setting of Bayesian updating for model calibration, it is chosen according to expert opinion and possible prior calibration attempts. A typical choice is to select a reasonable, although sufficiently large range (lower/upper bounds) for each parameter.

**Posterior distribution  $\pi(\mathbf{x} | \mathbf{y})$ :** the posterior distribution is the conditional distribution of the parameters given the observations. It can be regarded as the state of information about the parameters  $\mathbf{X} | \mathbf{y}$  after (*i.e. posterior* to) making observations.

Thus, the computation of the posterior distribution  $\pi(\mathbf{x} | \mathbf{y})$  can be considered as the solution of the calibration problem. Since it is a probability distribution rather than a single value, it encompasses all information specified by the prior distribution and the newly

observed data. Already conceptually it is thus a much broader way of defining calibration than single value estimators.

Another probability distribution of interest in the Bayesian inference setting is the *posterior predictive distribution*  $\pi(\mathbf{y}^*|\mathbf{y})$ . It is defined as

$$\pi(\mathbf{y}^*|\mathbf{y}) = \int \mathcal{L}(\mathbf{x}; \mathbf{y}^*)\pi(\mathbf{x}|\mathbf{y}) \, d\mathbf{x}. \quad (9)$$

This distribution expresses beliefs about future (*i.e.* predictive) observations  $\mathbf{y}^*$  given the already observed ones  $\mathbf{y}$ . If it is possible to sample from the posterior distribution, *i.e.*  $(\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon}) \sim \pi(\mathbf{x}|\mathbf{y})$ , a sample from the posterior predictive distribution is obtained by drawing

$$\mathbf{Y}^{\text{postpred}} \sim \mathcal{N}(\mathbf{y}^*|\mathcal{M}(\mathbf{x}_{\mathcal{M}}), \Sigma(\mathbf{x}_{\varepsilon})), \quad \text{where } (\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon}) \sim \pi(\mathbf{x}|\mathbf{y}). \quad (10)$$

The posterior predictive distribution allows to assess the predictive capabilities of the model following calibration. It contains the uncertainty about the model parameters  $\mathbf{X}_{\mathcal{M}}$  and the mismatch parameters  $\mathbf{X}_{\varepsilon}$ . Because this distribution is defined in the space where the data  $\mathbf{y}$  are collected, it can be used to visually check the calibration results.

### 3.2 Sampling from the posterior distribution

The analytical computation of the posterior distribution is typically not possible. This is mostly due to difficulties in evaluating the normalizing constant  $Z$  defined in Eq. (8). Its computation typically relies on estimating an integral in the parameter space, which is in most cases intractable.

A breakthrough technique called *Markov chain Monte Carlo* (MCMC) sampling, originally developed by Metropolis et al. (1953) and Hastings (1970), completely avoids the need to evaluate this high-dimensional integral. It is a type of stochastic simulation technique that constructs Markov chains that are guaranteed to produce samples distributed according to the posterior distribution. Posterior characteristics (*e.g.* quantities of interest, expected values, marginal distributions etc.) can then be estimated using this sample.

Following the initial development of the MH algorithm (Metropolis et al., 1953; Hastings, 1970), recent developments to improve the algorithm’s efficiency strived towards adaptive proposal distributions (Haario et al., 2001; Roberts and Rosenthal, 2009) and the utilization of gradient information (Rosky et al., 1978; MacKay, 2003). One common flaw of these algorithms, however, is the requirement to tune them using a set of tuning parameters. This is a particularly tedious task that is a major source of error in practical applications.

The *affine-invariant ensemble sampler* (AIES, (Goodman and Weare, 2010)) is a fairly recent MCMC algorithm that performs particularly well in this respect. This algorithm requires only a single tuning parameter and its performance is invariant to affine transformations of the target distribution. This property makes it particularly useful for real-world applications,

where strong correlations between individual parameters often hinder conventional MCMC algorithms.

The AIES algorithm relies on a set of parallel chains where proposal samples are obtained by moving in the direction of a randomly chosen conjugate sample from a different chain. The pseudo-code for the implementation used in this paper is given in Algorithm 1.

A property that makes MCMC algorithms especially suitable for Bayesian computations is that they do not require the explicit computation of the normalization constant  $Z$  (from Eq. (8)), as only a posterior ratio, called *acceptance ratio*, is required (Step 9 of Algorithm 1). In this ratio,  $Z$  cancels out. However, the computationally expensive forward model must be evaluated each time this acceptance ratio is computed. This necessity of many runs of computationally expensive models has spurred the idea of constructing a *surrogate model* that, after successful construction, can be used in the MCMC algorithms in lieu of the original model, whereby the overall computational burden is reduced to feasible levels.

---

**Algorithm 1** Affine-invariant ensemble sampler (Goodman and Weare, 2010)

---

```

1: procedure AIES( $\mathbf{x}_0^{(1)}, \dots, \mathbf{x}_0^{(L)}, a, \text{NSteps}$ )
2:   for  $i \leftarrow 1, \text{NSteps}$  do
3:     for  $l \leftarrow 1, L$  do
4:        $\tilde{\mathbf{x}} \leftarrow \mathbf{x}_{i-1}^{(l)}$ 
5:        $\mathbf{x}^* \leftarrow \mathbf{x}_{i-1}^{(k)}$  with  $k \in \{1, \dots, L\} \setminus \{l\}$  chosen randomly
6:       Sample  $z \leftarrow \mathbf{Z} \sim g(z) \propto \begin{cases} \frac{1}{\sqrt{z}} & \text{if } z \in [\frac{1}{a}, a] \\ 0 & \text{otherwise} \end{cases}$ 
7:        $\hat{\mathbf{x}} \leftarrow \tilde{\mathbf{x}} + z(\mathbf{x}^* - \tilde{\mathbf{x}})$ 
8:       Sample  $u \leftarrow U \sim \mathcal{U}(0, 1)$ 
9:       if  $z^{M-1} \frac{\pi(\hat{\mathbf{x}}|\mathbf{y})}{\pi(\tilde{\mathbf{x}}|\mathbf{y})} > u$  then  $\triangleright M$  is dimension of  $\mathbf{x} \in \mathbb{R}^M$ 
10:         $\mathbf{x}_i^{(l)} \leftarrow \hat{\mathbf{x}}$ 
11:       else
12:         $\mathbf{x}_i^{(l)} \leftarrow \tilde{\mathbf{x}}$ 
13:       end if
14:     end for
15:   end for
16: end procedure

```

---

### 3.3 Surrogate modelling of the temperature time series

Often, sampling based techniques (*e.g.* MCMC algorithms) are considered infeasible because of the high number of computationally expensive model runs  $\mathcal{M}(\mathbf{X}_{\mathcal{M}})$  required. Surrogate

modelling techniques try to solve this problem by constructing a computationally cheap emulator that can be used instead of the original model.

Non-intrusive approaches to construct surrogate models are solely based on realizations of the model parameters and corresponding model outputs (Choi et al., 2004). The set of parameters used for constructing the surrogate is referred to as *experimental design*, which means here a *set of computer experiments* and shall not be confused with physical experiments. Following the assembly of the experimental design, the constructed surrogate model aims to approximate the original model predictions denoted by  $\mathcal{M}^{\text{PC}}$ ,

$$\mathcal{M}^{\text{PC}}(\mathbf{X}_{\mathcal{M}}) \approx \mathcal{M}(\mathbf{X}_{\mathcal{M}}). \quad (11)$$

This section details the construction of a surrogate model combining *polynomial chaos expansions* (PCE) with the *principal component analysis* (PCA).

### 3.3.1 Polynomial Chaos Expansions

*Polynomial chaos expansions* (PCE) are a surrogate modelling technique that has been used extensively in the engineering disciplines (Xiu and Karniadakis, 2002; Soize and Ghanem, 2004; Guo et al., 2018) to construct surrogate models of scalar-valued functions of random variables. A brief introduction to the method is presented next.

Assume a random vector  $\mathbf{X} = (X_1, \dots, X_M)$  with mutually independent components  $X_i \sim \pi_i(x_i)$ . Its joint probability density function is thus given by

$$\pi(\mathbf{x}) = \prod_{i=1}^M \pi_i(x_i). \quad (12)$$

The functional inner product of two polynomials  $\psi_k^i, \psi_l^i : x_i \in \mathcal{D}_{X_i} \mapsto \mathbb{R}$  of degree  $k$  and  $l$  respectively, is then defined by

$$\langle \psi_k^i, \psi_l^i \rangle_{\pi_i} \stackrel{\text{def}}{=} \int_{\mathcal{D}_{X_i}} \psi_k^i(x_i) \psi_l^i(x_i) \pi_i(x_i) dx_i. \quad (13)$$

By choosing these polynomials to fulfil  $\langle \psi_k^i, \psi_l^i \rangle_{\pi_i} = \delta_{k,l}$ , *i.e.*  $\delta_{k,l} = 1$  if  $k = l$  and 0 otherwise, these polynomials form a family of *univariate orthonormal polynomials*  $\{\psi_k^i\}_{k=0}^{\infty}$ . There exist well-known families of polynomial functions that fulfil the fundamental condition of Eq. (13) w.r.t. standard parametric probability distributions  $\pi_i$  (Askey and Wilson, 1985).

These univariate polynomials can be used to build *multivariate* polynomials by tensor product. Introducing the multi-indices  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M) \in \mathbb{N}^M$  the latter are defined by:

$$\Psi_{\boldsymbol{\alpha}}(\mathbf{x}) \stackrel{\text{def}}{=} \prod_{i=1}^M \psi_{\alpha_i}^i(x_i). \quad (14)$$

It can be shown that the univariate orthonormality property of Eq. (13) extends to the multivariate case and that the following holds:

$$\langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\beta}} \rangle_{\pi} \stackrel{\text{def}}{=} \int_{\mathcal{D}_{\mathbf{x}}} \Psi_{\boldsymbol{\alpha}}(\mathbf{x}) \Psi_{\boldsymbol{\beta}}(\mathbf{x}) \pi(\mathbf{x}) d\mathbf{x} = \delta_{\boldsymbol{\alpha}, \boldsymbol{\beta}}. \quad (15)$$

These polynomials  $\{\Psi_{\alpha}\}_{\alpha \in \mathbb{N}^M}$  form a so-called *orthonormal basis* of the space of *square integrable functions* with respect to the probability distribution  $\pi(\mathbf{x})$ . Any such function can be represented by:

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^M} a_{\alpha} \Psi_{\alpha}(\mathbf{x}), \quad (16)$$

where  $a_{\alpha} \in \mathbb{R}$  are the *coefficients* of the expansion.

In practical applications it is not feasible to compute the infinite number of coefficients  $a_{\alpha} \in \mathbb{N}^M$ . Instead, a *truncation scheme* is typically proposed that reduces the number of considered polynomials to a finite set. This truncated set denoted by  $\mathcal{A} \subset \mathbb{N}^M$  transforms the equality of Eq. (16) to an approximation

$$f(\mathbf{x}) \approx f^{\text{PCE}}(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{\alpha \in \mathcal{A}} a_{\alpha} \Psi_{\alpha}(\mathbf{x}). \quad (17)$$

In regression-based approaches, the coefficient vector  $\mathbf{a} \in \mathbb{R}^{\text{card}(\mathcal{A})}$  is typically estimated by least-squares analysis, as originally proposed in Berveiller et al. (2006). This corresponds to selecting a truncation set  $\mathcal{A}$  (Blatman and Sudret, 2011a) and using an *experimental design*  $\mathcal{X} \stackrel{\text{def}}{=} \{\mathbf{x}^{(i)}, i = 1, \dots, K\}$  to minimize the expression

$$\tilde{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathbb{R}^{\text{card}(\mathcal{A})}} \frac{1}{K} \sum_{i=1}^K \left( f(\mathbf{x}^{(i)}) - \sum_{\alpha \in \mathcal{A}} a_{\alpha} \Psi_{\alpha}(\mathbf{x}^{(i)}) \right)^2. \quad (18)$$

By storing the function evaluations at  $\mathcal{X}$  in a vector  $\mathcal{Y} \stackrel{\text{def}}{=} \{f(\mathbf{x}^{(1)}), \dots, f(\mathbf{x}^{(K)})\}$  the solution of Eq. (18) reads:

$$\tilde{\mathbf{a}} = (\mathbf{B}^{\top} \mathbf{B})^{-1} \mathbf{B}^{\top} \mathcal{Y}, \quad (19)$$

where  $\mathbf{B} = \{B_{ij} \stackrel{\text{def}}{=} \Psi_j(\mathbf{x}^{(i)}), i = 1, \dots, K, j = 1, \dots, \text{card}(\mathcal{A})\}$  are the evaluations of the basis polynomials  $\Psi_{\alpha}$  on the experimental design  $\mathcal{X}$ .

To assess the accuracy of the obtained polynomial chaos expansion, the so-called *generalization error*  $\mathbb{E}[(f(\mathbf{X}) - f^{\text{PCE}}(\mathbf{X}))^2]$  shall be evaluated. A robust error measure can be obtained by using the *leave-one-out* (LOO) cross validation technique. This estimator is obtained by

$$\varepsilon_{\text{LOO}} = \frac{1}{K} \sum_{i=1}^K \left( f(\mathbf{x}^{(i)}) - f_{\sim i}^{\text{PCE}}(\mathbf{x}^{(i)}) \right)^2, \quad (20)$$

where  $f_{\sim i}^{\text{PCE}}$  is constructed by leaving out the  $i$ -th point from the experimental design. After some algebraic manipulation, it can be shown that the LOO error can be computed as a mere post-processing of the PCE expansion as follows

$$\varepsilon_{\text{LOO}} = \frac{1}{K} \sum_{i=1}^K \left( \frac{f(\mathbf{x}^{(i)}) - f^{\text{PCE}}(\mathbf{x}^{(i)})}{1 - h_i} \right)^2, \quad (21)$$

where  $h_i$  is the  $i^{\text{th}}$  component of the vector given by:

$$\mathbf{h} = \text{diag}(\mathbf{B}(\mathbf{B}^{\top} \mathbf{B})^{-1} \mathbf{B}^{\top}), \quad (22)$$

for more details refer to Blatman and Sudret (2010).



This section outlined the approach to use PCE for approximating scalar quantities. Since the heat transfer model  $\mathbf{Y} = \mathcal{M}(\mathbf{X}_{\mathcal{M}})$  considered in this paper returns a vector of interface temperatures at 601 time steps, a pure PCE approach would require the construction of  $N = 601$  independent *polynomial chaos expansions*. Instead, a dimensionality reduction technique on the output is applied before using the PCE technique.

### 3.3.2 Principal Component Analysis

Because the discretized temperature evolution  $\mathbf{Y}$  is expected to be rather smooth (see Figure 11), considerable correlation between the individual time steps is expected. This correlation can be exploited to reduce the dimensionality of the output in the context of surrogate modelling.

There exist numerous so-called *dimensionality reduction techniques* (van der Maaten et al., 2008), one of which is *principal component analysis* (PCA, Jolliffe (2002)). The latter utilizes an orthogonal transformation to express  $\mathbf{Y}$  in a new basis of uncorrelated *principal components*  $\mathbf{Z}$ .

In practice, PCA is carried out by computing estimators of the expectation  $\boldsymbol{\mu}_{\mathbf{Y}} \approx \mathbb{E}[\mathbf{Y}]$  and the covariance matrix  $\boldsymbol{\Sigma}_{\mathbf{Y}} \approx \text{Cov}[\mathbf{Y}]$ . The  $N$  eigenvectors of this covariance matrix are denoted by  $\boldsymbol{\phi}_p$  for  $p = 1, \dots, N$ . The associated eigenvalue  $\lambda_p$  corresponds to the variance of  $\mathbf{Y}$  in direction of the  $p$ -th principal component. Thereby the random vector  $\mathbf{Y}$  can be expressed through its  $N$  principal components  $z_p(\mathbf{X}_{\mathcal{M}})$  as  $\mathbf{Y} = \boldsymbol{\mu}_{\mathbf{Y}} + \sum_{p=1}^N z_p(\mathbf{X}_{\mathcal{M}})\boldsymbol{\phi}_p$ .

The model output can then be compressed to a lower dimensional subspace by retaining only those  $N'$  principal components with the highest variance:

$$\mathbf{Y} \approx \mathbf{Y}^{\text{PCA}} = \boldsymbol{\mu}_{\mathbf{Y}} + \sum_{p=1}^{N'} z_p(\mathbf{X}_{\mathcal{M}})\boldsymbol{\phi}_p. \quad (23)$$

The number of terms  $N'$  is selected such that  $\sum_{p=1}^{N'} \lambda_p = (1 - \varepsilon_0) \sum_{p=1}^N \lambda_p$ , with  $\varepsilon_0$  typically chosen as 0.01. This way, the model output  $\mathbf{Y} \in \mathbb{R}^N$  can be approximated by a linear transformation of the principal component vector  $\mathbf{Z} = (z_1(\mathbf{X}_{\mathcal{M}}), \dots, z_{N'}(\mathbf{X}_{\mathcal{M}}))^{\top}$  thereby reducing the problem dimensionality from  $N$  to  $N' \ll N$ .

### 3.3.3 Combining PCA with PCE

The combination of PCA with PCE gives rise to an efficient surrogate modelling technique as shown originally in Blatman and Sudret (2011b). Constructing  $N'$  polynomial chaos expansions of each retained principal component  $z_p(\mathbf{X}_{\mathcal{M}}) \approx z_p^{\text{PCE}}(\mathbf{X}_{\mathcal{M}}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_p} \tilde{a}_{p,\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{X}_{\mathcal{M}})$ , together with the PCA formulation from Eq. (23) yields a surrogate model relating the model parameters to the vector valued time series output of the transient heat transfer problem:

$$\mathbf{Y} \approx \mathcal{M}^{\text{PC}}(\mathbf{X}_{\mathcal{M}}) \stackrel{\text{def}}{=} \mathbf{Y}^{\text{PCA+PCE}} = \boldsymbol{\mu}_{\mathbf{Y}} + \sum_{p=1}^{N'} \left( \sum_{\boldsymbol{\alpha} \in \mathcal{A}_p} \tilde{a}_{p,\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{X}_{\mathcal{M}}) \right) \boldsymbol{\phi}_p, \quad (24)$$

which can be rewritten by introducing the union set  $\mathcal{A}^* \stackrel{\text{def}}{=} \bigcup_{p=1}^{N'} \mathcal{A}_p$ :

$$\mathbf{Y}^{\text{PCA+PCE}} = \boldsymbol{\mu}_Y + \sum_{\alpha \in \mathcal{A}^*} \sum_{p=1}^{N'} \tilde{a}_{p,\alpha} \Psi_{\alpha}(\mathbf{X}_{\mathcal{M}}) \phi_p. \quad (25)$$

For compactness, this equation can also be expressed in matrix form by letting  $\boldsymbol{\Phi} = (\phi_1, \dots, \phi_{N'})$  be a  $N \times N'$  matrix containing the retained eigenvectors  $\phi_p = (\phi_{p1}, \dots, \phi_{pN})^\top$ . For the PCE part of the equation the vector  $\boldsymbol{\Psi}(\mathbf{X}_{\mathcal{M}}) = \{\Psi_{\alpha}(\mathbf{X}_{\mathcal{M}}), \alpha \in \mathcal{A}^*\}$  is introduced that holds the individual multivariate orthogonal polynomials. Let  $\mathbf{A}$  be a  $\text{card}(\mathcal{A}^*) \times N'$  matrix that stores the corresponding PCE coefficients, then Eq. (24) can be written as

$$\mathbf{Y} \approx \boldsymbol{\mu}_Y + \boldsymbol{\Phi} (\mathbf{A}^\top \boldsymbol{\Psi}(\mathbf{X}_{\mathcal{M}})). \quad (26)$$

For completeness, the response can also be expressed for each random variable  $Y_t$  individually. For this, the row vector  $\phi_t^{\text{row}} = (\phi_{1t}, \dots, \phi_{N't})$ , taken from the  $t$ -th row of the matrix of eigenvectors  $\boldsymbol{\Phi}$ , is introduced:

$$Y_t \approx \mu_{Y_t} + \phi_t^{\text{row}} \mathbf{A}^\top \boldsymbol{\Psi}(\mathbf{X}_{\mathcal{M}}). \quad (27)$$

This surrogate model can then be used in lieu of the original computationally expensive forward model. The evaluation of the surrogate model is orders of magnitude faster than the original finite element model. For comparison, in our application example a single FE run takes about 1 min on a conventional computer, while in the same time  $10^7$  evaluations of the surrogate model can be made.

This reduction in computational time is a promising feature of the presented surrogate modelling technique. It does, however, come at the cost of a series of approximations that are introduced during the PCA and PCE computation. To ensure confidence in the produced surrogate model, a general error measure has to be devised. It includes the approximation error due to the PCA truncation and the truncated polynomial chaos expansion. Such an error measure  $\tilde{\eta}$  was derived in Blatman and Sudret (2013). For the sake of completeness the details are given in A.

### 3.4 Summary of the proposed method

In this section, a procedure to efficiently conduct Bayesian inference with expensive vector valued models was presented. It is assumed that the parametrization of the temperature-dependent effective material properties (see Section 2.3) is known. Bayesian inference then aims at determining the distribution of the parameters  $\mathbf{X}|\mathbf{y} \sim \pi(\mathbf{x}|\mathbf{y})$  after observations (*i.e.* experimental measurements) have been made. A brief step-by-step account of this procedure is given below for reference:

- Step 1** Choose a prior distribution  $\pi(\mathbf{x})$  on  $\mathbf{X}$  and construct an experimental design  $\mathcal{X}$  using  $K$  samples from this prior. Evaluate the forward model at  $\mathcal{X}$  and store the evaluations in  $\mathcal{Y}$ .

- Step 2** Approximate  $\mathcal{M}(\mathbf{X}_{\mathcal{M}})$  using the surrogate model  $\mathcal{M}^{\text{PC}}(\mathbf{X}_{\mathcal{M}})$  from Eq. (26). This requires the combination of the dimensionality reduction technique PCA with the PCE uncertainty propagation technique.
- Step 3** Compute the error estimate  $\tilde{\eta}$  from Eq. (49). This error is only valid over the prior domain. If it is too large, enrich the experimental design by increasing the number of samples  $K$  and restart from **Step 1**. The size of the admissible error depends on the application but should typically not exceed 5%.
- Step 4** Define a likelihood function  $\mathcal{L}(\mathbf{x}; \mathbf{y})$  from Eq. (7) that captures the discrepancy between a model run and the observations.
- Step 5** Run the AIES defined in Algorithm 1 where the likelihood function uses the surrogate model  $\mathcal{M}^{\text{PC}}(\mathbf{X}_{\mathcal{M}})$  instead of the original model  $\mathcal{M}(\mathbf{X}_{\mathcal{M}})$  to obtain a sample from the posterior distribution  $\pi(\mathbf{x}|\mathbf{y})$ .
- Step 6** Verify the fit of the calibrated model using a sample from the posterior predictive distribution from Eq. (9). Samples from the posterior predictive distribution can be obtained by reusing parameter samples distributed according to the posterior distribution from **Step 5**.

This method works if the support domain of the prior distribution contains that of the posterior distribution. In this respect, sufficiently large prior ranges shall be selected based on the expert’s judgment.

The successful calibration of the parameters through the Bayesian inference approach gives insight into the model mismatch and correlation structure between individual parameters. The distribution of the parameters can further be used in probabilistic analysis using these models and, given new observations, can be updated to reflect beliefs incorporating the newly acquired information.

A fundamental ingredient of the presented approach is the necessity to define a parametrization of the thermal effective material properties as described in Section 2.3. To judge the quality of the parametrization, it can be helpful to assess the relative importance of a single model parameter with respect to the output. For this, it is necessary to resort to the field of sensitivity analysis.

## 4 Sensitivity analysis

### 4.1 PCE based Sobol’ indices

Global sensitivity analysis aims at finding which input parameters of a computer model (or combination thereof) explain at best the uncertainties in the model predictions. In this respect, variance decomposition techniques rely on assigning fractions of the model output

variance  $\text{Var}[Y] = \text{Var}[\mathcal{M}(\mathbf{X})]$  to the individual input parameters  $X_i$ . For simplicity, in this section the subscript  $(\cdot)_{\mathcal{M}}$  from the parameter vector  $\mathbf{X} = \mathbf{X}_{\mathcal{M}}$  is dropped.

Consider a scalar-valued computational model  $\mathcal{M} : \mathbf{X} \in [0, 1]^M \mapsto \mathcal{M}(\mathbf{X}) \in \mathbb{R}$ , which maps a vector of input parameters in the unit hypercube to the real numbers. This computational model can be decomposed into a sum of terms that only depend on a subset of the input parameters, *i.e.* a constant  $\mathcal{M}_0$ , univariate functions  $\{\mathcal{M}_i(X_i), i = 1, \dots, M\}$ , bivariate functions *etc.*

$$\mathcal{M}(\mathbf{X}) = \mathcal{M}_0 + \sum_{i=1}^M \mathcal{M}_i(X_i) + \sum_{1 \leq i < j \leq M} \mathcal{M}_{ij}(X_i, X_j) + \dots + \mathcal{M}_{1,2,\dots,M}(X_1, \dots, X_M). \quad (28)$$

This decomposition is called the *Hoeffding-Sobol' decomposition* and is unique for any function  $\mathcal{M}$  that is square-integrable over the unit hypercube (Sobol', 1993).

Denoting by  $\mathbf{u} \stackrel{\text{def}}{=} \{i_1, \dots, i_s\} \subset \{1, \dots, M\}$  a subset of indices, Eq. (28) can be written in short:

$$\mathcal{M}(\mathbf{X}) = \mathcal{M}_0 + \sum_{\mathbf{u} \subset \{1, \dots, M\}} \mathcal{M}_{\mathbf{u}}(\mathbf{X}_{\mathbf{u}}). \quad (29)$$

It can be shown that the terms of this equation called *summands*, are orthogonal (Sobol', 1993). The variance of each term  $\mathcal{M}_{\mathbf{u}}(\mathbf{X}_{\mathbf{u}})$ , called *partial variance*, is obtained by:

$$D_{\mathbf{u}} \stackrel{\text{def}}{=} \int_{[0,1]^{\text{card}(\mathbf{u})}} \mathcal{M}_{\mathbf{u}}^2(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}}. \quad (30)$$

Due to the orthogonality of the terms in this equation, the total variance of the model output  $D = \text{Var}[\mathcal{M}(\mathbf{X})]$  is finally obtained as the sum of the partial variances

$$D = \sum_{\mathbf{u} \subset \{1, \dots, M\}} D_{\mathbf{u}}. \quad (31)$$

Each partial variance describes the amount of the output variance that can be attributed to the interaction of the input variables  $\mathbf{X}_{\mathbf{u}}$ . In particular,  $D_i$  describes the fraction of the variance that can be attributed to one input variable  $X_i$  taken separately.

Moreover, the total contribution to the variance attributable to a single input parameter  $X_i$  is captured in the sum of the partial variances  $D_{\mathbf{u}}$  that contain the  $i$ -th input variable. The sum of these partial variances normalized by the total variance is called the  $i$ -th *total Sobol' index* and is defined as

$$S_i^T \stackrel{\text{def}}{=} \frac{1}{D} \sum_{\mathbf{u} \supset \{i\}} D_{\mathbf{u}}. \quad (32)$$

It is noted here that the sum of all total Sobol' indices, *i.e.*,  $\sum_{i \in \{1, \dots, M\}} S_i^T$ , is larger than one because the same interaction effect contributes to multiple total Sobol' indices. Usually the integral in Eq. (30) can be computed through Monte Carlo integration. However, if the model  $\mathcal{M}$  is expressed in an orthogonal basis (as is the case for PCE, Eq. (16)), the Sobol' indices can be computed analytically by post-processing the PCE coefficients  $a_{\alpha}$  (Sudret, 2006, 2008):

$$S_i^T = \frac{1}{D} \sum_{\alpha \in \mathcal{A}_{i>0}} a_{\alpha}^2, \quad \text{with} \quad \mathcal{A}_{i>0} = \{\alpha \in \mathcal{A} : \alpha_i > 0\}, \quad (33)$$

*i.e.*  $\mathcal{A}_{i>0}$  is the set of multivariate polynomials that are non-constant in the  $i$ -th input parameter  $X_i$ . For scalar-valued models, this yields a measure of the variance fraction that can be attributed to a certain input parameter. In the following section, this concept is extended to models with multiple outputs.

## 4.2 PCA-based Sobol' indices

In the present paper models with multiple outputs (*i.e.* time-series of computed temperatures) are considered. By using a surrogate model that combines PCE with PCA, as discussed in Section 3.3.3, the total Sobol' indices for each output vector component (*i.e.* time step) can also be computed analytically (Marelli and Sudret, 2015; Nagel, Rieckermann, and Sudret, Nagel et al.).

For this, the partial variances of the model response components  $Y_t$  are computed by using the expression from Eq. (27). The total Sobol' index for the  $t$ -th component of the output random vector then reads

$$S_{i,t}^T = 1 - \frac{\sum_{\alpha \in \mathcal{A}_{i=0}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p,\alpha} \right)^2}{\sum_{\alpha \in \mathcal{A}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p,\alpha} \right)^2}. \quad (34)$$

where  $\mathcal{A}_{i=0}^*$  is the subset of  $\mathcal{A}^*$  for which  $\alpha_i = 0$ . The interested reader is referred to B for the derivations.

## 5 Results

In this section, the procedure presented in Sections 3 and 4 is applied to calibrate the temperature-dependent material properties of gypsum based insulation boards. The experimental data stems from experiments conducted by Breu (2016) and Just (2016) that were presented in Section 2.1. As explained in Section 2.3, the material properties are parameterized with a set of 6 parameters. In the Bayesian inference framework introduced in Section 3.1, determining the posterior distribution of these parameters constitutes the calibration of the temperature-dependent material properties.

To further investigate the effects of the introduced parametrization, the surrogate models  $\mathcal{M}^{\text{PC}}(\mathbf{X}_{\mathcal{M}})$  used for calibration are reused to conduct time-dependent sensitivity analyses (Section 4). These analyses show the influence each model parameter  $X_i$  has on the simulation output. They deliver valuable insights and can be used to further refine the model parametrization.

Finally, the calibrated time-dependent material properties are validated by simulating insulation panels in a different measurement setup and comparing these simulation results with actual measurements.

## 5.1 Calibration of material properties for gypsum boards

In this section the general calibration procedure from Section 3.4 is applied to the specific problem of calibrating heat transfer models describing the experiments of specimens E1-E4 (Test 1) presented in Section 2.1.

The model parameters  $\mathbf{X}_{\mathcal{M}}$  are assumed to be priorly independent and uniformly distributed with the lower and upper bounds ( $\underline{x}_i$  and  $\bar{x}_i$  respectively) defined by the ranges given in Table 3. The prior distribution of the model parameters is thus given by

$$\pi(\mathbf{x}_{\mathcal{M}}) = \prod_{i=1}^6 \mathcal{U}(x_i; \underline{x}_i, \bar{x}_i). \quad (35)$$

Since multiple measurements  $\mathbf{y}^{(s)}$  are available for each experiment, the formulation for the likelihood Eq. (6) has to be slightly adapted. Under the assumption of independence between the individual measurement locations, it can be written as the product

$$\mathcal{L}(\mathbf{x}_{\mathcal{M}}, \mathbf{x}_{\varepsilon}; \mathbf{y}) = \prod_{s=1}^S \mathcal{N}(\mathbf{y}^{(s)}; \mathcal{M}^{\text{PC}}(\mathbf{x}_{\mathcal{M}}), \boldsymbol{\Sigma}(\mathbf{x}_{\varepsilon})), \quad (36)$$

which generalizes Eq. (6) where only a single time series of measurements was considered. Consequently, the posterior distribution obtained from Bayes' theorem should strictly be written as  $\pi(\mathbf{x}|\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(S)})$ , but for notational simplicity the superscript  $^{(s)}$  is again dropped.

The covariance matrix  $\boldsymbol{\Sigma}(\mathbf{x}_{\varepsilon})$  is parametrized by

$$\boldsymbol{\Sigma}(\mathbf{x}_{\varepsilon}) = \{\Sigma(\mathbf{x}_{\varepsilon})_{ij} \stackrel{\text{def}}{=} \sigma_i \sigma_j R(t_i, t_j, \theta), \quad i, j = 1, \dots, N\}, \quad (37)$$

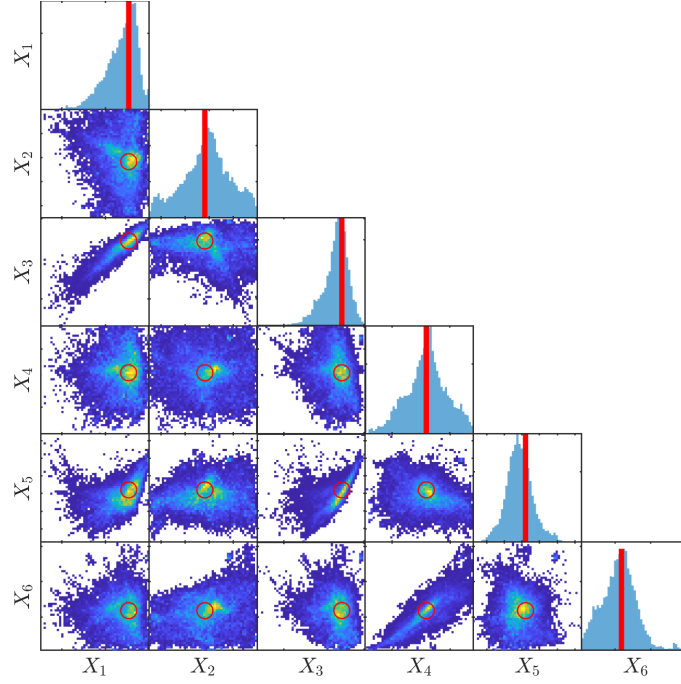
where we choose a so-called *Matérn* 5/2 autocorrelation function ( $h = t_i - t_j$ ):

$$R(h, \theta) = \left( 1 + \frac{\sqrt{5}|h|}{\theta} + \frac{5h^2}{3\theta^2} \right) \exp\left( -\frac{\sqrt{5}|h|}{\theta} \right). \quad (38)$$

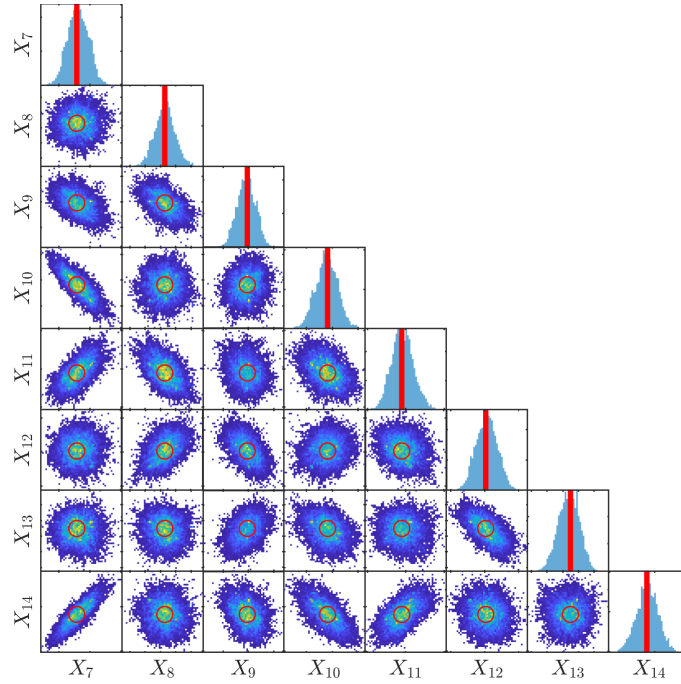
In this autocorrelation function,  $\theta$  is the correlation length and  $\sigma_i$  is the standard deviation at the  $i$ -th time step. To reduce the number of calibration parameters, it is assumed that the standard deviation  $\sigma(t)$  follows a degree-6 polynomial function

$$\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N), \quad \text{with} \quad \sigma_i \stackrel{\text{def}}{=} \sigma(t_i) = \sum_{k=0}^6 \varpi_k \psi_k \left( 2 \frac{t_i - t_1}{t_N} - 1 \right), \quad (39)$$

where  $\psi_k$  is the  $k$ -th Legendre polynomial (defined over  $[-1, 1]$ ) and  $\varpi_k$  are the coefficients to be calibrated. As a summary, there are 8 parameters to define our discrepancy term, namely  $\{\varpi_0, \dots, \varpi_6\}$  which define the non-stationary variance of the model discrepancy, and the autocorrelation length  $\theta$ . Using the notation from Section 3.1, we pose  $\mathbf{X}_{\varepsilon} = (\varpi_0, \dots, \varpi_6, \theta)$ . To complete the prior information, a uniform distribution is assumed for the discrepancy parameters with non-restrictive bounds. A summary of the full prior distribution is given in Table 4.



(a)  $\pi(\mathbf{x}_{\mathcal{M}}|\mathbf{y})$



(b)  $\pi(\mathbf{x}_{\varepsilon}|\mathbf{y})$

Figure 6: Univariate and bivariate marginals from the posterior distribution of the model parameters  $\pi(\mathbf{x}_{\mathcal{M}}|\mathbf{y})$  and discrepancy parameters  $\pi(\mathbf{x}_{\varepsilon}|\mathbf{y})$  calibrated using the data from Product A (E1). The vertical line (dot) indicates the MAP parameter  $\mathbf{x}^{\text{MAP}}$  defined in Eq. (40).

Table 4: Summary of the prior distribution  $\pi(\mathbf{x}) = \prod_{i=1}^{14} \pi_i(x_i)$  for the parameter vector  $\mathbf{X} = (X_1, \dots, X_{14})^\top$

		$\pi_i(x_i)$	$\mu$	$\sigma$	c.o.v.	units
$\mathbf{X}_{\mathcal{M}}$ :	$X_1$	$\mathcal{U}(300, 800)$	$5.50 \cdot 10^2$	$1.20 \cdot 10^1$	$2.18 \cdot 10^{-2}$	$^\circ\text{C}$
	$X_2$	$\mathcal{U}(0.1, 1)$	$5.50 \cdot 10^{-1}$	$5.10 \cdot 10^{-1}$	$9.27 \cdot 10^{-1}$	-
	$X_3$	$\mathcal{U}(0.1, 0.25)$	0.175	6.58	$3.76 \cdot 10^1$	W/mK
	$X_4$	$\mathcal{U}(0.1, 1.2)$	0.65	17.8	$2.74 \cdot 10^1$	W/mK
	$X_5$	$\mathcal{U}(1.4 \cdot 10^4, 6.5 \cdot 10^4)$	$3.95 \cdot 10^4$	$1.21 \cdot 10^2$	$3.07 \cdot 10^{-3}$	J/kgK
	$X_6$	$\mathcal{U}(10^3, 8 \cdot 10^4)$	$4.05 \cdot 10^4$	$1.51 \cdot 10^2$	$3.73 \cdot 10^{-3}$	J/kgK
$\mathbf{X}_{\varepsilon}$ :	$X_7$	$\mathcal{U}(0, 20)$	$1.00 \cdot 10^1$	2.40	$2.40 \cdot 10^{-1}$	$^\circ\text{C}$
	$X_8$	$\mathcal{U}(-20, 20)$	0	3.40	-	$^\circ\text{C}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$X_{13}$	$\mathcal{U}(-20, 20)$	0	3.40	-	$^\circ\text{C}$
	$X_{14}$	$\mathcal{U}(0, 50)$	$2.50 \cdot 10^1$	3.80	$1.52 \cdot 10^{-1}$	s

Figure 6 shows the resulting sample points of the posterior distribution obtained for an exemplary calibration run for the E1 setup. This sample was produced using the previously presented AIES algorithm (Algorithm 1).

Despite the broad information contained in the full posterior plot, one is often interested in the set of parameters that *best* describe the observations. In accordance with the Bayesian interpretation of probabilities, this parameter set is located at the maximum value of the posterior distribution (*maximum a posteriori*, MAP). It can be found by solving the optimization problem

$$\mathbf{x}^{\text{MAP}} = \arg \max_{\mathbf{x}} \pi(\mathbf{x}|\mathbf{y}). \quad (40)$$

This problem can be approximately solved by picking the parameter point from the available posterior sample that maximizes the unnormalized posterior distribution  $\tilde{\pi}(\mathbf{X}|\mathbf{y}) = \mathcal{L}(\mathbf{x}; \mathbf{y})\pi(\mathbf{x}) \propto \pi(\mathbf{x}|\mathbf{y})$ . The resulting *maximum a posteriori* estimator is also shown in Figure 6.

The calibrated posterior parameters for Product A (E1) are summarized in Table 5. It gives an overview of the calibrated parameters, including a set of summary statistics.

A major advantage full samples have compared to point estimators is that they allow investigating characteristics of the posterior distribution. This provides a fuller picture of the calibrated parameter vector, *e.g.* by showing dependence between individual parameters  $X_i$ , allowing the computation of confidence intervals or revealing problems with identifiability.



Table 5: Posterior statistics for the calibration with Product A (E1). The values are computed from the available posterior sample and include the MAP estimate, the empirical mean  $\hat{\mu}$ , the empirical 95% confidence interval, the empirical standard deviation  $\hat{\sigma}$ , and the empirical coefficient of variation c.o.v.  $\stackrel{\text{def}}{=} \hat{\sigma}/\hat{\mu}$ . The prior statistics are shown in Table 4.

	MAP	$\hat{\mu}$	95% conf. interval	$\hat{\sigma}$	c.o.v.
$X_1$	$6.91 \cdot 10^2$	$6.81 \cdot 10^2$	$[5.29 \cdot 10^2, 7.88 \cdot 10^2]$	$6.60 \cdot 10^1$	$9.70 \cdot 10^{-2}$
$X_2$	$5.08 \cdot 10^{-1}$	$5.65 \cdot 10^{-1}$	$[1.48 \cdot 10^{-1}, 9.60 \cdot 10^{-1}]$	$2.27 \cdot 10^{-1}$	$4.01 \cdot 10^{-1}$
$X_3$	0.185	0.183	$[0.167, 0.196]$	$6.99 \cdot 10^{-3}$	$3.82 \cdot 10^{-2}$
$X_4$	0.799	0.798	$[0.410, 1.15]$	0.199	$2.49 \cdot 10^{-1}$
$X_5$	$3.77 \cdot 10^4$	$3.75 \cdot 10^4$	$[3.46 \cdot 10^4, 4.14 \cdot 10^4]$	$1.58 \cdot 10^3$	$4.21 \cdot 10^{-2}$
$X_6$	$2.25 \cdot 10^4$	$2.17 \cdot 10^4$	$[3.55 \cdot 10^3, 4.65 \cdot 10^4]$	$1.10 \cdot 10^4$	$5.09 \cdot 10^{-1}$
$X_7$	8.56	8.62	$[7.99, 9.25]$	$3.36 \cdot 10^{-1}$	$3.89 \cdot 10^{-2}$
$X_8$	$4.53 \cdot 10^{-1}$	$4.58 \cdot 10^{-1}$	$[1.99 \cdot 10^{-1}, 7.25 \cdot 10^{-1}]$	$1.33 \cdot 10^{-1}$	$2.90 \cdot 10^{-1}$
$X_9$	1.35	1.36	$[1.59, 1.13]$	$1.79 \cdot 10^{-1}$	$5.87 \cdot 10^{-2}$
$X_{10}$	2.39	2.41	$[2.69, 2.14]$	$2.11 \cdot 10^{-1}$	$3.92 \cdot 10^{-2}$
$X_{11}$	2.46	2.46	$[2.19, 2.75]$	$1.43 \cdot 10^{-1}$	$5.84 \cdot 10^{-2}$
$X_{12}$	$5.06 \cdot 10^{-1}$	$5.29 \cdot 10^{-1}$	$[2.34 \cdot 10^{-1}, 8.29 \cdot 10^{-1}]$	$1.56 \cdot 10^{-1}$	$2.95 \cdot 10^{-1}$
$X_{13}$	$2.07 \cdot 10^{-1}$	$2.16 \cdot 10^{-1}$	$[3.93 \cdot 10^{-1}, 4.96 \cdot 10^{-2}]$	$1.33 \cdot 10^{-1}$	$2.75 \cdot 10^{-1}$
$X_{14}$	$2.76 \cdot 10^1$	$2.77 \cdot 10^1$	$[2.67 \cdot 10^1, 2.86 \cdot 10^1]$	$4.87 \cdot 10^{-1}$	$1.76 \cdot 10^{-2}$

Additionally, the full parameter distribution explains why it can be hard to calibrate with the conventional approach. When strong correlations exist, such as for  $X_3$  and  $X_5$  in Figure 6, it is hard to move to a better guess by changing just one parameter.

In conclusion, the reduction of the standard deviation in all posterior parameters in conjunction with the unimodal posterior distribution can be seen as an indicator of a successful calibration.

## 5.2 Discussion of the calibration results

The results of the proposed calibration procedure for four different insulation materials are discussed next. The materials are all characterized by the temperature-dependent material properties  $(\lambda(T, \mathbf{X}_{\mathcal{M}}), c(T, \mathbf{X}_{\mathcal{M}}), \rho(t, \mathbf{X}_{\mathcal{M}}))$ . The calibration results for Product A are summarized in Figure 7 and Table 5. In C the results of the remaining products are presented. The discussion in this section refers to them at times. In this section we use the notation

$\mathbf{X}^{\text{prior}} \stackrel{\text{def}}{=} \mathbf{X}$  and  $\mathbf{X}^{\text{post}} \stackrel{\text{def}}{=} \mathbf{X}|\mathbf{y}$  to more clearly distinguish between the prior and posterior random variables.

For each material, Figures (a) and (b) show the temperature-dependent conductivity  $\lambda(T)$  and heat capacity  $c(T)$ . Since the full posterior distribution of the parameters is inferred, the plots do not only show one line for prior and posterior, but 1,000 samples each. The shown curves result from prior parameter draws  $\mathbf{X}^{\text{prior}} \sim \pi(\mathbf{x})$  (*i.e.* before calibration), posterior parameter draws  $\mathbf{X}^{\text{post}} \sim \pi(\mathbf{x}|\mathbf{y})$  (*i.e.* after calibration) and the MAP parameter  $\mathbf{x}^{\text{MAP}}$  (see Eq. (40)). The calibrated density  $\rho(T)$  is not shown, since its two governing parameters  $X_1$  and  $X_2$  can be seen also in the plots of the calibrated  $\lambda(T)$  and  $c(T)$ . It is obvious that the posterior samples have a smaller variance than the prior samples since the bandwidth of the 1,000 curves is much smaller.

A plot of the model predictions together with the measured data  $\mathbf{y}$  is presented in Figures (c). These plots show runs of the surrogate model before calibration ( $\mathcal{M}^{PC}(\mathbf{X}^{\text{prior}})$ ) and confidence intervals  $CI$  of the predictions  $\mathbf{Y}^{\text{postpred}}$  of the surrogate model following calibration (see Eq. (9)). The MAP parameter  $\mathbf{x}^{\text{MAP}}$  is propagated through the surrogate and the response of the original finite element model for  $\mathbf{x}^{\text{MAP}}$  is also plotted. Additionally, samples from the posterior are used to show the calibrated discrepancy standard deviation  $\sigma(\mathbf{X}^{\text{post}})$  (see Eq. (39)).

To emphasize the reduction in uncertainty by the presented calibration procedure, additionally posterior distributions of the predicted temperature at *snapshot* times  $t = 20$  min and  $t = 30$  min are shown in Figures (d) and (e). These plots show kernel density estimates of the output probability density function of the surrogate model before calibration  $\mathcal{M}_t^{PC}(\mathbf{X}^{\text{prior}})$  and after calibration  $Y_t^{\text{postpred}}$  at the respective times  $t$ . Additionally, the measured data  $y$  and the MAP predictions are displayed. These plots are now discussed in more detail.

The distributions of  $X_3$  (conductivity between 180°C and the start  $X_1$  of the second key process) and  $X_5$  (distribution of the heat capacity at 140°C), show a large variance reduction. This means that the information gained about these parameters through the conducted calibration is high.

Generally, it can be expected that the heat capacity and the conductivity correlate, since a higher heat flow due to a higher conductivity can be *compensated* by a higher heat capacity, resulting in the same temperature profile. This is indeed the case for the parameters  $X_3$  and  $X_5$  as well as for  $X_4$  and  $X_6$  (which describe the conductivity and the heat capacity at higher temperatures), as seen in Figure 6. The effect of this correlation and its *compensation effect* can be seen by comparing Figure 13 (E2) to Figures 7 (E1), 15 (E3) and 17 (E4). Despite the fact that parameters  $X_4$  and  $X_6$  show a higher posterior variance for specimens E1, E3 and E4 than for E2, the variance of the resulting model predictions (Figures (c)) is comparable.

Despite the fact that the shown prior thermal material properties in Figures (d) and (e)

are based on the same prior assumptions for the parameters  $\{X_i, i = 1, \dots, 6\}$ , the modes of the prior distributions estimated from  $\mathcal{M}_t^{PC}(\mathbf{X}^{\text{prior}})$  are significantly different for the individual specimens. This is caused by two factors: (1) the variable layer thickness and (2) the parametrization. The layer thickness influences the measured temperature simply by the fact that the further inside the specimen the measurement is placed, the colder it is at a certain time. This can easily be seen by comparing the different Figures (c), from 9.5 mm (Figure 13) to 12.5 mm (Figures 7 and 15) to 15 mm (Figure 17). Additionally, parameter  $X_1$ , the start time of the second key process, is responsible for the large variance of the start time of the temperature increase after the 100°C plateau. For higher values of  $X_1$ , the plateau is longer and the temperature increase starts later. Thus if the snapshots (Figures (d) and (e)) are taken where most prior temperatures are still at the plateau, the probability density functions have positive (right) skew. The later the snapshot is drawn, the more negatively skewed is the probability density function and the larger its variance gets.

Product B (E2) is the only product where the conductivity decreases after 400°C and where there is hardly a second key process according to the posterior curves. This shows that this product is made of a significantly different material.

Generally, all posterior predictive distributions  $\pi(\mathbf{y}^*|\mathbf{y})$  agree well with the measurements. This suggests that the used heat transfer model with the chosen parametrization is sufficiently accurate to reproduce the observations.

The calibrated temperature-dependent material properties show a higher variance reduction at lower temperatures. This is mostly due to the maximum system temperature at 40 min being only 885°C as seen in Figure 3.

It can be clearly seen in Figures (c) that the calibrated discrepancy standard deviation  $\sigma(\mathbf{X}^{\text{post}})$  varies similarly across all four experiments. It is influenced by two factors: (1) the temperature variance from the data series  $\mathbf{y}^{(s)}$  (*i.e.* the individual sensor recordings) and (2) model insufficiencies that hinder a fully accurate reproduction of the data. The latter can be reduced by changing the parametrization: an additional model parameter between approximately 200°C and 400°C could reduce the calibrated mismatch standard deviation for the temperature increase after the 100°C plateau. This temperature range currently lacks an independent parameter.

### 5.3 Time-dependent sensitivity analysis

Using the PCA+PCE surrogate model and the derivations from Section 4, the constructed surrogate model of the temperature evolution can be reused to compute the time-dependent Sobol' indices as measures for the individual model parameters' importance across the simulation time.

The time-dependent Sobol' indices  $S_i^T$  were computed for all four surrogate models con-

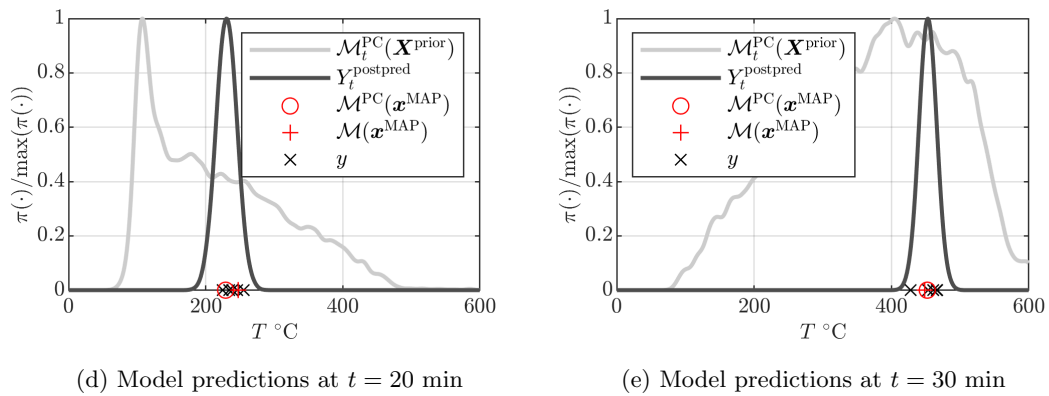
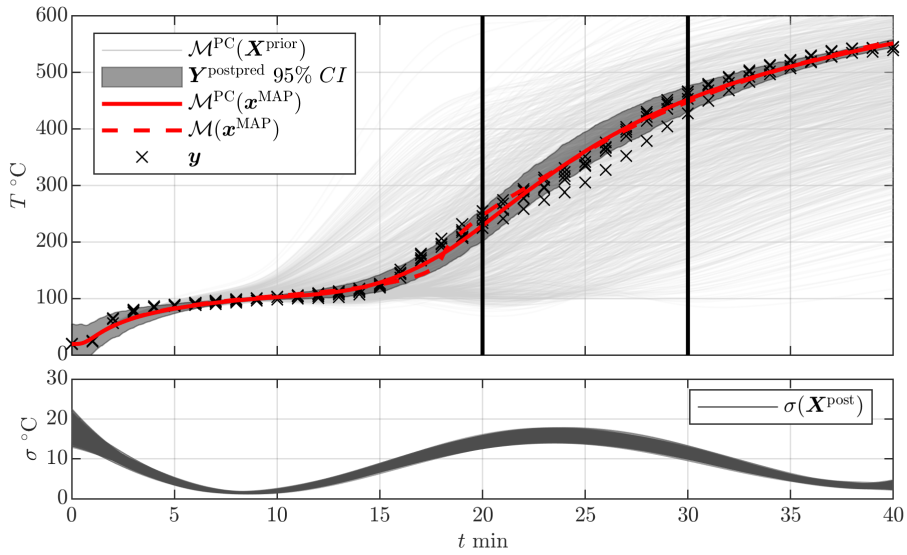
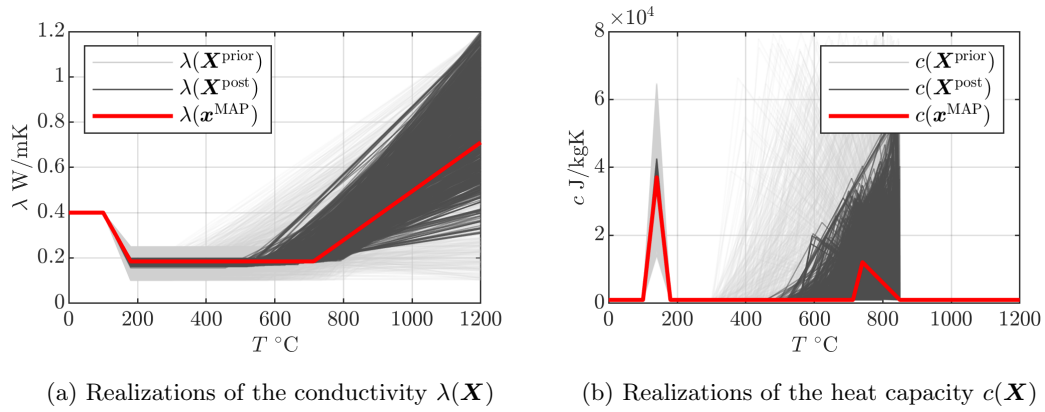


Figure 7: Calibration results for Product A 12.5 mm insulation (E1), experiments conducted by Just (2016).

structured for the experimental setups E1-E4. They are displayed in Figures 8 for Product A. In C the results of the remaining products are presented. The discussion in this section refers to them at times.

Generally there is no single parameter that clearly dominates the simulation output across the whole simulation time range. Instead, most parameters have a clear time range where they are important and other time ranges where their influence can be neglected. All parameters show a similar behaviour across the four models.

At early simulation times before  $t = 15$  min, the Sobol' indices  $S_3^T$  and  $S_5^T$  clearly dominate the sensitivity analysis across all four models. These indices correspond to the parameters  $X_3$  (low to mid temperature conductivity) and  $X_5$  (specific heat capacity at  $T = 140^\circ\text{C}$ ). This is not surprising, because the system temperature is monotonously increasing and low temperature effects like these are expected to have a higher impact at earlier times. The Sobol' index  $S_3^T$  carries on to influence the simulation at later times, while  $S_5^T$  decreases in importance towards the end of the simulation.

Starting from  $t = 15$  min, the Sobol' indices  $S_1^T$ ,  $S_4^T$  and  $S_6^T$  become dominant. This can be explained similarly, since they influence the mid to high temperature behaviour.  $X_1$  is the start of the second key process and thus heavily influences the high temperature behaviour of all temperature-dependent material properties.  $X_4$  (high temperature conductivity) and  $X_6$  (specific heat capacity at the second key process) are the high temperature material properties and as such are more important at later time steps of the simulation.

The second key process results in a temperature plateau (like the one at  $100^\circ\text{C}$ ). Because this process occurs at temperatures higher than the maximum temperature reached at the measurement location, the process has no direct effect there. The system does, however, reach higher temperatures at locations closer to the exposed surface and it is there that the second key process influences the system behaviour. As can be seen from the results of the sensitivity analysis, the magnitude of the second key process ( $X_6$ ) significantly influences the temperatures at the measurement location due to the locally lower thermal diffusivity. However, the temperature where this peak in the thermal diffusivity appears ( $X_2$ ), is not so important.

#### 5.4 Validation of the calibration using V1-V2

So far, all presented results were related to the experiments *Test 1* (specimens E1-E4) described in Section 2.1. It was shown that the calibrated material properties for the examined materials can be used to conduct computer simulations that agree well with the same experimental observations that were used for calibration (Section 5.1). Before these calibrated material properties can be used to make predictions about the material behaviour in other experimental setups, validation experiments need to be carried out to judge the accuracy for the new intended use (Oberkampf and Roy, 2010).

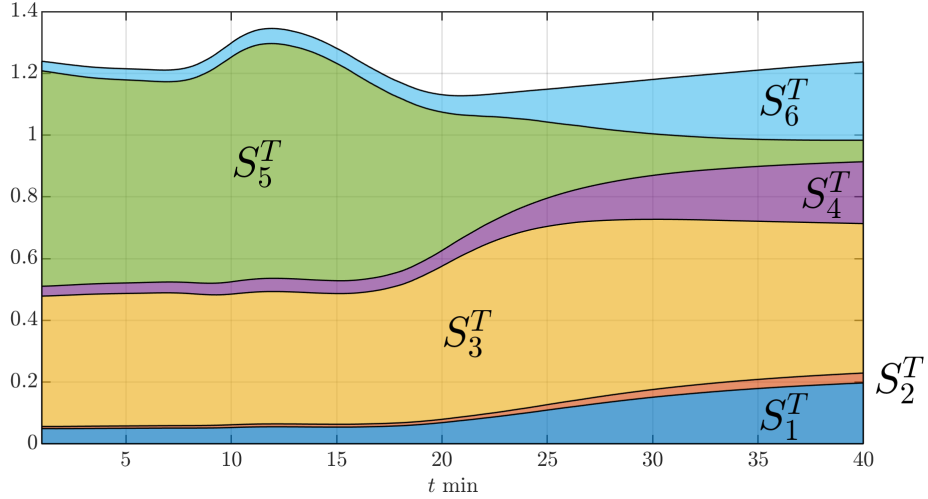


Figure 8: Time-dependent total Sobol' indices  $S_i^T$  for the surrogate model of Product A (E1).

In this work, the results from Test 2 (specimens V1 and V2) are used for this purpose (see Figure 3(e) and 3(f)). These tests were conducted using two of the materials for which calibrated material properties were obtained (Product C (E3) and Product D (E4)). To validate the calibration results, the posterior predictive distribution of this setup needs to be computed. It is defined identically to Eq. (9) and samples from it can be drawn by

$$\mathbf{Y}^{\text{postpred}} \sim \mathcal{N}(\mathbf{y} | \mathcal{M}_{\text{Test } 2}(\mathbf{x}_{\mathcal{M}}), \mathbf{\Sigma}(\mathbf{x}_{\varepsilon})), \quad \text{where} \quad \mathbf{x}_{\mathcal{M}}^{\text{post}} \sim \pi(\mathbf{x}_{\mathcal{M}} | \mathbf{y}_{\text{Test } 1}), \quad (41)$$

where the subscripts  $\text{Test } 1$  and  $\text{Test } 2$  were introduced to distinguish between quantities and models belonging to the respective setups. In this sense,  $\mathcal{M}_{\text{Test } 2}(\mathbf{X}_{\mathcal{M}})$  refers to the finite element model predicting the heat evolution in the two interfaces of Test 2 and  $\mathbf{y}_{\text{Test } 1}$  are measurements from Test 1.

Using parameters calibrated in Test 1 for the predictions in Test 2 requires careful consideration of the implications. On the one hand, the model parameters  $\mathbf{X}_{\mathcal{M}}$  of the insulation products can be reused without further considerations (since this is the point of the validation). On the other hand, parameters related to the discrepancy model  $\mathbf{X}_{\varepsilon}$  cannot be transferred so easily.

The discrepancy term captures measurement noise and model inadequacy and is assumed to follow a zero mean normal distribution (Eq. (4)). It cannot be directly applied in drawing predictive quantities as neither the measurement noise nor the model inadequacy can be expected to be identical between the two setups. Therefore, simplifying assumptions about the discrepancy covariance matrix  $\mathbf{\Sigma}(\mathbf{x}_{\varepsilon})$  have to be made.

The only information available about the discrepancy covariance matrix are the calibration results of E1-E4 (Section 5.1). There it was parameterized as described in Eq. (37). Across all posterior parameter distributions, the MAP estimator of the correlation length parameter  $\theta = X_{14}$  was  $\sim 30$  s. The discrepancy standard deviation  $\sigma(t)$ , however, did not

yield such a uniform result as can be clearly seen from Figures 13 to 17. To still parameterize the covariance matrix in the predictive draws from Eq. (41), a conservative choice of  $\theta = 30$  s and a constant  $\sigma(t) = 10^\circ\text{C}$  is thus made. This corresponds to setting the discrepancy parameter vector to

$$\mathbf{x}_\varepsilon = (10^\circ\text{C}, 0, 0, 0, 0, 0, 0, 30 \text{ s}) \quad (42)$$

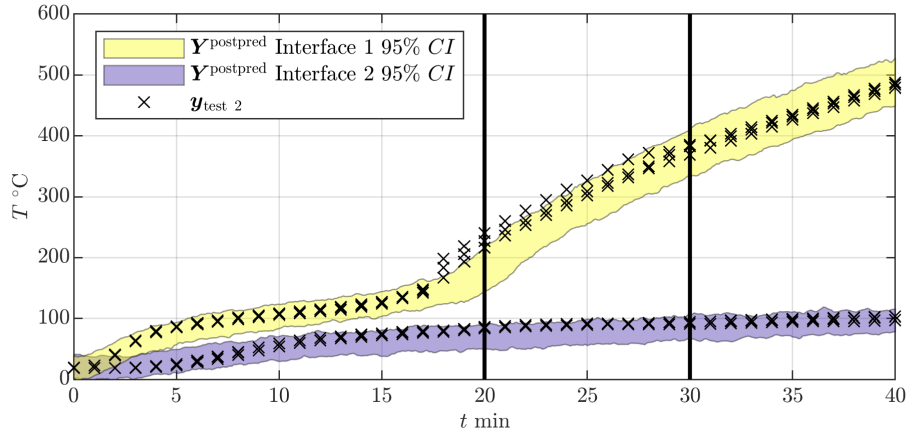
Figures 9 and 10 show the resulting confidence intervals from the posterior predictive distribution  $\mathbf{Y}^{\text{postpred}}$  defined in Eq. (41) for interface 1 and 2 of Test 2. To show the agreement of the predictions with the measurements,  $\mathbf{y}_{\text{Test } 2}$  are displayed as well. The figures contain one time-temperature plot in Figures (a) and summary statistics of the model predictions in Figure (b) at snapshot times  $t = 20$  min and  $t = 30$  min respectively.

Generally the simulations agree remarkably well with the experimental observations, but there are also time intervals, where small differences between the predicted and observed temperature evolutions are visible. The predictions at large time instants, which are of interest in practice, appear excellent.

## 6 Summary and conclusion

In this paper, a procedure to calibrate temperature-dependent effective material properties of fire insulation panels was presented. Available experimental temperature measurements were modelled using a 1D finite element heat transfer model. Because the associated material properties vary with temperature, they were parameterized using a set of model parameters. The actual calibration was then carried out using the well-known Bayesian inference framework. The necessary sampling from the posterior distribution was conducted with the advanced AIES (Affine invariant ensemble sampler) MCMC algorithm. In an effort to reduce the computational burden from the required repeated finite element simulations, a surrogate model of the heat transfer problem was constructed by combining polynomial chaos expansions (PCE) with the principal component analysis (PCA) technique. This surrogate model offered the possibility to additionally conduct a sensitivity analysis using the time-dependent Sobol' indices at no additional computational cost. Finally, the calibration was validated using a secondary set of experiments.

The proposed approach is superior to the previously used brute-force calibration approach because it automates the process, clearly defines the discrepancy and explicitly considers the uncertainties present in the model and measurements. Accordingly, it does not only deliver a single best fit property, but returns the full multivariate distribution of the calibrated properties and allows the computation of confidence intervals. Furthermore, the used Bayesian framework is a natural way to update information through the use of conditional random variables. It is well suited for engineering problems, where often expert knowledge is available that thereby can be directly integrated into the calibration procedure. However, it needs to



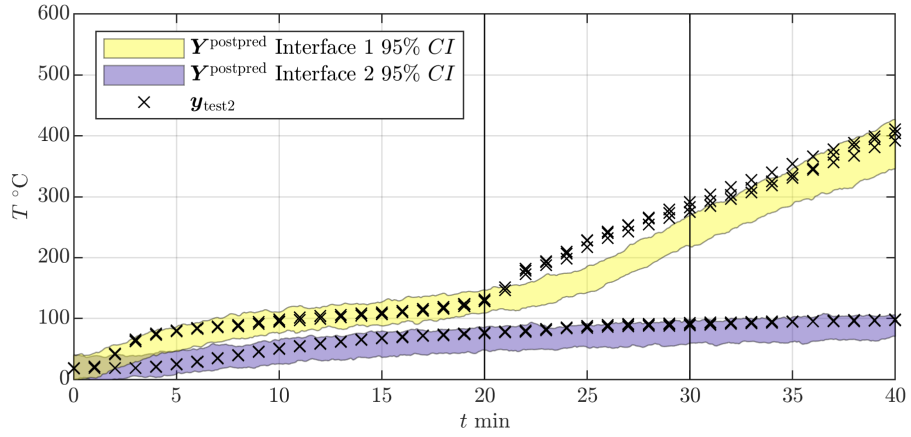
(a) Model predictions

	$t = 20$ min		$t = 30$ min	
( $^{\circ}\text{C}$ )	Interface 1	Interface 2	Interface 1	Interface 2
$y_t^{(s)}$	83.4	240.6	90.8	382.8
	85.8	227.1	94.5	369.0
	85.7	216.6	94.2	385.6
$\hat{\mu}$	70.1	185.4	84.0	370.7
$\hat{\sigma}$	10.7	18.1	10.2	20.3

(b) Statistics of model predictions at two snapshot times

Figure 9: Validation of the calibrated material properties of Product C and Product D using measurements from Test 2 (V1) including the empirical mean  $\hat{\mu}$  and the empirical standard deviation  $\hat{\sigma}$ .





(a) Model predictions

(°C)	$t = 20$ min		$t = 30$ min	
	Interface 1	Interface 2	Interface 1	Interface 2
$y_t^{(s)}$	75.3	129.0	88.9	283.0
	76.2	131.5	92.7	291.6
	77.2	127.9	90.4	274.8
$\hat{\mu}$	65.9	127.8	76.6	243.2
$\hat{\sigma}$	9.9	9.7	9.7	12.9

(b) Statistics of model predictions at two snapshot times

Figure 10: Validation of the calibrated material properties of Product D using measurements from Test 2 (V2) including the empirical mean  $\hat{\mu}$  and the empirical standard deviation  $\hat{\sigma}$ .

be considered that depending on the application less general calibration procedures might be more suitable (Mottershead et al., 2011; Patelli et al., 2017).

A valuable side product of the presented surrogate modelling technique is the *free* computation of the time-dependent Sobol' indices. This sensitivity analysis offers valuable insight into the time-dependent effect of the used parametrization.

Finally, it is worth emphasizing that the proposed method, which combines surrogate modelling (PCE) with dimensionality reduction (PCA), an advanced MCMC algorithm (AIES) and global sensitivity analysis is general and can be applied to any calibration problem involving complex computer codes. All the algorithms used are available in the UQLab uncertainty quantification software (Marelli and Sudret, 2014), especially the recently developed Bayesian inversion module (Wagner et al., 2019).

## Acknowledgement

The authors gratefully acknowledge the European COST Action FP1404 on the fire safe use of bio-based building products. Additionally, we would like to thank Prof. Alar Just and his team, as well as Dr. Joseph Nagel, who participated in the preliminary discussions about the case study definitions. The PhD thesis of the first author is supported by ETH grant #44 17-1.

## A Surrogate model approximation error

The following approximation error  $\tilde{\eta}$  was originally derived in Blatman and Sudret (2013). The surrogate model from Eq. (26) has a total  $L_2$ -approximation error that can be written by denoting the 2-norm as  $\|\cdot\|_2$  by

$$\varepsilon = \mathbb{E} \left[ \|\mathbf{Y} - \mathbf{Y}^{\text{PCA+PCE}}\|_2^2 \right] \quad (43)$$

$$= \mathbb{E} \left[ \|(\mathbf{Y} - \mathbf{Y}^{\text{PCA}}) + (\mathbf{Y}^{\text{PCA}} - \mathbf{Y}^{\text{PCA+PCE}})\|_2^2 \right]. \quad (44)$$

Through the Cauchy-Schwarz inequality this error is bounded by

$$\varepsilon \leq \left( \sqrt{\mathbb{E} \left[ \|\mathbf{Y} - \mathbf{Y}^{\text{PCA}}\|_2^2 \right]} + \sqrt{\mathbb{E} \left[ \|\mathbf{Y}^{\text{PCA}} - \mathbf{Y}^{\text{PCA+PCE}}\|_2^2 \right]} \right)^2 \quad (45)$$

$$\stackrel{\text{def}}{=} (\sqrt{\varepsilon_{\text{PCA}}} + \sqrt{\varepsilon_{\text{PCE}}})^2. \quad (46)$$

**PCA error**  $\varepsilon_{\text{PCA}}$ : this incorporates the error from estimating the mean  $\boldsymbol{\mu}_{\mathbf{Y}}$  and covariance matrix  $\boldsymbol{\Sigma}_{\mathbf{Y}}$  of the response as well as the dimensionality reduction error from leaving out  $N - N'$  dimensions. The former is neglected in this paper and for the latter the sum of the discarded eigenvalues  $\lambda_p$  can be directly used:

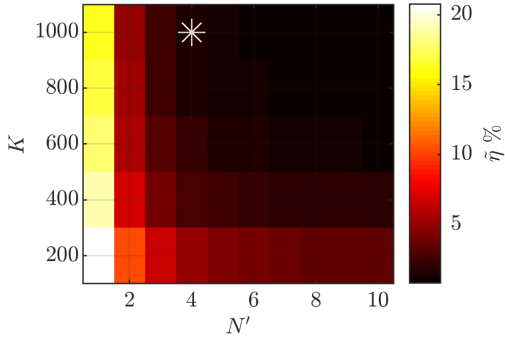
$$\varepsilon_{\text{PCA}} \approx \tilde{\varepsilon}_{\text{PCA}} = \sum_{p=N'+1}^N \lambda_p. \quad (47)$$

**PCE error**  $\varepsilon_{\text{PCE}}$ : this is the error of the polynomial chaos approximation. It can be estimated as the sum of the individual LOO errors for the  $N'$  scalar-valued principal component PCE's:

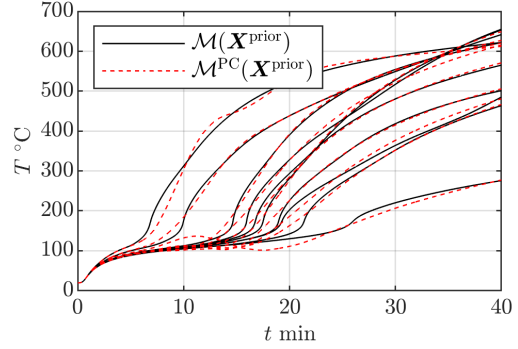
$$\varepsilon_{\text{PCE}} \approx \tilde{\varepsilon}_{\text{PCE}} = \sum_{p=1}^{N'} \varepsilon_{p,\text{LOO}}. \quad (48)$$

For practicality a relative error measure  $\eta$  is preferred. This can be obtained by dividing the estimator of the absolute error bound by an estimator of  $\mathbb{E} [\|\mathbf{Y}\|_2^2]$ , such as the trace of its estimated covariance matrix  $\text{Tr}(\boldsymbol{\Sigma}_{\mathbf{Y}})$ :

$$\tilde{\eta} = \frac{(\sqrt{\tilde{\varepsilon}_{\text{PCA}}} + \sqrt{\tilde{\varepsilon}_{\text{PCE}}})^2}{\text{Tr}(\boldsymbol{\Sigma}_{\mathbf{Y}})}. \quad (49)$$



(a) Relative error  $\tilde{\eta}$  of surrogate for increasing experimental design size  $K$  and included principal components  $N'$ .



(b) Comparison of forward model with surrogate model.

Figure 11: Convergence diagnostics for the surrogate model for setup E1, with  $K = 1,000$  and  $N' = 4$ .

Figure 11 shows the resulting error estimate for a set of experimental designs and included principal components using the setup E1. Additionally, the actual and surrogate model are run for a set of parameters and plotted for comparison.

## B Derivation of PCA-based Sobol' indices

This appendix contains the derivations for PCA-based Sobol' indices. To simplify the derivations, we introduce the notation  $\mathbf{X}_{\sim i} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_M)^\top$  to denote the random vector that contains all *but* the  $i$ -th random variable  $X_i$ . In the following derivations, the subscript in the expectation and variance operators  $\mathbb{E}_{X_i}$  and  $\text{Var}_{X_i}$  denotes the variable(s) with respect to which expectation and variance are computed, *i.e.*  $\mathbb{E}_{X_i}[\cdot] \stackrel{\text{def}}{=} \int(\cdot)\pi_i(x_i)dx_i$ .

The total Sobol' index  $S_{i,t}^T$  for the  $t$ -th element of a vector valued model output  $\mathbf{Y} = (Y_1, \dots, Y_N)^\top$  is defined as

$$S_{i,t}^T = \frac{\mathbb{E}_{\mathbf{X}_{\sim i}}[\text{Var}_{X_i}[Y_t]]}{\text{Var}[Y_t]} = 1 - \frac{\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]]}{\text{Var}[Y_t]}. \quad (50)$$

This can be used to write an expression for the variance of the expectation  $\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]]$  of the  $t$ -th model output as

$$\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]] = \mathbb{E}_{\mathbf{X}_{\sim i}}\left[(\mathbb{E}_{X_i}[Y_t])^2\right] - (\mathbb{E}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]])^2 \quad (51)$$

$$= \mathbb{E}_{\mathbf{X}_{\sim i}}\left[(\mathbb{E}_{X_i}[Y_t])^2\right] - (\mathbb{E}_{\mathbf{X}}[Y_t])^2. \quad (52)$$

By introducing the expression for the surrogate model from Eq. (27) and the fact that the expectation of the  $t$ -th response is  $\mathbb{E}_{\mathbf{X}}[Y_t] \approx \mu_{Y_t}$ , the following is obtained:

$$\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]] = \mathbb{E}_{\mathbf{X}_{\sim i}}\left[(\mathbb{E}_{X_i}[\mu_{Y_t} + \phi_t^{\text{row}} \mathbf{A}^\top \Psi(\mathbf{X})])^2\right] - \mu_{Y_t}^2 \quad (53)$$

$$= \mathbb{E}_{\mathbf{X}_{\sim i}}\left[(\mu_{Y_t} + \phi_t^{\text{row}} \mathbf{A}^\top \mathbb{E}_{X_i}[\Psi(\mathbf{X})])^2\right] - \mu_{Y_t}^2 \quad (54)$$

$$= \mathbb{E}_{\mathbf{X}_{\sim i}}\left[\mu_{Y_t}^2 + 2\mu_{Y_t} \phi_t^{\text{row}} \mathbf{A}^\top \mathbb{E}_{X_i}[\Psi(\mathbf{X})] + (\phi_t^{\text{row}} \mathbf{A}^\top \mathbb{E}_{X_i}[\Psi(\mathbf{X})])^2\right] - \mu_{Y_t}^2. \quad (55)$$

Because the expectation of all principal components vanishes ( $\mathbf{A}^\top \mathbb{E}[\Psi(\mathbf{X})] = \mathbb{E}[\mathbf{Z}] = \mathbf{0}$ ), one can write

$$\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]] = \mathbb{E}_{\mathbf{X}_{\sim i}}\left[(\phi_t^{\text{row}} \mathbf{A}^\top \mathbb{E}_{X_i}[\Psi(\mathbf{X})])^2\right]. \quad (56)$$

By switching to the summation notation, this can also be written as

$$\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]] = \mathbb{E}_{\mathbf{X}_{\sim i}}\left[\left(\sum_{\alpha \in \mathcal{A}^*} \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p,\alpha} \mathbb{E}_{X_i}[\Psi_\alpha(\mathbf{X})]\right)^2\right]. \quad (57)$$

Because the inner sum is only over the coefficients  $\phi_{pt} \tilde{a}_{p,\alpha}$ , this can be further simplified by substituting  $c_\alpha \stackrel{\text{def}}{=} \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p,\alpha}$  to obtain:

$$\text{Var}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y_t]] = \mathbb{E}_{\mathbf{X}_{\sim i}}\left[\left(\sum_{\alpha \in \mathcal{A}^*} c_\alpha \mathbb{E}_{X_i}[\Psi_\alpha(\mathbf{X})]\right)^2\right] \quad (58)$$

$$= \mathbb{E}_{\mathbf{X}_{\sim i}}\left[\left(\sum_{\alpha \in \mathcal{A}^*} c_\alpha \mathbb{E}_{X_i}[\Psi_\alpha(\mathbf{X})]\right) \left(\sum_{\beta \in \mathcal{A}^*} c_\beta \mathbb{E}_{X_i}[\Psi_\beta(\mathbf{X})]\right)\right] \quad (59)$$

$$= \sum_{\alpha, \beta \in \mathcal{A}^*} c_\alpha c_\beta \mathbb{E}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[\Psi_\alpha(\mathbf{X})] \mathbb{E}_{X_i}[\Psi_\beta(\mathbf{X})]]. \quad (60)$$

Due to the orthonormality of the polynomial basis  $\{\Psi_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}^*}$ , the conditional expectation in this equation can be expressed analytically as

$$\mathbb{E}_{\mathbf{X} \sim_i} [\mathbb{E}_{X_i} [\Psi_{\boldsymbol{\alpha}}(\mathbf{X})] \mathbb{E}_{X_i} [\Psi_{\boldsymbol{\beta}}(\mathbf{X})]] = \delta_{\boldsymbol{\alpha}, \boldsymbol{\beta}, i}$$

$$\text{with } \delta_{\boldsymbol{\alpha}, \boldsymbol{\beta}, i} = \begin{cases} 1, & \text{if } \boldsymbol{\alpha} = \boldsymbol{\beta} \text{ and } \alpha_i = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (61)$$

Therefore, the variance of the conditional expectation from Eq. (50) becomes

$$\text{Var}_{\mathbf{X} \sim_i} [\mathbb{E}_{X_i} [Y_t]] = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{i=0}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p, \boldsymbol{\alpha}} \right)^2, \quad (62)$$

where  $\mathcal{A}_{i=0}^* = \{\boldsymbol{\alpha} \in \mathcal{A}^* : \alpha_i = 0\}$  is the subset that contains only those polynomials  $\Psi_{\boldsymbol{\alpha}}$  with  $\alpha_i = 0$ . For completeness the total variance in the denominator of Eq. (50) reads:

$$\text{Var} [Y_t] = \sum_{\boldsymbol{\alpha} \in \mathcal{A}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p, \boldsymbol{\alpha}} \right)^2. \quad (63)$$

The total PCA-based index for the  $t$ -th component of the output vector  $\mathbf{Y}$  is thus obtained by plugging these results into Eq. (50):

$$S_{i,t}^T = 1 - \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{A}_{i=0}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p, \boldsymbol{\alpha}} \right)^2}{\sum_{\boldsymbol{\alpha} \in \mathcal{A}^*} \left( \sum_{p=1}^{N'} \phi_{pt} \tilde{a}_{p, \boldsymbol{\alpha}} \right)^2}. \quad (64)$$

Table 6: Posterior statistics for the calibration with Product B (E2). The values are computed from the available posterior sample and include the MAP estimate, the empirical mean  $\hat{\mu}$ , the empirical 95% confidence interval, the empirical standard deviation  $\hat{\sigma}$ , and the empirical coefficient of variation c.o.v.  $\stackrel{\text{def}}{=} \hat{\sigma}/\hat{\mu}$ . The prior statistics are shown in Table 4.

	MAP	$\hat{\mu}$	95% conf. interval	$\hat{\sigma}$	c.o.v.
$X_1$	$5.38 \cdot 10^2$	$5.05 \cdot 10^2$	$[4.25 \cdot 10^2, 5.79 \cdot 10^2]$	$4.25 \cdot 10^1$	$8.41 \cdot 10^{-2}$
$X_2$	$3.97 \cdot 10^{-1}$	$4.02 \cdot 10^{-1}$	$[3.37 \cdot 10^{-1}, 4.75 \cdot 10^{-1}]$	$3.65 \cdot 10^{-2}$	$9.07 \cdot 10^{-2}$
$X_3$	0.224	0.226	[0.217, 0.239]	$5.39 \cdot 10^{-3}$	$2.38 \cdot 10^{-2}$
$X_4$	0.107	0.115	[0.1, 0.138]	$1.12 \cdot 10^{-2}$	$9.74 \cdot 10^{-2}$
$X_5$	$5.12 \cdot 10^4$	$5.16 \cdot 10^4$	$[4.93 \cdot 10^4, 5.37 \cdot 10^4]$	$1.14 \cdot 10^3$	$2.21 \cdot 10^{-2}$
$X_6$	$1.08 \cdot 10^3$	$1.08 \cdot 10^3$	$[1.00 \cdot 10^3, 1.30 \cdot 10^3]$	$8.20 \cdot 10^1$	$7.56 \cdot 10^{-2}$
$X_7$	8.80	8.83	[8.30, 9.40]	$2.83 \cdot 10^{-1}$	$3.21 \cdot 10^{-2}$
$X_8$	2.02	2.04	[2.30, 1.78]	$1.95 \cdot 10^{-1}$	$4.28 \cdot 10^{-2}$
$X_9$	1.17	1.15	[1.39, 0.91]	$1.81 \cdot 10^{-1}$	$0.70 \cdot 10^{-1}$
$X_{10}$	$5.17 \cdot 10^{-1}$	$5.34 \cdot 10^{-1}$	$[0.75, 3.39 \cdot 10^{-1}]$	$1.56 \cdot 10^{-1}$	$1.31 \cdot 10^{-1}$
$X_{11}$	4.27	4.29	[3.93, 4.67]	$1.93 \cdot 10^{-1}$	$4.50 \cdot 10^{-2}$
$X_{12}$	1.89	1.90	[2.14, 1.68]	$1.78 \cdot 10^{-1}$	$4.16 \cdot 10^{-2}$
$X_{13}$	$4.04 \cdot 10^{-1}$	$4.22 \cdot 10^{-1}$	$[2.04 \cdot 10^{-1}, 6.41 \cdot 10^{-1}]$	$1.11 \cdot 10^{-1}$	$2.63 \cdot 10^{-1}$
$X_{14}$	$2.64 \cdot 10^1$	$2.64 \cdot 10^1$	$[2.56 \cdot 10^1, 2.71 \cdot 10^1]$	$3.81 \cdot 10^{-1}$	$1.44 \cdot 10^{-2}$

## C Additional results

This section presents the calibration and sensitivity analysis results for Product B, Product C and Product D.

## References

- Dassault Systèmes (Ed.) (2017). *ABAQUS/Standard User's Manual, Version 6.14*.
- Arwade, S. R., M. Moradi, and A. Louhghalam (2010). Variance decomposition and global sensitivity for structural systems. *Engineering Structures* 32(1), 1–10.
- Askey, R. and J. Wilson (1985). Some basic hypergeometric polynomials that generalize Jacobi polynomials. *Memoirs of the American Mathematical Society* 54(319), 1–57.

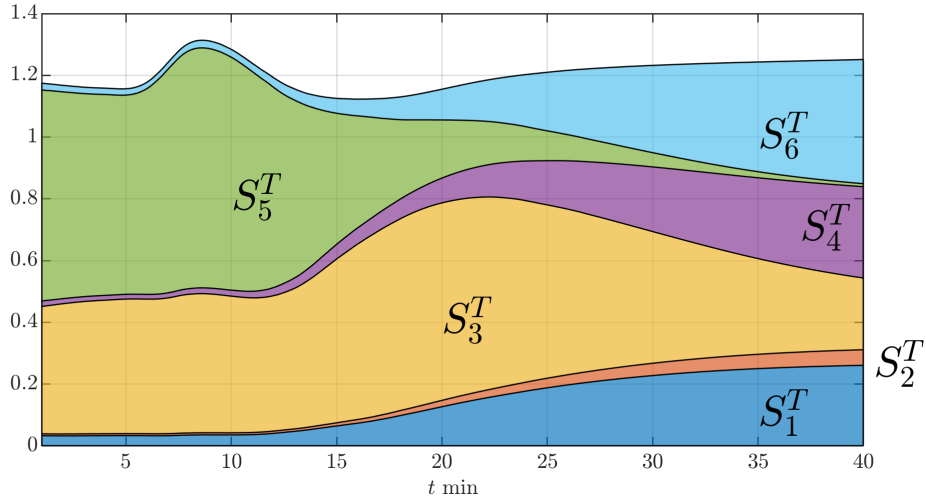


Figure 12: Time-dependent total Sobol' indices  $S_i^T$  for the surrogate model of Product B (E2).

Table 7: Posterior statistics for the calibration with Product C (E3). The values are computed from the available posterior sample and include the MAP estimate, the empirical mean  $\hat{\mu}$ , the empirical 95% confidence interval, the empirical standard deviation  $\hat{\sigma}$ , and the empirical coefficient of variation c.o.v.  $\stackrel{\text{def}}{=} \hat{\sigma}/\hat{\mu}$ . The prior statistics are shown in Table 4.

	MAP	$\hat{\mu}$	95% conf. interval	$\hat{\sigma}$	c.o.v.
$X_1$	$5.50 \cdot 10^2$	$5.63 \cdot 10^2$	$[4.63 \cdot 10^2, 7.10 \cdot 10^2]$	$6.25 \cdot 10^1$	$1.11 \cdot 10^{-1}$
$X_2$	$8.36 \cdot 10^{-1}$	$7.93 \cdot 10^{-1}$	$[4.24 \cdot 10^{-1}, 9.95 \cdot 10^{-1}]$	$1.60 \cdot 10^{-1}$	$2.01 \cdot 10^{-1}$
$X_3$	0.165	0.160	[0.133, 0.184]	$1.28 \cdot 10^{-2}$	$7.98 \cdot 10^{-2}$
$X_4$	0.648	0.701	[0.539, 0.93]	0.106	$1.51 \cdot 10^{-1}$
$X_5$	$2.16 \cdot 10^4$	$2.06 \cdot 10^4$	$[1.64 \cdot 10^4, 2.49 \cdot 10^4]$	$2.18 \cdot 10^3$	$1.06 \cdot 10^{-1}$
$X_6$	$1.05 \cdot 10^4$	$1.12 \cdot 10^4$	$[3.99 \cdot 10^3, 2.65 \cdot 10^4]$	$4.99 \cdot 10^3$	$4.44 \cdot 10^{-1}$
$X_7$	$1.08 \cdot 10^1$	$1.08 \cdot 10^1$	[9.95, $1.17 \cdot 10^1$ ]	$4.53 \cdot 10^{-1}$	$4.18 \cdot 10^{-2}$
$X_8$	1.43	1.39	[1.03, 1.72]	$1.80 \cdot 10^{-1}$	$1.30 \cdot 10^{-1}$
$X_9$	1.32	1.31	[1.56, 1.06]	$1.82 \cdot 10^{-1}$	$6.20 \cdot 10^{-2}$
$X_{10}$	1.62	1.62	[1.90, 1.36]	$2.00 \cdot 10^{-1}$	$5.51 \cdot 10^{-2}$
$X_{11}$	1.58	1.65	[1.37, 1.94]	$1.53 \cdot 10^{-1}$	$9.29 \cdot 10^{-2}$
$X_{12}$	2.15	2.24	[1.89, 2.63]	$1.94 \cdot 10^{-1}$	$8.64 \cdot 10^{-2}$
$X_{13}$	1.25	1.24	[1.43, 1.05]	$1.44 \cdot 10^{-1}$	$5.19 \cdot 10^{-2}$
$X_{14}$	$3.06 \cdot 10^1$	$3.06 \cdot 10^1$	$[2.94 \cdot 10^1, 3.18 \cdot 10^1]$	$5.93 \cdot 10^{-1}$	$1.94 \cdot 10^{-2}$

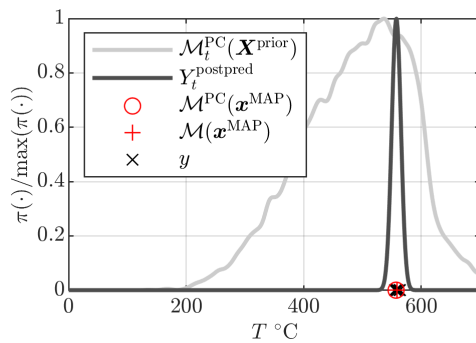
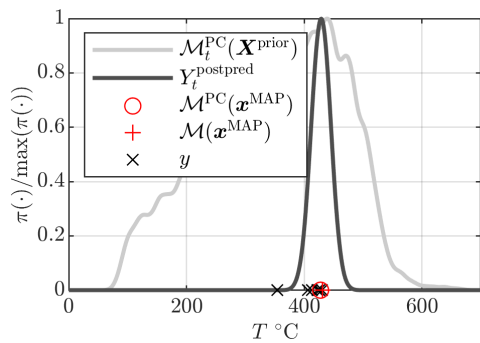
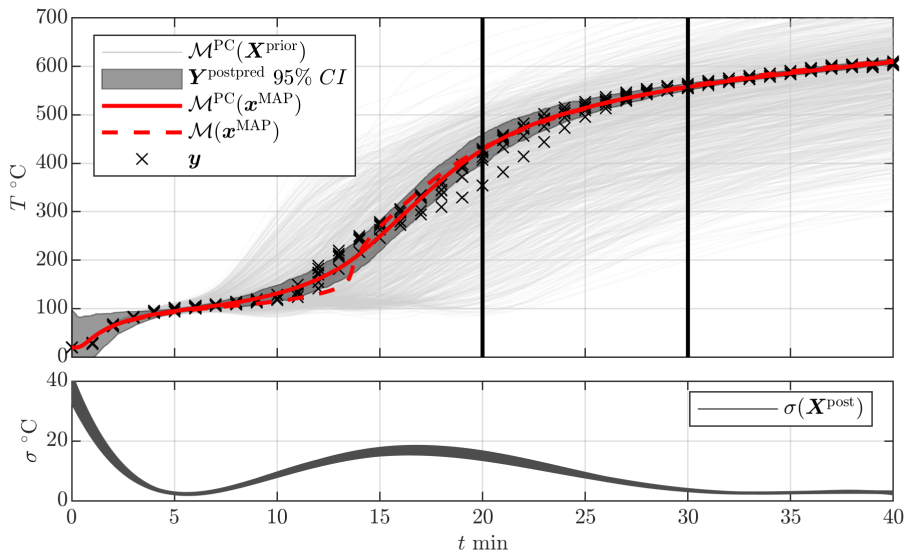
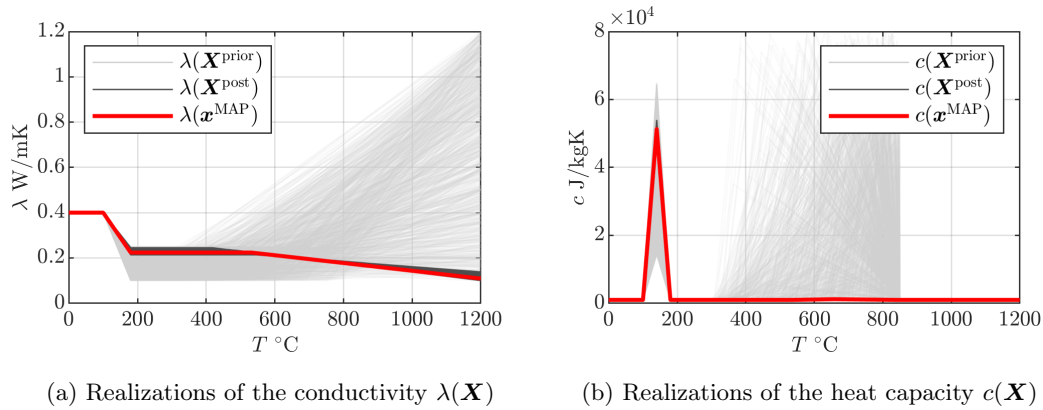


Figure 13: Calibration results for Product B 9.5 mm insulation (E2), experiments conducted by Just (2016).



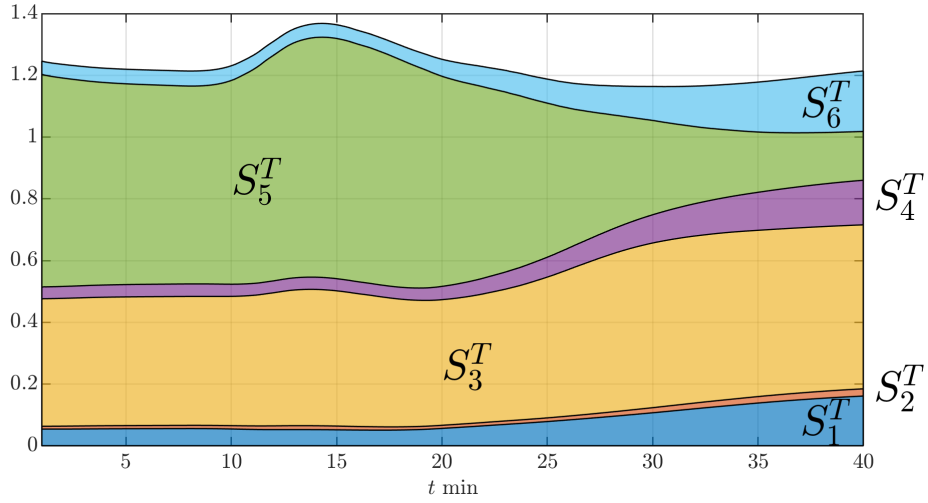


Figure 14: Time-dependent total Sobol' indices  $S_i^T$  for the surrogate model of Product C (E3).

Table 8: Posterior statistics for the calibration with Product D (E4). The values are computed from the available posterior sample and include the MAP estimate, the empirical mean  $\hat{\mu}$ , the empirical 95% confidence interval, the empirical standard deviation  $\hat{\sigma}$ , and the empirical coefficient of variation c.o.v.  $\stackrel{\text{def}}{=} \hat{\sigma}/\hat{\mu}$ . The prior statistics are shown in Table 4.

	MAP	$\hat{\mu}$	95% conf. interval	$\hat{\sigma}$	c.o.v.
$X_1$	$7.07 \cdot 10^2$	$6.90 \cdot 10^2$	$[6.41 \cdot 10^2, 7.43 \cdot 10^2]$	$2.62 \cdot 10^1$	$3.79 \cdot 10^{-2}$
$X_2$	$9.82 \cdot 10^{-1}$	$8.07 \cdot 10^{-1}$	$[5.42 \cdot 10^{-1}, 9.95 \cdot 10^{-1}]$	$1.25 \cdot 10^{-1}$	$1.55 \cdot 10^{-1}$
$X_3$	0.157	0.155	$[0.14, 0.173]$	$9.21 \cdot 10^{-3}$	$5.93 \cdot 10^{-2}$
$X_4$	1.16	1.06	$[0.843, 1.2]$	$9.75 \cdot 10^{-2}$	$9.16 \cdot 10^{-2}$
$X_5$	$2.04 \cdot 10^4$	$2.08 \cdot 10^4$	$[1.81 \cdot 10^4, 2.42 \cdot 10^4]$	$1.70 \cdot 10^3$	$8.16 \cdot 10^{-2}$
$X_6$	$5.18 \cdot 10^3$	$4.91 \cdot 10^3$	$[1.60 \cdot 10^3, 8.17 \cdot 10^3]$	$1.64 \cdot 10^3$	$3.34 \cdot 10^{-1}$
$X_7$	8.15	8.39	$[7.77, 8.99]$	$3.28 \cdot 10^{-1}$	$3.91 \cdot 10^{-2}$
$X_8$	2.88	2.97	$[2.61, 3.33]$	$1.86 \cdot 10^{-1}$	$6.28 \cdot 10^{-2}$
$X_9$	$6.76 \cdot 10^{-1}$	$7.09 \cdot 10^{-1}$	$[4.65 \cdot 10^{-1}, 9.73 \cdot 10^{-1}]$	$1.31 \cdot 10^{-1}$	$1.84 \cdot 10^{-1}$
$X_{10}$	1.67	1.69	$[1.90, 1.49]$	$1.57 \cdot 10^{-1}$	$4.14 \cdot 10^{-2}$
$X_{11}$	0.76	0.75	$[0.91, 5.75 \cdot 10^{-1}]$	$1.25 \cdot 10^{-1}$	$0.75 \cdot 10^{-1}$
$X_{12}$	1.38	1.41	$[1.21, 1.64]$	$1.09 \cdot 10^{-1}$	$7.75 \cdot 10^{-2}$
$X_{13}$	$5.10 \cdot 10^{-1}$	$5.35 \cdot 10^{-1}$	$[6.55 \cdot 10^{-1}, 4.15 \cdot 10^{-1}]$	$9.24 \cdot 10^{-2}$	$0.77 \cdot 10^{-1}$
$X_{14}$	$2.96 \cdot 10^1$	$2.97 \cdot 10^1$	$[2.88 \cdot 10^1, 3.07 \cdot 10^1]$	$4.73 \cdot 10^{-1}$	$1.59 \cdot 10^{-2}$

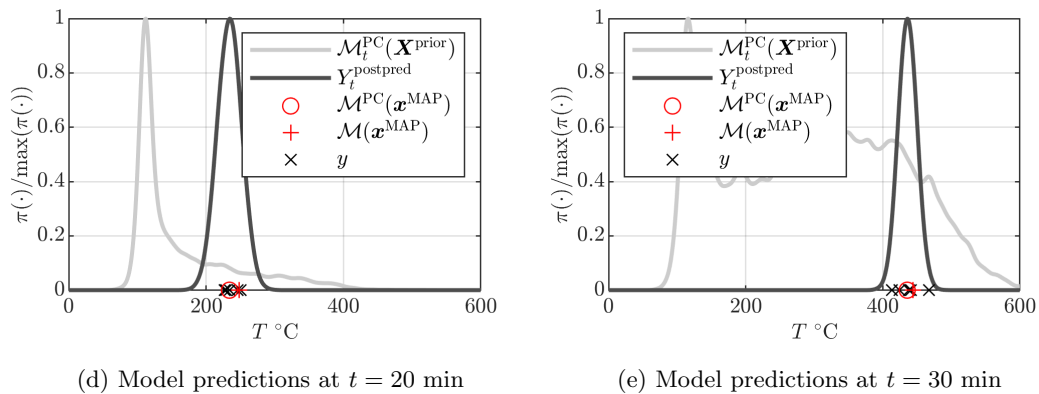
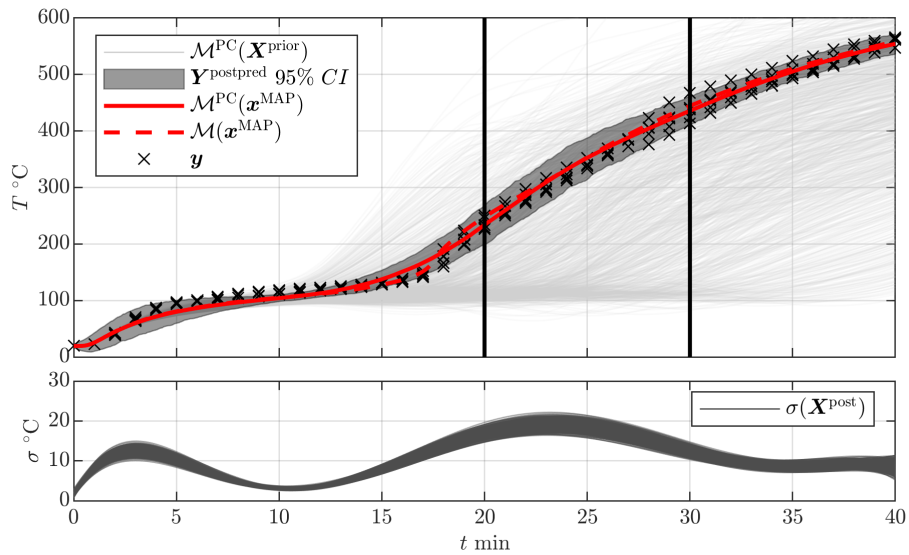
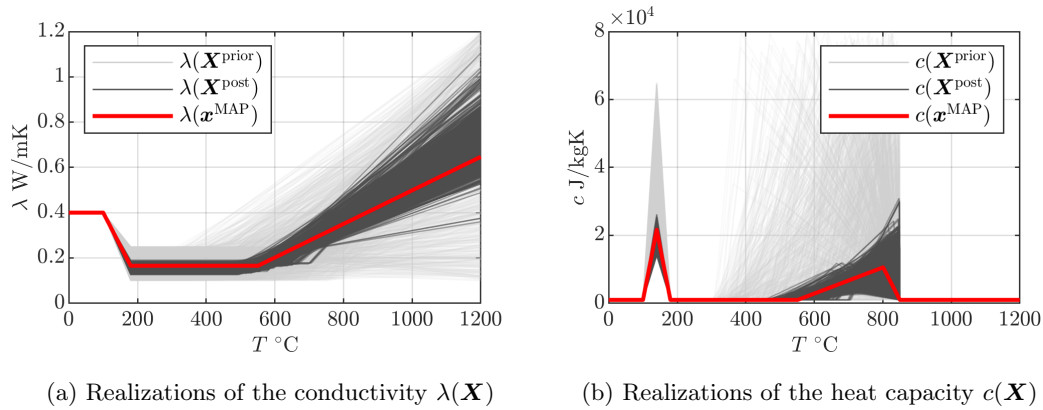


Figure 15: Calibration results for Product C 12.5 mm insulation (E3), experiments conducted by Just (2016).

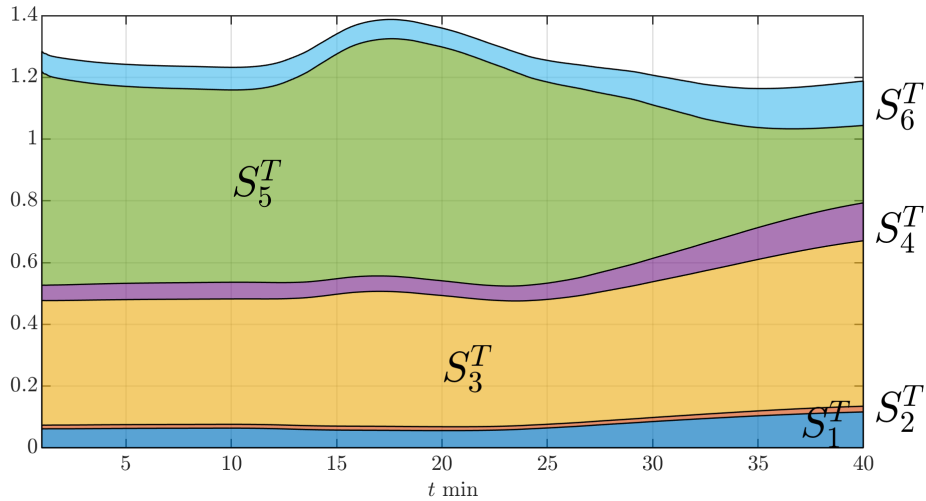


Figure 16: Time-dependent total Sobol' indices  $S_i^T$  for the surrogate model of Product D (E4).

Beck, J. L. and L. S. Katafygiotis (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechanics* 124(4), 455–461.

Berveiller, M., B. Sudret, and M. Lemaire (2006). Stochastic finite elements: a non intrusive approach by regression. *European Journal of Computational Mechanics* 15(1-3), 81–92.

Blatman, G. (2009). *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*. Ph. D. thesis, Université Blaise Pascal, Clermont-Ferrand.

Blatman, G. and B. Sudret (2010). An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics* 25, 183–197.

Blatman, G. and B. Sudret (2011a). Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *Journal of Computational Physics* 230, 2345–2367.

Blatman, G. and B. Sudret (2011b). Principal component analysis and Least Angle Regression in spectral stochastic finite element analysis. In M. Faber, J. Köhler, and K. Nishijima (Eds.), *Proceedings 11th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP11), Zurich, Switzerland*.

Blatman, G. and B. Sudret (2013). Sparse polynomial chaos expansions of vector-valued response quantities. In G. Deodatis (Ed.), *Proceedings 11th International Conference on Structural Safety and Reliability (ICOSSAR'2013), New York, USA*.

Breu, R. D. (2016). Improved component additive method for the separating function - development of a testing and calculation procedure. Master's thesis, ETH Zürich, Switzerland.

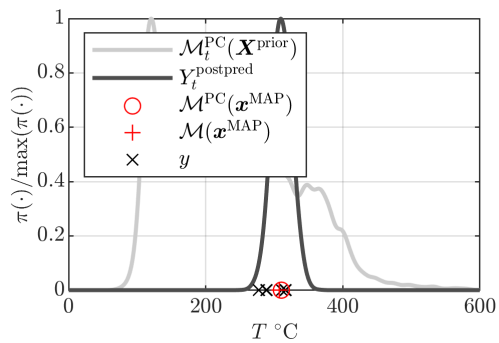
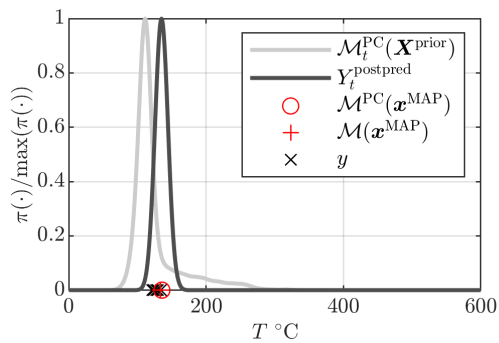
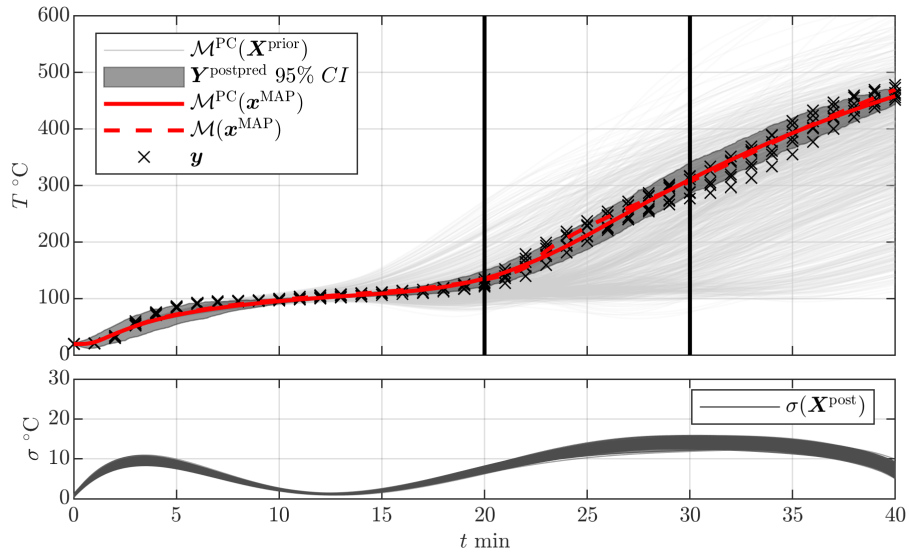
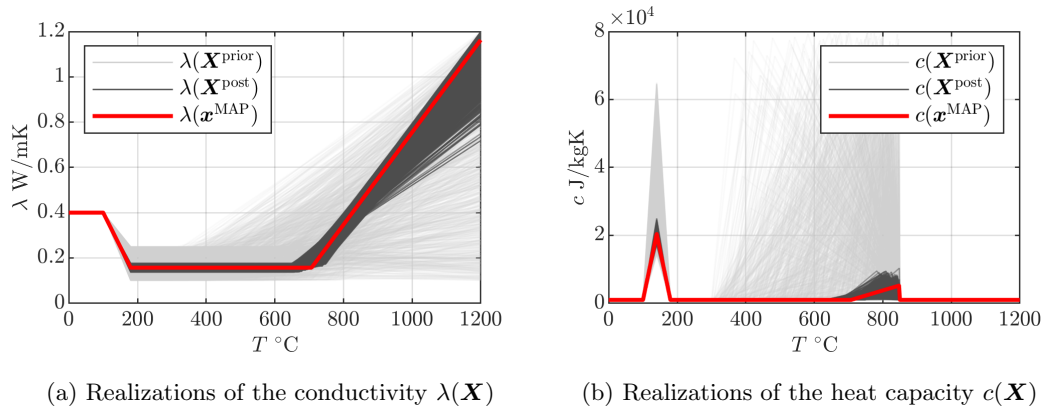


Figure 17: Calibration results for Product D 15 mm insulation (E4), experiments conducted by Just (2016).

- Choi, S., R. Grandhi, R. Canfield, and C. Pettit (2004). Polynomial chaos expansion with Latin Hypercube sampling for estimating response variability. *AIAA Journal* 45, 1191–1198.
- EN 1363-1:2012 (2012). Feuerwiderstandsprüfungen - Teil 1: Allgemeine Anforderungen. Technical report, Deutsches Institut für Normung, Berlin, DE.
- EN 1991-1-2:2002 (2002). Eurocode 1: Actions on structures - Part 1-2: General actions - Actions on structures exposed to fire. Technical report, European Committee for Standardization, Brussels, BE.
- Frangi, A., V. Schleifer, and M. Fontana (2010). Design model for the verification of the separating function of light timber frame assemblies. *Engineering Structures* 32(4), 1184–1195.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2014). *Bayesian Data Analysis* (3 ed.). Texts in Statistical Science. CRC Press.
- Gelman, A., X.-L. Meng, and H. Stern (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 6(4), 733–760.
- Goodman, J. and J. Weare (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science* 5(1), 65–80.
- Guo, X., D. Dias, C. Carvajal, L. Peyras, and P. Breul (2018). Reliability analysis of embankment dam sliding stability using the sparse polynomial chaos expansion. *Engineering Structures* 174(1), 295–307.
- Haario, H., E. Saksman, and J. Tamminen (2001). An adaptive Metropolis algorithm. *Bernoulli* 7(2), 223–242.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57(1), 97–109.
- ISO 834-1:1999 (1999). Fire-resistance tests - Elements of building construction. Technical report, International Organization for Standardization, Geneva, CH.
- Jolliffe, I. T. (2002). *Principal Component Analysis* (2 ed.). Springer Series in Statistics. Springer-Verlag.
- Just, A. (2016). Model scale fire tests of four gypsum plasterboards of Gyproc and stone wool. Test report 5P08165-03, SP Technical Research Institute of Sweden.
- Just, A. and J. Schmid (2018). Guidance for implementation of materials and products in fire design methods of timber frame assemblies. Technical report, COST Action FP1404, Zürich, Switzerland.

- MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.
- Mäger, K. N., A. Just, J. Schmid, N. Werther, M. Klippel, D. Brandon, and A. Frangi (2017). Procedure for implementing new materials to the component additive method. *Fire Safety Journal* 107, 149–160.
- Marelli, S. and B. Sudret (2014). UQLab: A framework for uncertainty quantification in Matlab. In *Vulnerability, Uncertainty, and Risk (Proceedings 2nd International Conference on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom)*, pp. 2554–2563.
- Marelli, S. and B. Sudret (2015). Compressive polynomial chaos expansion for multi-dimensional model maps. In T. Haukaas (Ed.), *Proceedings 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12), Vancouver, Canada*. Paper #209.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* 21(6), 1087–1092.
- Mottershead, J. E., M. Link, and M. I. Friswell (2011). The sensitivity method in finite element model updating: A tutorial. *Mechanical Systems and Signal Processing* 25(7), 2275–2296.
- Nagel, J., J. Rieckermann, and B. Sudret. Uncertainty quantification in urban drainage simulation: Fast surrogates for sensitivity analysis and model calibration. Submitted.
- Oberkampf, W. and C. Roy (2010). *Verification and Validation in Scientific Computing*. Cambridge University Press.
- Oberkampf, W., T. Trucano, and C. Hirsch (2004). Verification, validation, and predictive capability in computational engineering and physics. *Applied Mechanics Reviews* 57(5), 345–384.
- Patelli, E., Y. Govers, M. Broggi, H. Gomes, M. Link, and J. Mottershead (2017). Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data. *Archive of Applied Mechanics* 87, 905–925.
- Roberts, G. O. and J. S. Rosenthal (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics* 18(2), 349–367.
- Rosky, P. J., J. D. Doll, and H. L. Friedmann (1978). Brownian dynamics as smart Monte Carlo simulation. *The Journal of Chemical Physics* 69(10), 4628–4633.

- Saltelli, A., K. Chan, and E. Scott (Eds.) (2000). *Sensitivity analysis*. J. Wiley & Sons.
- Schleifer, V. (2009). *Zum Verhalten von raumabschliessenden mehrschichtigen Holzbauteilen im Brandfall*. Ph. D. thesis, ETH Zürich, Switzerland.
- Sobol', I. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling & Computational Experiment 1*, 407–414.
- Soize, C. and R. Ghanem (2004). Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing 26*(2), 395–410.
- Sudret, B. (2006). Global sensitivity analysis using polynomial chaos expansions. In P. Spanos and G. Deodatis (Eds.), *Proceedings 5th International Conference on Computational Stochastic Mechanics (CSM5)*, Rhodos, Greece.
- Sudret, B. (2007). Uncertainty propagation and sensitivity analysis in mechanical models – contributions to structural reliability and stochastic spectral methods. Technical report. Habilitation à diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France (229 pages).
- Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety 93*, 964–979.
- van der Maaten, L., E. Postma, and H. van den Herik (2008). Dimensionality reduction: a comparative review. *Journal of Machine Learning Research 10*, 66–71.
- Wagner, P.-R., J. Nagel, S. Marelli, and B. Sudret (2019). UQLab user manual – Bayesian inference for model calibration and inverse problems. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich. Report # UQLab-V1.2-113.
- Xiu, D. and G. Karniadakis (2002). The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing 24*(2), 619–644.
- Yu, B., R. Tang, and B. Li (2019). Probabilistic calibration for development length models of deformed reinforcing bar. *Engineering Structures 182*(1), 279–289.