

Bayesian inference to identify parameters in viscoelasticity

Hussein Rappel · Lars A. A. Beex · Stéphane
P. A. Bordas

Received: date / Accepted: date

Abstract This contribution discusses Bayesian inference (BI) as an approach to identify parameters in viscoelasticity. The aims are (i) to show that the prior has a substantial influence for viscoelasticity, (ii) to show that this influence decreases for an increasing number of measurements and (iii) to show how different types of experiments influence the identified parameters and their uncertainties. The standard linear solid model is the material description of interest and a relaxation test, a constant strain-rate test and a creep test are the tensile experiments focused on. The experimental data are artificially created, allowing us to make a one-to-one comparison between the input parameters and the identified parameter values. Besides dealing with the aforementioned aims, we believe that this contribution forms a comprehensible start for those interested in applying BI in viscoelasticity.

Keywords Bayesian inference · Bayes' theorem · Statistical identification · Parameter identification · Viscoelasticity

H. Rappel

Faculty of Science, Technology and Communication, University of Luxembourg, Campus Kirchberg, 6, Rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg.

Computational & Multiscale Mechanics of Materials (CM3), Department of Aerospace and Mechanical Engineering, University of Liège, Quartier Polytech 1, Allée de la Découverte 9, B-4000 Liège, Belgium.
E-mail: hussein.rappel@uni.lu

L.A.A. Beex

Faculty of Science, Technology and Communication, University of Luxembourg, Campus Kirchberg, 6, Rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg.

E-mail: lars.beex@uni.lu

S.P.A. Bordas

Faculty of Science, Technology and Communication, University of Luxembourg, Campus Kirchberg, 6, Rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg.

School of Engineering, Cardiff University, Queens Buildings, The Parade, Cardiff CF243AA, Wales, UK.
Adjunct Professor, Intelligent Systems for Medicine Laboratory School of Mechanical and Chemical Engineering, The University of Western Australia, 35 Stirling Highway, Crawley/Perth WA 6009, Australia.
E-mail: stephane.bordas@uni.lu

1 Introduction

A frequently employed approach to identify model parameters from a set of experimental data is the least squares method (LSM, see e.g. Beex et al. 2013). The standard LSM formulates the squared difference between the measured data and the model response in terms of the parameters of interest and minimises this difference with respect to the parameters of interest. Those parameters that minimise the difference are the resulting parameter values. The residual is a measure for how well the response fits the measurement data, but does not give insight in the uncertainty of the identified parameter values.

Some enhanced LSM formulations incorporate the statistical information of the measurement noise, in contrast to the standard LSM (Magorou et al. 2002; Genovese et al. 2005). Linearised approximations are however needed to extract the most probable parameter values and the covariance (i.e. a measure that shows how the random variables depend on each other). These approaches furthermore assume that the measurement noise is symmetrically distributed with a constant variance. In other words, they assume that the noise is independent of the measured value and that a particular positive noise realisation is equally possible as a negative realisation of the same magnitude (Elster and Wübbeler 2016).

Bayesian inference (BI) can be used to formulate an alternative identification approach, which may account for the fact that only a limited number of measurement data is available. It achieves this by incorporating an *a-priori* assumption on the parameter values. In practise, this entails that the user has to specify a prior distribution (i.e. the prior). The influence of this prior distribution can influence the identified parameter values significantly, but its influence decreases for an increasing number of measurement data (Madireddy et al. 2015; Rappel et al. 2016).

An intermediate result of BI is a probability density function (PDF) in terms of the parameters of interest, called the posterior distribution (or the posterior). Once the posterior distribution is established, its statistical characteristics need to be evaluated, e.g. the mean, the covariance matrix and the parameter values at which the PDF is maximum (called the ‘maximum-a-posteriori-probability’ or ‘MAP’ point). Most often, numerical methods like Markov chain Monte Carlo (MCMC, Higdon et al. 2002; Wang and Zabarar 2004) are necessary to extract the statistical characteristics from the posterior distribution. For posteriors with C_2 -continuity, Laplace approximations can be applied to evaluate the quantities of the interest. The posterior is then approximated by a Gaussian distribution (Beck and Katafygiotis 1998; Oh et al. 2008).

The application of BI to identify (elastic) material parameters was started by Isenberg (1979), to the best of our knowledge. For two decades afterwards, BI was not used for material parameter identification. When the developments started again, it was amongst others used for the identification of elastic constants from dynamic responses (Alvin 1997; Beck and Katafygiotis 1998; Marwala and Sibusiso 2005; Daghia et al. 2007; Abhinav and Manohar 2015), the elastic constants of composite and laminate plates (Lai and Ip 1996; Daghia et al. 2007; Nichols et al. 2010; Gogu et al. 2013) and spatially varying elastic constants (Koutsourelakis 2012). An introduction to identify Young’s moduli using BI is presented in Gogu et al. (2010).

Approaches based on BI are also used to identify material parameters in elastoplasticity. Some examples are the studies of Fitzenz et al. (2007), Muto and Beck (2008), Most (2010), Rosić et al. (2013) and Liu and Au (2013). Rappel et al. (2016) have recently written an introduction of how BI can be employed for elastoplastic models. In another study, Sarkar et al. (2012) used the Bayesian approach to identify thermodynamical parameters of cementitious materials.

These days, BI is also used as an approach to evaluate the quality of different mechanical models with respect to each other, to which is referred as 'model selection'. Some examples are the studies of Madireddy et al. (2015) for hyperelastic constitutive models for tissue, Oden et al. (2013) for phenomenological models for tumour growth, Chiachío et al. (2015) for models of damage progression in composites due to fatigue, and Babuška et al. (2016) for fatigue descriptions of metals. Model selection is however out of the scope of the current study.

BI is also used to identify viscoelastic parameters. Zhang et al. (2013) used BI to characterise the Young's modulus of a viscoelastic polymer layer in a laminated structure. Their work includes a validation. Mehrez et al. (2015) employed BI to identify the viscoelastic properties of aged and unaged asphalt. Miles et al. (2015) applied BI to characterise the viscoelastic parameters of a dielectric elastomer undergoing finite deformation. Hernandez et al. (2015) employed the Bayesian approach for the probabilistic identification of five viscoelastic parameters. Zhao and Pelegri (2016) used BI together with a finite element model to identify the time constant of a Voigt-based tissue model. All mentioned studies (except Zhao and Pelegri 2016) are based on actual experimental data. Kenz et al. (2013) compared asymptotic theory, bootstrapping, and Bayesian estimation for a viscoelastic wave propagation model. Also in this study, experimental data was used (for a homogeneous tissue-mimicking gel).

None of the mentioned BI studies for viscoelasticity have focused on the influence of the prior distribution however. Furthermore, none of them have investigated the influence of the number of measurement data. Except for Zhao and Pelegri (2016) none of the studies for viscoelasticity are able to compare the resulting values with the input values. Finally, none of the aforementioned studies have compared the estimated parameter values for different types of tests (e.g. relaxation test, constant strain rate test and creep test). Hence, BI was used in the past to identify viscoelastic parameters, but if other prior distributions were selected, the number of measurements were different or other tests were used, the results of the aforementioned studies would be different.

The aim of this study is to show how viscoelastic parameters identified by BI are dependent on the selected prior distribution, the number of measurement data and the type of test. The standard linear solid (SLS) model is the employed viscoelastic description and a tensile relaxation test, a tensile constant strain-rate test and the creep test are the experiments of interest. The measurements are generated artificially, including the noise realisations. This allows us to compare the identified values with the input values and with the LSM values, in a controlled and thorough manner. In order to investigate if better parameter values are obtained by different types of tests, the values identified based on different tests are also compared to each other.

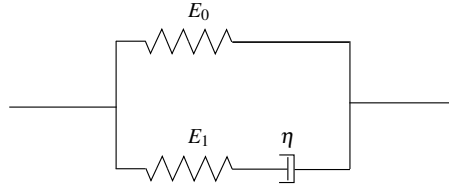


Fig. 1: The one-dimensional spring-dashpot representation of the SLS model.

The structure of the paper is as follows. Section 2 briefly discusses the SLS model. Section 3 presents the theoretical fundamentals of BI, the final expression of the posterior distribution for the SLS model and the three types of experiments. Section 4 discusses the MCMC method to numerically extract the most probable parameter values from the posterior distribution (including the associated covariances). Section 5 briefly discusses the LSM and genetic optimisation as the minimisation approach to find the parameter values identified by LSM. Section 6 presents the examples. This contribution is closed with some conclusions (section 7).

2 Material model

Viscoelasticity can be used to describe the stress-strain-time relation of time-dependent materials. These materials show both viscous and elastic behaviour under deformation (Banks et al. 2011). Well-known viscoelastic material models are the Maxwell model, the Kelvin-Voigt model, the standard linear solid (SLS) model and the generalised Maxwell model (Banks et al. 2011). As the goal of this contribution is to identify the parameters of the SLS model, the remainder of this section considers this description.

The SLS model describes stress-relaxation and creep phenomena in viscoelastic systems with only one rate-dependent parameter (Orosz et al. 1997). In one dimension, the model can schematically be represented using two springs and a dashpot (see Fig. 1).

For uniaxial tension, the stress-strain-time relation of the SLS can be described by the following differential equation (Liu et al. 2011):

$$\sigma + \frac{\eta}{E_1} \frac{\partial \sigma}{\partial t} = E_0 \varepsilon + (E_0 + E_1) \frac{\eta}{E_1} \frac{\partial \varepsilon}{\partial t}, \quad (1)$$

where σ is the stress, ε is the strain, $\frac{\partial}{\partial t}$ denotes the derivative with respect to time, E_0 is the stiffness of the parallel spring, E_1 is the stiffness of the spring in series with the dashpot and η is the viscosity of the dashpot (see Fig. 1).

In a tensile relaxation experiment in which strain ε_0 is applied infinitely fast and then kept constant, the stress-time relation during relaxation resulting from Eq. (1) reads:

$$\sigma(\mathbf{x}, t) = E_0 \varepsilon_0 + E_1 \varepsilon_0 \exp\left(-\frac{E_1 t}{\eta}\right), \quad (2)$$

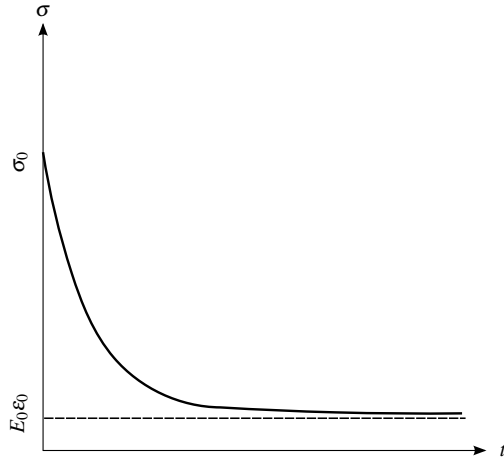


Fig. 2: A schematic stress-time curve for one-dimensional stress relaxation of the SLS model. It is assumed that applied strain ϵ_0 is prescribed infinitely fast, resulting in $\sigma_0 = E_0\epsilon_0 + E_1\epsilon_0$.

where $\mathbf{x} = [E_0 \ E_1 \ \eta]^T$ is the parameter vector. Fig. 2 shows the relaxation behaviour of Eq. (2) schematically.

For a constant strain-rate experiment in tension, in which the specimen is elongated with a constant strain-rate $\dot{\epsilon}_0$, the stress-time relation reads:

$$\sigma(\mathbf{x}, t) = E_0\dot{\epsilon}_0 t + \eta\dot{\epsilon}_0 \left(1 - \exp\left(-\frac{E_1 t}{\eta}\right)\right). \quad (3)$$

The stress-time response of the constant strain-rate test is shown in Fig. 3 schematically.

In a creep test in tension a constant tensile stress σ_0 is applied to the specimen infinitely fast, which is then kept constant. Strain is the quantity that is measured in this experiment. The strain-time relation for a creep test of the SLS model reads:

$$\epsilon(\mathbf{x}, t) = \frac{\sigma_0}{E_0} + \sigma_0 \left(\frac{1}{E_0 + E_1} - \frac{1}{E_0}\right) \exp\left(-\frac{E_0 E_1 t}{\eta(E_0 + E_1)}\right). \quad (4)$$

Fig. 4 shows the behaviour of Eq. (4) schematically.

3 Bayesian inference

Identification approaches based on BI are less common than those based on the LSM. An intermediate result of Bayesian approaches is a probability density function (PDF) for the parameters to be identified (the so-called ‘posterior distribution’ or ‘posterior’). The posterior includes one’s prior knowledge about the parameters, as a PDF (the so-called ‘prior distribution’ or ‘prior’). Numerical techniques are most often required to extract the mean, the variance and MAP point from the posterior. In this contribution, MCMC is employed for this, which is discussed in section 4.

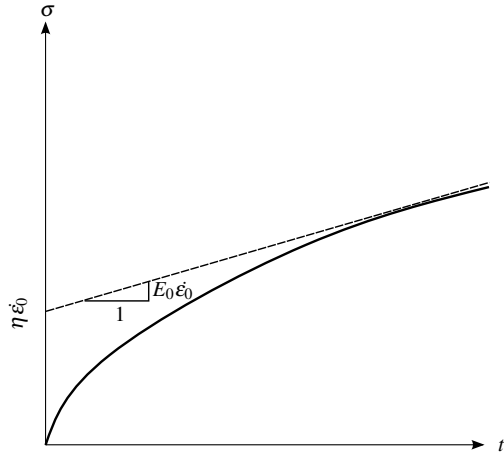


Fig. 3: A schematic stress-time curve for a one-dimensional constant strain-rate test of the SLS model. The applied strain-rate is $\dot{\epsilon}_0$.

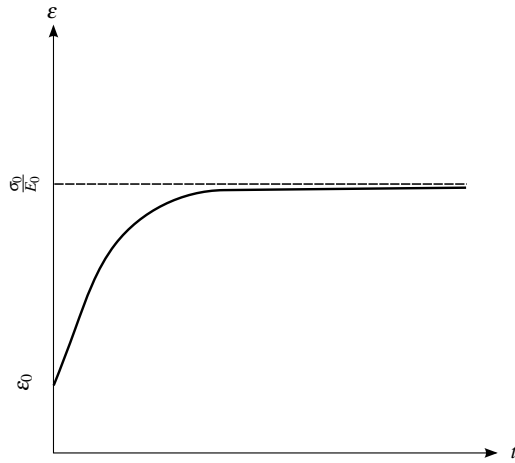


Fig. 4: A schematic strain-time curve for a one-dimensional creep test in tension of the SLS model. It is assumed that the constant stress σ_0 is applied infinitely fast and then remains constant. The initial strain is $\epsilon_0 = \frac{\sigma_0}{E_0 + E_1}$.

This section focuses on the fundamentals of BI for continuous events (subsection 3.1) and BI for the specific application to the SLS model and the tensile uniaxial relaxation, constant strain-rate and creep tests (subsection 3.2).

Throughout this section capital letters denote random variables and bold ones denote vectors *and* matrices. An important issue for this contribution is to recognise that we only consider a statistical noise in the measured output, not in the experimental input. We thus consider the time measurements to be exact. In Rappel et al. (2016), we have treated double uncertainties however.

3.1 Fundamentals of Bayesian inference

3.1.1 Bayes' theorem

Consider two continuous random variables $\mathbf{X} \in \mathbb{R}^n$ and $\mathbf{Y} \in \mathbb{R}^k$ with associated PDFs $\pi(\mathbf{x})$ and $\pi(\mathbf{y})$. Joint probability density function $\pi(\mathbf{x}, \mathbf{y})$ expresses the probability that *both* \mathbf{x} and \mathbf{y} occur. This joint PDF can be expanded as:

$$\pi(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x}) = \pi(\mathbf{y})\pi(\mathbf{x}|\mathbf{y}), \quad (5)$$

where $\pi(\mathbf{y}|\mathbf{x})$ and $\pi(\mathbf{x}|\mathbf{y})$ are conditional PDFs. Conditional PDF $\pi(\mathbf{x}|\mathbf{y})$ expresses the probability that \mathbf{x} occurs, if it is certain that \mathbf{y} occurs. Using Eq. (5), the simplest form of Bayes' theorem can be written as:

$$\pi(\mathbf{x}|\mathbf{y}) = \frac{\pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x})}{\pi(\mathbf{y})}. \quad (6)$$

If we now consider \mathbf{x} to be the vector with n identifiable parameters and \mathbf{y} to be the vector with k measurements, then $\pi(\mathbf{x})$, $\pi(\mathbf{y}|\mathbf{x})$ and $\pi(\mathbf{x}|\mathbf{y})$ are the prior distribution (i.e. the PDF that represents the original belief or prior knowledge), the likelihood function (i.e. the PDF of the measured data \mathbf{y} , given unknown parameters \mathbf{x}) and the posterior distribution (i.e. the PDF of unknown parameters \mathbf{x} , given measured data \mathbf{y}). Furthermore, $\pi(\mathbf{y})$ is called the evidence and can be calculated using the law of total probabilities (Ulrych et al. 2001) as follows:

$$\pi(\mathbf{y}) = \int_{\mathbb{R}^n} \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x})d\mathbf{x}. \quad (7)$$

As the measured data (\mathbf{y}) is already known after the experiment, the evidence in Eq. (6) ($\pi(\mathbf{y})$) is a constant number ($C \in \mathbb{R}^+$). It can thus be considered as a normalisation factor that ensures that the integral of the posterior ($\pi(\mathbf{x}|\mathbf{y})$) equals 1.

As the MAP point, the mean and the covariance matrix of the posterior are independent of this normalisation factor, it suffices to rewrite Eq. (6) as follows:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{x})\pi(\mathbf{y}|\mathbf{x}). \quad (8)$$

Eq. (8) is the expression for the posterior in which we are interested. Based on Eq. (8) it may be clear that in order to obtain the posterior distribution ($\pi(\mathbf{x}|\mathbf{y})$), one needs to define the prior ($\pi(\mathbf{x})$) and the likelihood function ($\pi(\mathbf{y}|\mathbf{x})$). The prior is the PDF that expresses one's *a-priori* knowledge. It can have a significant influence on the estimated MAP point, mean value and covariance matrix if the number of measurement data is small (see e.g. Kaipio and Somersalo 2006; Rappel et al. 2016 and the results section of this contribution). On the other hand, if the number of measurement data is large, the influence of the prior decreases (Madireddy et al. 2015).

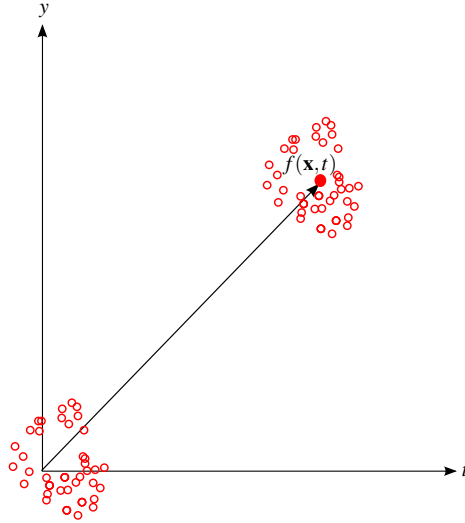


Fig. 5: Schematic of additive noise. The noise around the known theoretical response $f(\mathbf{x}, t)$ is independent of $f(\mathbf{x}, t)$ (Calvetti and Somersalo 2007). In this representation, an error in the time measurements is included as well.

3.1.2 Likelihood function

In order to construct the likelihood function, one needs to formulate the noise model (i.e. the uncertainty model) and determine the type of noise distribution (π_{noise}), including its parameters. This is often based on a calibration procedure of the experimental equipment. The most common noise model in the literature is the additive noise model, in which the noise distribution is independent of the theoretical model ($f(\mathbf{x}, t)$ in Fig. 5). Consequently, the noise only shifts around the theoretical model (see Fig. 5 again).

The additive noise model is also employed in this work. For a consideration of other types of noise models, the readers are referred to Kaipio and Somersalo (2006).

Given $\mathbf{X} \in \mathbb{R}^n$ as the vector with the parameters to be identified, $\mathbf{Y} \in \mathbb{R}^k$ as the vector with the measurements, $\boldsymbol{\Omega} \in \mathbb{R}^k$ as the noise vector, $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ as the model dependent on the unknown parameters and using the additive noise model, one can write the relation between the measurements and the noise (i.e. the error in the measurements) as follows:

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}) + \boldsymbol{\Omega}. \quad (9)$$

It can be noted that the measurements, \mathbf{Y} , are made at a specific experimental input \mathbf{t} . In the *relaxation* and *constant strain-rate* tests, the *stress* measurements (\mathbf{Y}) can be recognised as the experimental output and the *strain* measurements for the *creep* test. The experimental input can be recognised as the time at which each stress or strain measurement is made (\mathbf{t}). In Eq. (9) we should thus have written $\mathbf{f}(\mathbf{X}|\mathbf{t})$. We have however chosen not to do this in order to simplify the notation.

If realisation $\mathbf{X} = \mathbf{x}$ is given for Eq. (9), the only random variable on the right-hand side is noise vector ($\boldsymbol{\Omega}$). Assuming that the noise vector and the unknown parameters are statistically independent, the noise distribution is the likelihood function:

$$\pi(\mathbf{y}|\mathbf{x}) = \pi_{\text{noise}}(\mathbf{y} - \mathbf{f}(\mathbf{x})), \quad (10)$$

where $\pi_{\text{noise}}(\boldsymbol{\omega})$ is the noise PDF (based on calibration experiments, see subsection 6.1) and $\boldsymbol{\Omega} = \boldsymbol{\omega}$ is the noise realisation. Substituting Eq. (10) in Eq. (8) yields the posterior:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{x})\pi_{\text{noise}}(\mathbf{y} - \mathbf{f}(\mathbf{x})). \quad (11)$$

Once the posterior is established, one can calculate the statistical quantities of interest (e.g. the MAP point, the mean and the covariance matrix).

3.2 Bayesian inference for the uniaxial tensile relaxation, constant strain-rate and creep tests

In this subsection, we will employ the theory of the previous subsection to construct the posterior for the one-dimensional SLS model and our experiments of interest in uniaxial tension.

The parameters of the SLS model to be identified are E_0 , E_1 and η . Hence, the parameter vector thus reads $\mathbf{x} = [E_0 \ E_1 \ \eta]^T$.

First we start with the relaxation test, in which the model ($\mathbf{f}(\mathbf{X})$ in Eq. (9)) is given by Eq. (2) for one measurement. Employing these relations and considering only one measurement point (measurement point i), Eq. (9) reads:

$$y_i = E_0 \varepsilon_0 + E_1 \varepsilon_0 \exp\left(-\frac{E_1 t_i}{\eta}\right) + \omega_i, \quad (12)$$

where y_i is the measured stress, t_i is the time at which y_i is measured and ω_i originates from realisation $\boldsymbol{\Omega} = \boldsymbol{\omega}$.

To identify the statistical distribution of the measurement noise, calibration experiments must be performed (discussed in subsection 6.1). For now we assume that these calibration experiments have exposed that a normal distribution with standard deviation S_{noise} can be used:

$$\pi_{\text{noise}}(\boldsymbol{\omega}) = \frac{1}{\sqrt{2\pi}S_{\text{noise}}} \exp\left(-\frac{\boldsymbol{\omega}^2}{2S_{\text{noise}}^2}\right), \quad (13)$$

where $\boldsymbol{\Omega} = \boldsymbol{\omega}$ is a realisation for the noise drawn from the distribution given by Eq. (13).

Using Eqs. (10) and (12) the likelihood function for measurement point i reads:

$$\pi(y_i|\mathbf{x}) = \pi_{\text{noise}}(y_i - E_0 \varepsilon_0 - E_1 \varepsilon_0 \exp(-\frac{E_1 t_i}{\eta})) = \frac{1}{\sqrt{2\pi}S_{\text{noise}}} \exp\left(-\frac{\left(y_i - E_0 \varepsilon_0 - E_1 \varepsilon_0 \exp(-\frac{E_1 t_i}{\eta})\right)^2}{2S_{\text{noise}}^2}\right). \quad (14)$$

The next step is to choose the prior distribution ($\pi(\mathbf{x})$). As the material parameters cannot be negative, a modified normal distribution is selected:

$$\pi(\mathbf{x}) \propto \begin{cases} \exp\left(-\frac{(\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})}{2}\right) & \text{if } E_0 \geq 0 \text{ and } E_1 \geq 0 \text{ and } \eta \geq 0, \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

where $\bar{\mathbf{x}}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{prior}}$ are the mean vector and the covariance matrix of the prior, respectively. The values in $\bar{\mathbf{x}}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{prior}}$ must be selected by the user. After all, they are assumptions (as is the shape of the priori distribution).

Substituting Eqs. (14) and (15) in Eq. (8) yields the posterior for measurement point i as follows:

$$\pi(\mathbf{x}|y_i) \propto \exp\left(-\left[\frac{(\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})}{2} + \frac{\left(y_i - E_0 \varepsilon_0 - E_1 \varepsilon_0 \exp(-\frac{E_1 t_i}{\eta})\right)^2}{2S_{\text{noise}}^2}\right]\right). \quad (16)$$

Since the posterior for measurement point i can serve as the prior for measurement point $i + 1$, the posterior for k measurement points reads:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\left[\frac{(\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})}{2} + \frac{\sum_{i=1}^k \left(y_i - E_0 \varepsilon_0 - E_1 \varepsilon_0 \exp(-\frac{E_1 t_i}{\eta})\right)^2}{2S_{\text{noise}}^2}\right]\right), \quad (17)$$

where $\pi(\mathbf{x}|\mathbf{y}) = \pi(\mathbf{x}|y_1, y_2, \dots, y_k)$.

For the constant strain-rate test the only change is the employed model ($\mathbf{f}(\mathbf{X})$ in Eq. (9)) which is given by Eq. (3). Hence, for the same noise distribution (Eq. (13)) and the prior given in Eq. (15), the posterior distribution for k stress measurements reads:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\left[\frac{(\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})}{2} + \frac{\sum_{i=1}^k \left(y_i - E_0 \dot{\epsilon}_0 t_i - \eta \dot{\epsilon}_0 \left(1 - \exp\left(-\frac{E_1 t_i}{\eta}\right)\right)\right)^2}{2S_{\text{noise}}^2}\right]\right), \quad (18)$$

where $\dot{\epsilon}_0$ is the constant applied strain-rate and $\pi(\mathbf{x}|\mathbf{y}) = \pi(\mathbf{x}|y_1, y_2, \dots, y_k)$.

As was mentioned previously the output for the creep test (\mathbf{Y} in Eq. (9)) is the strain measurements. The model $\mathbf{f}(\mathbf{X})$ for this case is given by Eq. (4). We furthermore assume that the noise distribution is of the same for as for the previous two cases (Eq. (13)). Hence, using Eqs. (4), (11), (13) and (15) the final form of the posterior for the creep test with k strain measurements reads:

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\left[\frac{(\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{\text{prior}})}{2} + \frac{\sum_{i=1}^k \left(y_i - \frac{\sigma_0}{E_0} - \sigma_0 \left(\frac{1}{E_0 + E_1} - \frac{1}{E_0}\right) \exp\left(-\frac{E_0 E_1 t_i}{\eta(E_0 + E_1)}\right)\right)^2}{2S_{\text{noise}}^2}\right]\right), \quad (19)$$

where y_i is the measured strain at time t_i , \mathbf{x} is the parameter vector $[E_0 \ E_1 \ \eta]^T$ and σ_0 is the constant applied tensile stress.

Now the posterior is established for all three cases (relaxation, constant strain-rate and creep test), we will use the MCMC approach to identify the mean parameter values, the parameter values at which the posterior is maximum (i.e. the ‘MAP’ point) and the covariance matrix of the posterior. The covariance matrix is a measure for the possible spread of the parameter values, as well as how each parameter value depends on the others.

4 Markov chain Monte Carlo method (MCMC)

After establishing the posterior, one needs to determine the statistical properties of the posterior, such as the MAP point, the mean and the covariance matrix. These statistical quantities can be found analytically for straightforward cases, but often the use of a numerical technique cannot be avoided. One such a technique is Markov chain Monte Carlo sampling (Beck and Au 2002; Marzouk et al. 2007; Kristensen and Zabarar 2014). MCMC techniques are based on drawing samples from a target distribution (here the posterior) and numerically approximating the quantities of interest (e.g. the mean). Below, the Monte Carlo method and the adaptive Metropolis algorithm are discussed as means to obtain appropriate samples.

4.1 Monte Carlo method

The main goal of the Monte Carlo approach is to numerically approximate integrals of the following form:

$$\mathbf{I} = \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x})\pi(\mathbf{x})d\mathbf{x}, \quad (20)$$

where π is the target distribution (here posterior) and $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is an integrable function over \mathbb{R}^n . Having a set of samples $\{\mathbf{x}_i\}_i^N$ drawn from target PDF $\pi(\mathbf{x})$, the integral \mathbf{I} can be numerically approximated as follows:

$$\hat{\mathbf{I}} = \frac{1}{N} \sum_{i=1}^N \mathbf{g}(\mathbf{x}_i), \quad (21)$$

where $\hat{\mathbf{I}}$ represents the numerical approximation of \mathbf{I} . The law of large numbers entails that $\hat{\mathbf{I}}$ converges as follows (Andrieu et al. 2003):

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N \mathbf{g}(\mathbf{x}_i) = \mathbf{I}. \quad (22)$$

The numerical approximation of the components of the covariance matrix for $\mathbf{g}(\mathbf{x})$ ($\widehat{\mathbf{cov}}_{\mathbf{g}}$) reads (Brooks et al. 2011):

$$(\widehat{\mathbf{cov}}_{\mathbf{g}})_{jm} = \frac{1}{N} \sum_{i=1}^N \left(g_j(\mathbf{x}_i) - I_j \right) \left(g_m(\mathbf{x}_i) - I_m \right), \quad j = 1, 2, \dots, k, \quad m = 1, 2, \dots, k. \quad (23)$$

In our case, we substitute $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ and $\pi = \pi_{post}$ in the previous equations in order to obtain the following relation for the mean of the posterior:

$$\bar{\mathbf{x}}_{post} = \int_{\mathbb{R}^n} \mathbf{x}\pi_{post}(\mathbf{x})d\mathbf{x} = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i. \quad (24)$$

The components of the covariance matrix of the posterior ($\widehat{\mathbf{cov}}_{post}$) are furthermore approximated as follows:

$$(\widehat{\mathbf{cov}}_{post})_{jm} = \frac{1}{N} \sum_{i=1}^N \left((x_i)_j - (\bar{x}_{post})_j \right) \left((x_i)_m - (\bar{x}_{post})_m \right), \quad j = 1, 2, \dots, k, \quad m = 1, 2, \dots, k. \quad (25)$$

Assuming that the number of the drawn samples is large (i.e. N is large), the MAP point can be approximated as follows (Andrieu et al. 2003):

$$\widehat{\mathbf{MAP}} = \underset{\mathbf{x}_{i=1, \dots, N}}{\operatorname{argmax}} \pi(\mathbf{x}_i). \quad (26)$$

In next subsection the standard and adaptive Metropolis algorithms are discussed as means to obtain appropriate samples. The adaptive Metropolis algorithm is employed in this contribution as the sampling approach.

4.2 The standard Metropolis algorithm

The Metropolis-Hastings approach is an often employed technique for sampling (Andrieu et al. 2003). Its goal is to explore the target PDF (i.e. the posterior) by making a random walk through parameter space \mathbf{x} . The procedure accomplishes this by basing each new sample on the previous sample.

We start by the consideration of sample \mathbf{x}_i . The posterior is evaluated for this sample (resulting in $\pi_{\text{post}}(\mathbf{x}_i)$). Subsequently, a new sample \mathbf{x}_p is proposed using a proposal distribution (q in Algorithm 1). If the value of the PDF at the proposed sample is larger than the PDF at the previous sample or they are equal ($\pi_{\text{post}}(\mathbf{x}_i) \leq \pi_{\text{post}}(\mathbf{x}_p)$), the proposed sample will be accepted as the new sample. Otherwise, the sample will be accepted *or* rejected, based on the ratio of the PDF evaluated at the proposed and current sample (r in Algorithm 1) and its comparison with a random number generated from a uniform PDF. If the ratio is less than the random number drawn from the uniform distribution, the proposed sample is rejected. If the ratio is larger, it is accepted. Algorithm 1 shows the standard Metropolis-Hastings algorithm in practice.

Algorithm 1 The standard Metropolis-Hastings algorithm

```

1: select the initial sample  $\mathbf{x}_0 \in \mathbb{R}^n$  and  $\gamma$ 
2: for  $i = 0, 1, 2, \dots, N-1$  do
3:   draw  $\mathbf{x}_p \in \mathbb{R}^n$  from the proposal distribution  $q(\mathbf{x}_p|\mathbf{x}_i)$  in Eq. (29)
4:   calculate the ratio  $r(\mathbf{x}_i, \mathbf{x}_p) = \min\left(1, \frac{\pi(\mathbf{x}_p)q(\mathbf{x}_i|\mathbf{x}_p)}{\pi(\mathbf{x}_i)q(\mathbf{x}_p|\mathbf{x}_i)}\right)$ 
       $\triangleright \pi(\cdot)$  denotes the target distribution (i.e. posterior).
5:   draw  $u \in [0, 1]$  from uniform probability density
6:   if  $r(\mathbf{x}_i, \mathbf{x}_p) \geq u$  then
7:      $\mathbf{x}_{i+1} = \mathbf{x}_p$ 
8:   else
9:      $\mathbf{x}_{i+1} = \mathbf{x}_i$ 
10:  end if
11: end for

```

For a symmetric proposal distribution (i.e. transition kernel, q in Algorithm 1), one can write:

$$q(\mathbf{x}_i|\mathbf{x}_p) = q(\mathbf{x}_p|\mathbf{x}_i). \quad (27)$$

Using Eq. (27), line 4 of the algorithm can then be rewritten as:

$$r(\mathbf{x}_i, \mathbf{x}_p) = \min\left(1, \frac{\pi(\mathbf{x}_p)}{\pi(\mathbf{x}_i)}\right). \quad (28)$$

The convergence and stability of the algorithm can be checked by focusing for instance on the evolution of the mean value and the covariance matrix as a function of the number of the drawn samples (Sarkar et al. 2012). The approximated distribution converges to the posterior if the approximated statistical quantities remain the same for an increase of the number of samples.

Factors that influence the efficiency of the algorithm are the proposal distribution (q) and the initial sample (\mathbf{x}_0 , Kaipio and Somersalo 2006). The most common proposal distribution for the Metropolis-Hastings algorithm (as employed in this work) is a normal distribution:

$$q(\mathbf{x}_i|\mathbf{x}_p) = q(\mathbf{x}_p|\mathbf{x}_i) \propto \exp\left(-\frac{1}{2\gamma^2} \|\mathbf{x}_i - \mathbf{x}_p\|^2\right), \quad (29)$$

where γ is the standard deviation, which can be tuned to increase the efficiency. In the work of Gelman et al. (1996), $\gamma = \frac{2.38}{\sqrt{n}}$ is given as an efficient starting value, where n is the dimension of the posterior (i.e. the number of unknown parameters).

4.3 The adaptive Metropolis algorithm

The adaptive proposal (AP) is introduced by Haario et al. (1999) to automatically tune γ and by doing so, increase the efficiency of the Metropolis algorithm. AP bases its updates on the knowledge of the posterior, gathered from previously drawn samples. The proposal distribution in the adaptive Metropolis algorithm reads:

$$q(\mathbf{x}_p|\mathbf{x}_i) \sim N(\mathbf{x}_i, \gamma^2 \mathbf{R}_i), \quad (30)$$

where $N(\mathbf{x}_i, \gamma^2 \mathbf{R}_i)$ denotes a normal distribution with mean \mathbf{x}_i and covariance matrix $\gamma^2 \mathbf{R}_i$, of size $n \times n$. γ is the initially selected standard deviation and \mathbf{R}_i is updated based on the previous samples. Matrix \mathbf{R}_i is established by first storing all i previous samples in matrix \mathbf{K} of size $i \times n$. \mathbf{R}_i is then computed as:

$$\mathbf{R}_i = \frac{1}{i-1} \tilde{\mathbf{K}}^T \tilde{\mathbf{K}}, \quad (31)$$

where $\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{K}_{\text{mean}}$ and \mathbf{K}_{mean} reads:

$$\mathbf{K}_{\text{mean}} = \begin{bmatrix} \mathbf{k}_{\text{mean}} \\ \mathbf{k}_{\text{mean}} \\ \vdots \\ \mathbf{k}_{\text{mean}} \end{bmatrix}_{i \times n}, \quad (32)$$

with \mathbf{k}_{mean} as a $1 \times n$ vector of the following form:

$$\mathbf{k}_{\text{mean}} = \frac{1}{i} \left[\sum_{j=1}^i (K)_{j1} \quad \sum_{j=1}^i (K)_{j2} \quad \cdots \quad \sum_{j=1}^i (K)_{jn} \right]. \quad (33)$$

In terms of $\tilde{\mathbf{K}}$, one can rewrite the proposal distribution as follows:

$$N(\mathbf{x}_i, \gamma^2 \mathbf{R}_i) = \mathbf{x}_i + \frac{\gamma}{\sqrt{i-1}} \tilde{\mathbf{K}}^T N(0, \mathbf{I}_i), \quad (34)$$

where \mathbf{I}_i is the identity matrix of size $i \times i$ and $N(0, \mathbf{I}_i)$ denotes an i -dimensional normal distribution. In this contribution, the proposal distribution is updated once per thousand samples, as it is inefficient to update the proposal distribution too often. Algorithm 2 shows the Metropolis-Hastings algorithm with the symmetric AP proposal (Eq. (34)).

Algorithm 2 The Metropolis-Hastings algorithm with symmetric AP proposal

```

1: select the initial sample  $\mathbf{x}_0 \in \mathbb{R}^n$  and set  $\gamma = \frac{2.38}{\sqrt{n}}$ 
2: for  $i = 0, 1, 2, \dots, N - 1$  do
3:   draw  $\mathbf{x}_p \in \mathbb{R}^n$  from the proposal distribution  $q(\mathbf{x}_p | \mathbf{x}_i)$  in Eq. (34)
4:   calculate the ratio  $r(\mathbf{x}_i, \mathbf{x}_p) = \min\left(1, \frac{\pi(\mathbf{x}_p)}{\pi(\mathbf{x}_i)}\right)$ 
       $\triangleright \pi(\cdot)$  denotes the target distribution (i.e. posterior).
5:   draw  $u \in [0, 1]$  from uniform probability density
6:   if  $r(\mathbf{x}_i, \mathbf{x}_p) \geq u$  then
7:      $\mathbf{x}_{i+1} = \mathbf{x}_p$ 
8:   else
9:      $\mathbf{x}_{i+1} = \mathbf{x}_i$ 
10:  end if
11:  per 1000 samples
12:    update matrix  $\tilde{\mathbf{K}}$ 
13: end for

```

5 Least squares method and genetic minimisation**5.1 The least squares method**

The LSM is based on measuring the squared difference between the measurement data and the response of the model for the same experiment. Consequently, the squared difference is a function of the parameters to be identified. In order to obtain the parameter values that give the best model response for the measurement data, the squared difference is minimized with respect to the parameters. Those parameter values that minimise the squared difference are the identified values. The difference that remains at the identified values (i.e. the residual) is a measure for how well the model fits the experimental data.

The squared difference $J(\mathbf{x}|\mathbf{y})$ for the one-dimensional SLS model and relaxation experiment in tension is:

$$J(\mathbf{x}|\mathbf{y}) = \frac{1}{2} \sum_{i=1}^k \left(y_i - E_0 \varepsilon_0 - E_1 \varepsilon_0 \exp\left(-\frac{E_1 t_i}{\eta}\right) \right)^2, \quad (35)$$

In contrast to the Bayesian formulation, the noise distribution does not require to be known explicitly in the standard LSM of Eq. (35). This saves calibration efforts. It does however entail two assumptions about the noise: (1) it assumes an additive noise model and (2) it assumes that the noise is symmetrically distributed.

The parameter values that give the best match between the measurement data and the model response are denoted by \mathbf{x}_{LSM} . The squared difference must be minimized to obtain them. We are thus interested in the following:

$$\mathbf{x}_{\text{LSM}} = \operatorname{argmin}_{\mathbf{x}} J(\mathbf{x}|\mathbf{y}), \quad (36)$$

where $\operatorname{argmin}_{\mathbf{x}}$ gives us the values of \mathbf{x} that minimise $J(\mathbf{x}|\mathbf{y})$.

5.2 Genetic optimisation

Numerous approaches can be employed to tackle the minimization problem of Eq. (36). In this contribution, we employ genetic optimization (GO). GO has the advantage that no initial guess is required. Instead, it searches for the optimum in a domain. This may be considered convenient for identification problems, because an appropriate initial guess for the parameters is not always trivial to choose. Another advantage is that GO finds the optimum in an entire, even it contains several local optima (Goldberg 1989). A particular advantage compared to the Newton-Raphson method is that GO is entirely derivative-free. Hence, no issues with the conditioning of Hessians occur (i.e. a squared matrix of second-order partial derivatives of a scalar valued function, Goldberg 1989).

The derivative-free character of GO entails some clear advantages as discussed above, but it also makes GO substantially slower than optimization approaches using derivatives (Chaparro et al. 2008). This can make the approach computationally prohibitive for large optimization problems. For the case in this contribution however, this is not an issue, because the model is an equation (Eqs. (2) and (36)) and only three parameters are to be identified.

The GO algorithm for the problem defined in Eq. (36) starts with the seeding of numerous vectors in the parameter space. In fact, we start with 1000 vectors, given by $E_0 \in \{0.5, 1.5556, \dots, 10\}$, $E_1 \in \{0.5, 1, \dots, 5\}$ and $\eta \in \{50, 63.3333, \dots, 170\}$. The squared difference (J) is computed for each of these vectors, and they are ranked based on their squared difference. The first 100 with the smallest squared difference are subsequently used to create new vectors, whilst the other ones are ignored. Each of the first 100 vectors will randomly choose 10 other vectors out of these 100 vectors (which may include the same vector more than once). With each partner vector j , vector i creates a new vector v , of which the values are established as follows:

$$E_0^v = (r_1^{E_0} E_0^i + (1 - r_1^{E_0}) E_0^j) r_2^{E_0}, \quad (37)$$

$$E_1^v = (r_1^{E_1} E_1^i + (1 - r_1^{E_1}) E_1^j) r_2^{E_1}, \quad (38)$$

and

$$\eta^v = (r_1^\eta \eta^i + (1 - r_1^\eta) \eta^j) r_2^\eta, \quad (39)$$

where $r_1^{E_0}$, $r_1^{E_1}$ and r_1^η are independently chosen from a uniform distribution in domain $[0, 1]$ and $r_2^{E_0}$, $r_2^{E_1}$ and r_2^η are independently chosen from a uniform distribution in domain $[0.95, 1.05]$.

For the newly created vectors, the squared difference is computed and the newly created vectors are ranked, together with the old vectors, based on their squared difference. Those 100 with the smallest squared difference (this can include old and new vectors) are then again used to create new vectors and the process is repeated for 50 generations.

6 Results

This section presents the results of the Bayesian framework for the SLS model. The main points we focus on are how the number of measurement points affects the results and the influence of the prior, and how the three different tests influence the results. However, we start with the calibration of the measurement noise for the relaxation experiment.

6.1 Noise calibration

In this subsection, we briefly discuss a procedure to identify the noise model, the noise distribution and its parameters. The noise is artificially created and hence, this subsection only aims to outline a procedure of how to identify the noise. The noise here is generated using an additive noise model and a normal distribution with a zero mean and $S_{\text{noise}} = 0.5$ MPa. The aim here is thus to find this back. We only focus on the identification of the noise in the relaxation test, because the procedure is principally the same for the other two tests (note that the output y for the creep test is the measured strain).

We start with a test without specimens. The stress-time measurements of these tests show that the noise in the stress behaves according to a normal distribution with a zero mean and a standard deviation S_{noise} (see Fig. 6(a)). No noise in the time is observed.

Now we would like to know if this distribution depends on the measured stress. Therefore, relaxation tests on a calibration specimen (of which parameters E_0 , E_1 and η are known) are performed. Some schematic results are shown in Fig. 6(b). The results show that the noise distribution does not evolve and only shifts with the exact response of the calibration specimen (shown by the bold black curve in Fig. 6(b)).

The artificially created calibration experiments show that the additive model can be used to describe the uncertainty and a normal distribution with a zero mean and a standard deviation of S_{noise} pollutes the stress measurements.

6.2 Identification of the SLS parameters using the relaxation test and two measurements

Bayesian inference

In this section we consider a relaxation experiment in tension on a specimen with $E_0 = 3.9455$ MPa, $E_1 = 2.9636$ MPa and $\eta = 136.8035$ MPa.s. The noise distribution is furthermore identified in subsection 6.1. The standard deviation of this normal distribution is identified as $S_{\text{noise}} = 0.5$ MPa.

In the first example, we consider only two measurements: $y_1 = 4.1598$ MPa and $y_2 = 1.9220$ MPa, measured at $t_1 = 10$ s and $t_2 = 80$ s, respectively. The measurements are made in the fast decaying part of the stress-time response (i.e. the

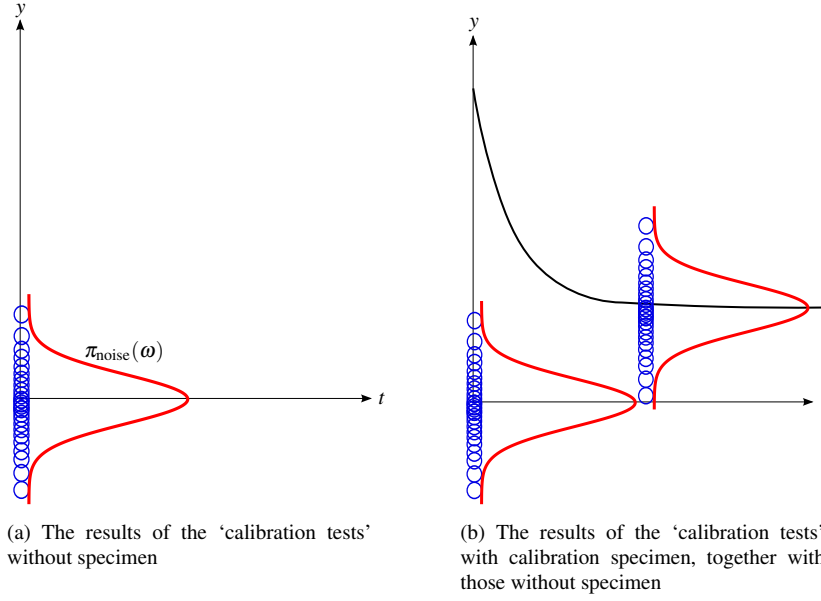


Fig. 6: Schematic of the stress-time measurements (blue circles) of the ‘calibration tests’, including the noise distributions (red curves). The stress-time relation of the calibration specimen is presented by the bold black curve in (b). The noise distribution shifts to the curve of the calibration sample, meaning that the additive noise model can be used to describe the uncertainty.

nonasymptotic part of the curve in Fig. 2). The prior is chosen to follow Eq. (15) with the following mean and covariance matrix:

$$\bar{\mathbf{x}}_{\text{prior}} = \begin{bmatrix} 5.5 \\ 2 \\ 110 \text{ s} \end{bmatrix} \text{MPa}, \quad \mathbf{\Gamma}_{\text{prior}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.4444 & 0 \\ 0 & 0 & 400 \text{ s}^2 \end{bmatrix} \text{MPa}^2. \quad (40)$$

This yields a posterior in the form of Eq. (17). We run an MCMC chain of 10^4 samples and we ‘burn’ the first 3000 samples (meaning that the first 3000 samples are not used to determine the mean, the covariance matrix and the MAP estimate). The result of this reads:

$$\bar{\mathbf{x}}_{\text{post}} = \begin{bmatrix} 4.7082 \\ 2.4488 \\ 106.9068 \text{ s} \end{bmatrix} \text{MPa}, \quad \mathbf{\Gamma}_{\text{post}} = \begin{bmatrix} 0.3253 & -0.0573 & -1.4514 \text{ s} \\ -0.0573 & 0.3487 & 0.3816 \text{ s} \\ -1.4514 \text{ s} & 0.3816 \text{ s} & 373.7074 \text{ s}^2 \end{bmatrix} \text{MPa}^2, \quad (41)$$

and

$$\mathbf{MAP} = \begin{bmatrix} 4.6841 \\ 2.3969 \\ 106.6689 \text{ s} \end{bmatrix} \text{MPa}. \quad (42)$$

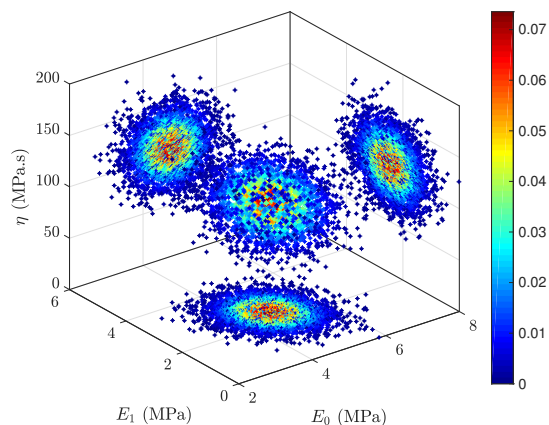


Fig. 7: Samples drawn by the adaptive MCMC approach to approximate the posterior distribution. The projections of the samples on three planes are also presented.

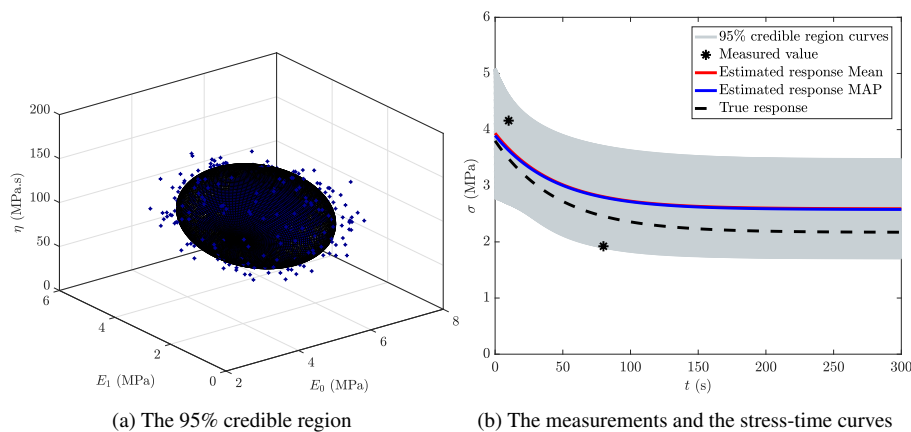


Fig. 8: The 95% credible region of the posterior distribution (a) and the stress-time response associated with the 95% credible region, together with the two measurements, the true response, the response associated with the mean and the response associated with the MAP point (b). The 95% credible region (approximated by an ellipsoid) contains 95% of the posterior. Note that the responses associated with the mean and the MAP point are practically on top of each other.

Fig. 7 shows the samples drawn by the adaptive MCMC approach and their projections on the planes $E_0 - E_1$, $E_1 - \eta$ and $\eta - E_0$. Fig. 8 shows the 95% credible region (i.e. the region that contains 95% of the posterior, approximated by an ellipsoid) and the associated stress-time responses. The wide credible region is partially caused by the large standard deviation of the noise distribution ($S_{\text{noise}} = 0.5$ MPa).

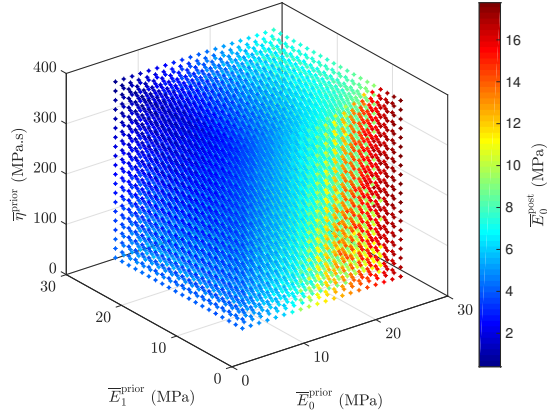


Fig. 9: The mean value of E_0 of the posterior (\bar{E}_0^{post}) as a function of the mean values of the prior (\bar{E}_0^{prior} , \bar{E}_1^{prior} and $\bar{\eta}^{\text{prior}}$). \bar{E}_0^{post} is mostly dependent on the selected value of \bar{E}_0^{prior} and \bar{E}_1^{prior} , whereas $\bar{\eta}^{\text{prior}}$ has substantially less influence.

The prior in BI may have a significant effect on the posterior for a small number of measurements (as is the case here). To illustrate this influence, we now change the mean of the prior, while keeping the covariance matrix of the prior constant. The following covariance matrix is used for the prior:

$$\mathbf{\Gamma}_{\text{prior}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.4444 & 0 \\ 0 & 0 & 277.7778 \text{ s}^2 \end{bmatrix} \text{MPa}^2. \quad (43)$$

Figs. 9 to 11 show that the mean value of a particular parameter in the posterior mostly depends on the mean value of that same parameter in the prior (e.g. \bar{E}_1^{post} depends mostly on the selected value of \bar{E}_1^{prior}). For \bar{E}_0^{post} however, \bar{E}_0^{prior} and \bar{E}_1^{prior} both have an influence.

$\bar{\eta}^{\text{post}}$ almost entirely depends on $\bar{\eta}^{\text{prior}}$. This is caused by the fact that η has relatively the smallest influence on the stress-time relation. After all, E_0 is responsible for the plateau stress and E_1 (together with E_0) is responsible for the initial stress at the start of the relaxation (see Fig. 2). Parameters η and E_1 are however together responsible for the time scale at which the initial stress relaxes to the plateau stress (see Eq. (2)). Consequently, the sensitivity of the stress-time relation is the smallest for η of the three parameters.

Fig. 12 shows the effect of the mean of the prior on the diagonal components of the posterior ($\mathbf{\Gamma}_{\text{post}}$). The results show that the mean of the prior only affects $(\mathbf{\Gamma}_{\text{post}})_{33}$ in a systematic manner. An increasing $\bar{\eta}^{\text{prior}}$ leads to an increase of $(\mathbf{\Gamma}_{\text{post}})_{33}$, whereas an opposite effect can be observed for \bar{E}_0^{prior} and \bar{E}_1^{prior} .

The effect of the selected mean of the prior on the off-diagonal components of the $\mathbf{\Gamma}_{\text{post}}$ is shown in Fig. 13. The prior's mean hardly has a systematic influence on the off-diagonal components.

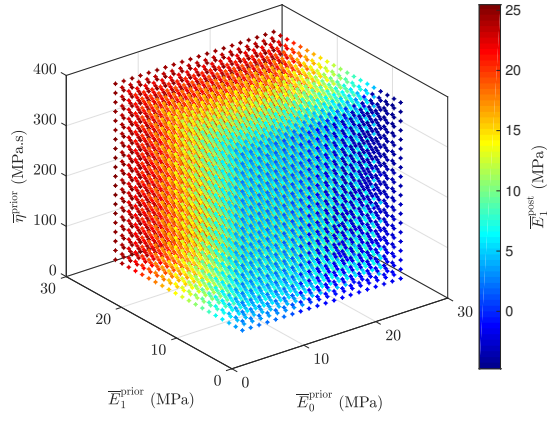


Fig. 10: The mean value of E_1 of the posterior (\bar{E}_1^{post}) as a function of the mean values of the prior (\bar{E}_0^{prior} , \bar{E}_1^{prior} and $\bar{\eta}^{\text{prior}}$). \bar{E}_1^{post} is mostly dependent on the selected value of \bar{E}_1^{prior} , compared to \bar{E}_0^{prior} and $\bar{\eta}^{\text{prior}}$.

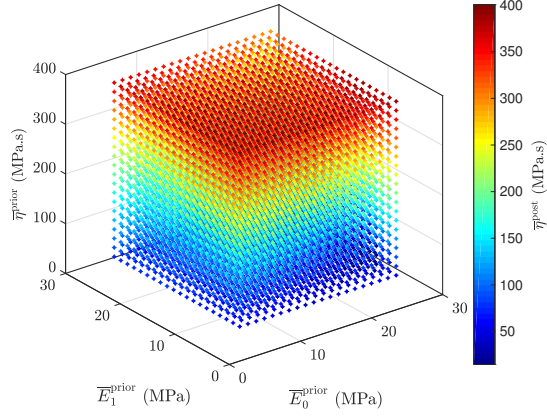


Fig. 11: The mean value of η of the posterior ($\bar{\eta}^{\text{post}}$) as a function of the mean values of the prior (\bar{E}_0^{prior} , \bar{E}_1^{prior} and $\bar{\eta}^{\text{prior}}$). $\bar{\eta}^{\text{post}}$ depends highly on the selected value of $\bar{\eta}^{\text{prior}}$, compared to the values of \bar{E}_0^{prior} and \bar{E}_1^{prior} .

Now we will look at the effect of the prior's covariance matrix on the posterior. We will keep the mean of the prior constant in this exercise. To investigate the effect of the prior's covariance matrix, the modified normal distribution given in Eq. (15) is selected with following mean:

$$\bar{\mathbf{x}}_{\text{prior}} = \begin{bmatrix} 6 \\ 3 \\ 110 \text{ s} \end{bmatrix} \text{ MPa}, \quad (44)$$

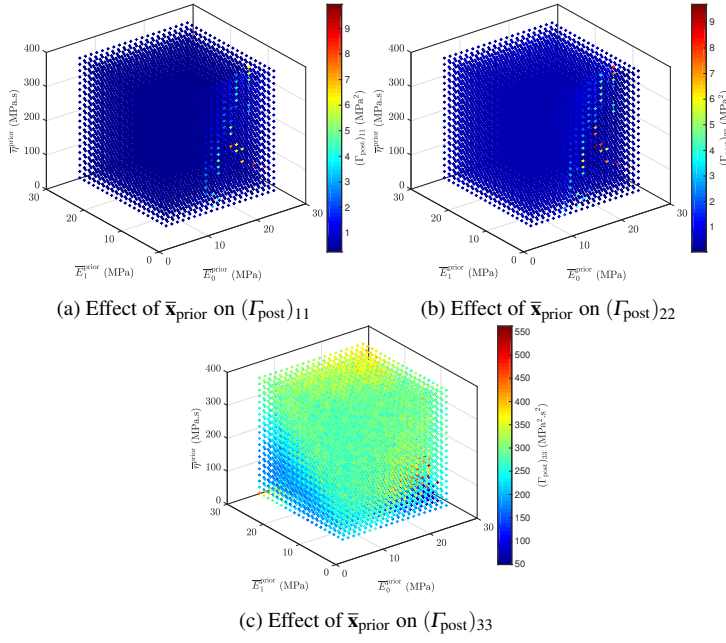


Fig. 12: The effect of the mean of the prior on the diagonal components of the posterior's covariance matrix. Only $(\Gamma_{\text{post}})_{33}$ is systematically influenced by the mean of the prior.

We have chosen this prior's mean, because it is located relatively far from $\bar{\mathbf{x}}_{\text{prior}} = [0 \ 0 \ 0]^T$ so that we can investigate the influence of relative large components of the prior's covariance matrices.

Figs. 14 to 16 show the effect of the different diagonal components of the prior covariance matrix ($(\Gamma_{\text{prior}})_{11}$, $(\Gamma_{\text{prior}})_{22}$ and $(\Gamma_{\text{prior}})_{33}$) on the mean of the posterior. The results show that \bar{E}_0^{post} mostly depends on $(\Gamma_{\text{prior}})_{11}$ and for an increase of $(\Gamma_{\text{prior}})_{11}$, the influence on \bar{E}_0^{post} decreases. \bar{E}_1^{post} is hardly influenced by $(\Gamma_{\text{prior}})_{33}$. All components of the prior's covariance matrix are of influence on $\bar{\eta}^{\text{post}}$.

The effect of the prior's covariance matrix on the components of the posterior's covariance matrix Γ_{post} is shown in Figs. 17 and 18. The diagonal components of the posterior's covariance matrix clearly depend mostly on their equivalents in the prior's covariance matrix. For the off-diagonal components Fig. 18 shows that $(\Gamma_{\text{post}})_{12}$ is mostly influenced by $(\Gamma_{\text{prior}})_{11}$ and $(\Gamma_{\text{prior}})_{22}$. $(\Gamma_{\text{post}})_{13}$ depends on all diagonal components, and $(\Gamma_{\text{post}})_{23}$ is influenced by $(\Gamma_{\text{prior}})_{22}$ and $(\Gamma_{\text{prior}})_{33}$.

Least squares method

As the number of parameters to be identified is three and we only have two measurements, the minimisation of the objective function in the LSM ($J(\mathbf{x}|\mathbf{y})$ in Eqs. (35) and (36)) is ill-posed. This means that no unique solution to the minimisation prob-

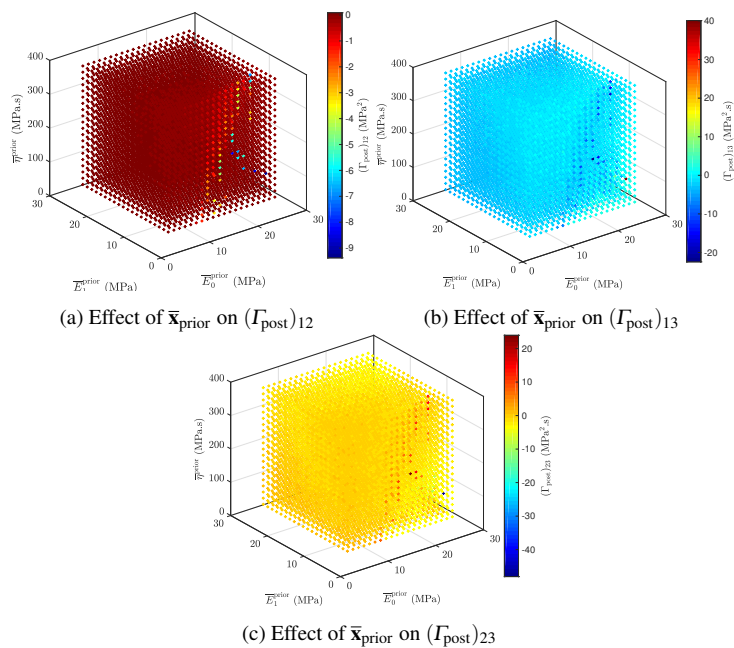


Fig. 13: The effect of the mean of the prior on the off-diagonal components of the posterior’s covariance matrix. The prior’s mean does not seem to have a systematic influence.

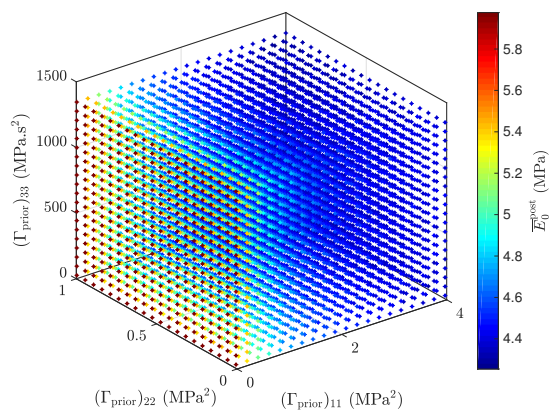


Fig. 14: The effect of Γ_{prior} on \bar{E}_0^{post} . Except for $(\Gamma_{\text{prior}})_{11}$, the diagonal components of the prior’s covariance matrix hardly have an influence on \bar{E}_0^{post} .

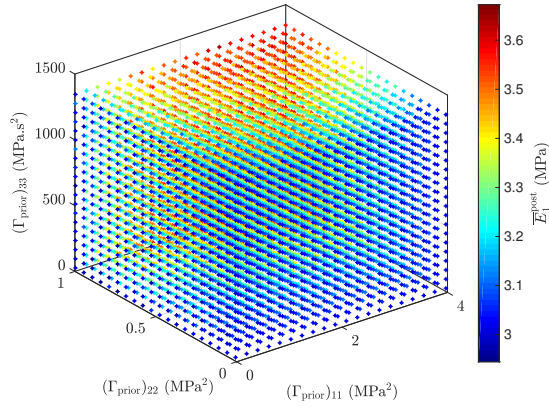


Fig. 15: The effect of $\mathbf{\Gamma}_{\text{prior}}$ on \bar{E}_1^{post} . \bar{E}_1^{post} is influenced by $(\Gamma_{\text{prior}})_{11}$ and $(\Gamma_{\text{prior}})_{22}$, but not by $(\Gamma_{\text{prior}})_{33}$.

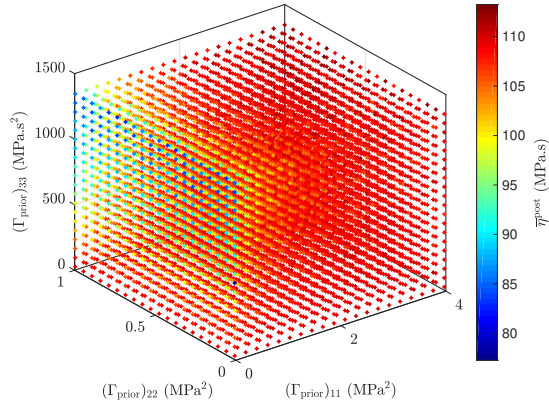


Fig. 16: The effect of $\mathbf{\Gamma}_{\text{prior}}$ on the estimated value for $\bar{\eta}^{\text{post}}$. All diagonal components of the prior's covariance matrix are of influence.

lem of Eq. (32) exists. Consequently, a wide range of parameter sets can be obtained. Fig. 19 shows some possible curves generated based on the results of running the GO routine described in subsection 5.2 several times. Note that the various curves obtained for the LSM do not represent any uncertainty as the curves created using the 95% credible region in the Bayesian approach. The bounds between which any possible curve of the LSM is located are presented by the two blue curves in Fig. 19.

Hence, for two measurements a large difference between the Bayesian approach and the LSM occurs. After all, the LSM case is ill-posed, whereas the prior in BI regularises the Bayesian framework.

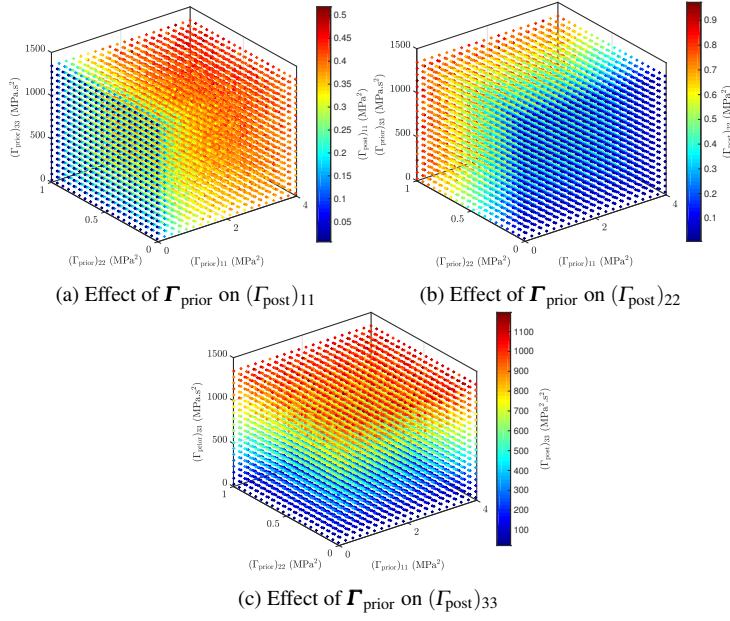


Fig. 17: The effect of $\mathbf{\Gamma}_{\text{prior}}$ on the diagonal components of the posterior's covariance matrix. Each diagonal component is mostly dependent on its equivalent in the prior covariance matrix.

6.3 Identification of the SLS parameters using the relaxation test and more than two measurements

In the next example, we add another three measurements to the two measurements of the previous subsection and investigate how this influences the results of the Bayesian framework as well as those of the LSM. The modified normal distribution given in Eq. (15) with the mean and covariance matrix given in Eq. (40) is selected as the prior distribution. Running the MCMC chain for 10^4 samples, whilst burning the first 3000 samples, yields:

$$\bar{\mathbf{x}}_{\text{post}} = \begin{bmatrix} 4.8354 \\ 2.4069 \\ 109.7946 \text{ s} \end{bmatrix} \text{ MPa}, \quad \mathbf{\Gamma}_{\text{post}} = \begin{bmatrix} 0.1453 & -0.0328 & -1.4935 \text{ s} \\ -0.0328 & 0.3392 & 0.6262 \text{ s} \\ -1.4935 \text{ s} & 0.6262 \text{ s} & 361.5418 \text{ s}^2 \end{bmatrix} \text{ MPa}^2, \quad (45)$$

and

$$\mathbf{MAP} = \begin{bmatrix} 4.8463 \\ 2.4558 \\ 110.6589 \text{ s} \end{bmatrix} \text{ MPa}. \quad (46)$$

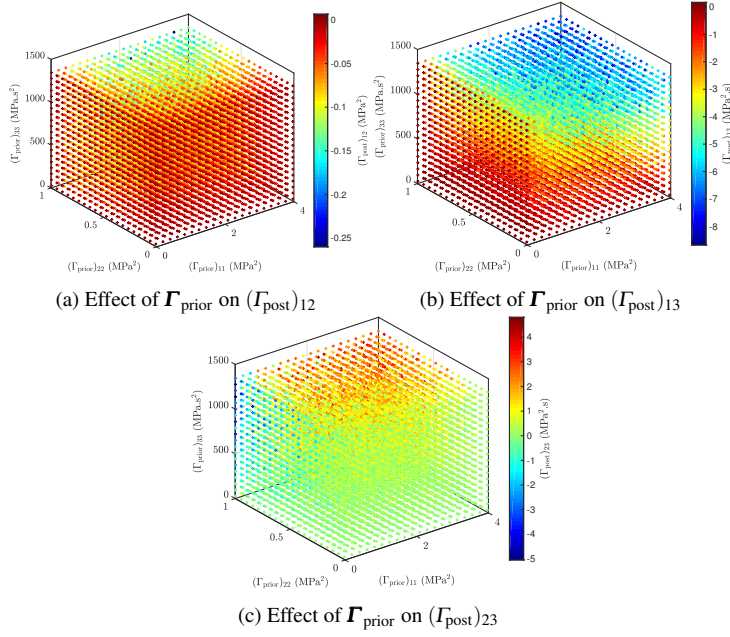


Fig. 18: The effect of Γ_{prior} on the off-diagonal components of the posterior's covariance matrix. $(\Gamma_{\text{post}})_{12}$ is mostly influenced by $(\Gamma_{\text{prior}})_{11}$ and $(\Gamma_{\text{prior}})_{22}$. $(\Gamma_{\text{post}})_{13}$ depends on all diagonal components and $(\Gamma_{\text{post}})_{23}$ is influenced by $(\Gamma_{\text{prior}})_{22}$ and $(\Gamma_{\text{prior}})_{33}$.

The values identified by the LSM are:

$$\mathbf{x}_{\text{LSM}} = \begin{bmatrix} 4.1156 \\ 4.6433 \\ 199.5536 \text{ s} \end{bmatrix} \text{ MPa.} \quad (47)$$

Comparing Eqs. (45), (47) and the true values, one can see that the errors of E_1^{LSM} and η^{LSM} are larger than of \bar{E}_1^{post} and $\bar{\eta}^{\text{post}}$. This is thanks to the prior knowledge used in the Bayesian framework. One should however note that if the number of measurements is small, the selection of the prior's mean far from the true values and small components for the prior's covariance matrix would lead to a larger error. After all, the Bayesian framework aims to account for the fact that only a limited number of measurement points are available by incorporating a prior distribution. It is therefore questionable if a direct comparison between the LSM results and the posterior's mean and MAP point is truly valid at all.

The stress-time curves associated with the 95% credible region are given in Fig. 20. Increasing the number of measurements clearly leads to a narrower credible region (i.e. a smaller uncertainty, cf. Fig. 8(b)).

To study the effect of the prior's mean on the posterior's mean, we again show the posterior's mean values as functions of the prior's mean, see Fig. 21. Comparing

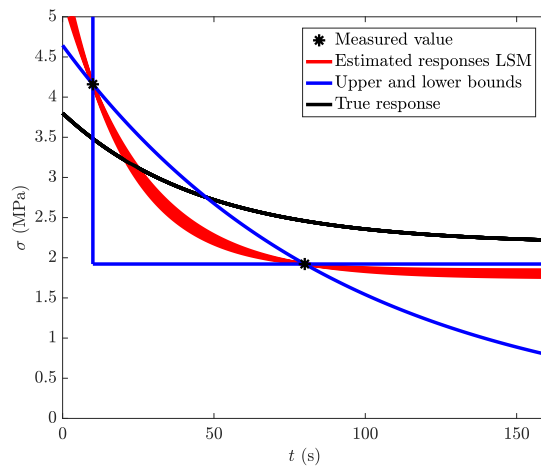


Fig. 19: Different curves (red) associated with the parameters identified by the LSM due to ill-posedness. All the curves are made for the same measurements and each time that the LSM is applied as an identification approach the identified values lead to a different curve. However, in the Bayesian approach the problem is regularised by the prior and hence, it is not ill-posed. Note that the estimated responses by the LSM do not represent any uncertainty.

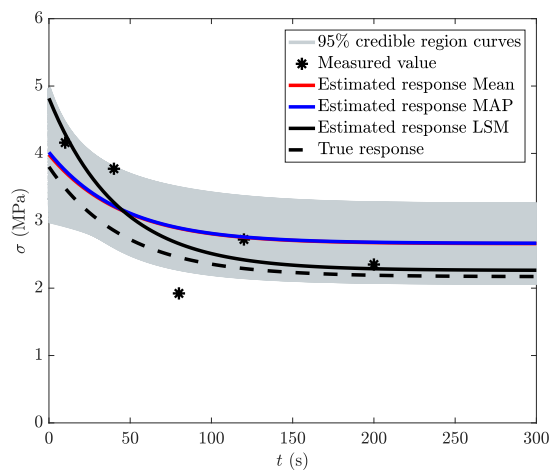


Fig. 20: The stress-time curves for the relaxation test with five measurements. Increasing the number of the measurements leads to a narrower region (cf. Fig. 8(b)). This means that the uncertainty decreases as the number of measurements increases. Note that the responses associated with the mean and the MAP point are almost on top of each other.

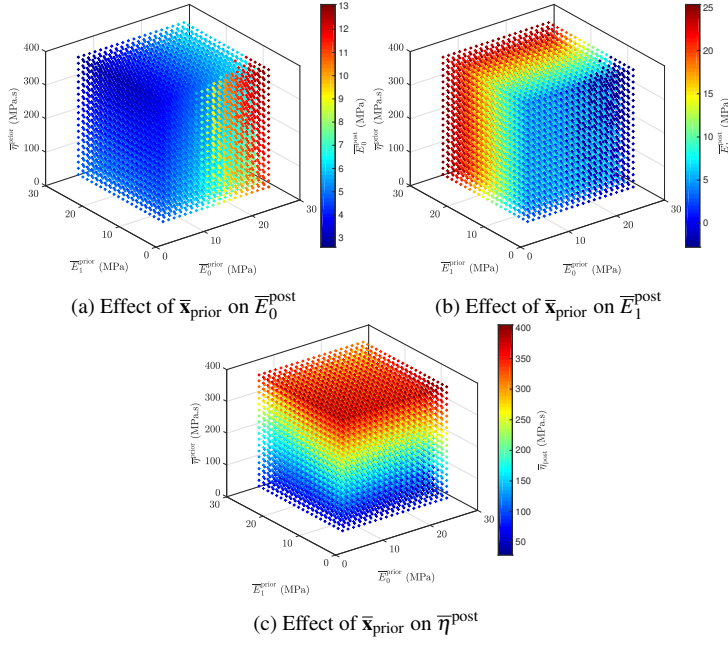


Fig. 21: The effect of $\bar{\mathbf{x}}_{\text{prior}}$ on $\bar{\mathbf{x}}_{\text{post}}$ for five measurements. An increase of the number of measurements results in a smaller dependency of \bar{E}_0^{post} on $\bar{\mathbf{x}}_{\text{prior}}$. This is also the case for \bar{E}_1^{post} and $\bar{\eta}^{\text{post}}$, but this is more difficult to observe by comparing this figure with Fig. 9.

Fig. 21 with 9 shows that an increasing number of measurements results in a smaller dependency of \bar{E}_0^{post} on $\bar{\mathbf{x}}_{\text{prior}}$. This is also true for \bar{E}_1^{post} and η^{post} , but this cannot well be observed by comparing Figs. 21 and 9.

In the next example the number of measurements is increased to ten. Assuming the same prior distribution given in Eq. (15) with $\bar{\mathbf{x}}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{prior}}$ given in Eq. (40) the mean, covariance matrix and MAP point read:

$$\bar{\mathbf{x}}_{\text{post}} = \begin{bmatrix} 4.6381 \\ 2.4154 \\ 107.4595 \text{ s} \end{bmatrix} \text{ MPa}, \quad \mathbf{\Gamma}_{\text{post}} = \begin{bmatrix} 0.0812 & -0.0143 & -1.4925 \text{ s} \\ -0.0143 & 0.3043 & 0.8244 \text{ s} \\ -1.4925 \text{ s} & 0.8244 \text{ s} & 359.5813 \text{ s}^2 \end{bmatrix} \text{ MPa}^2, \quad (48)$$

and

$$\mathbf{MAP} = \begin{bmatrix} 4.6507 \\ 2.4299 \\ 106.5064 \text{ s} \end{bmatrix} \text{ MPa}. \quad (49)$$

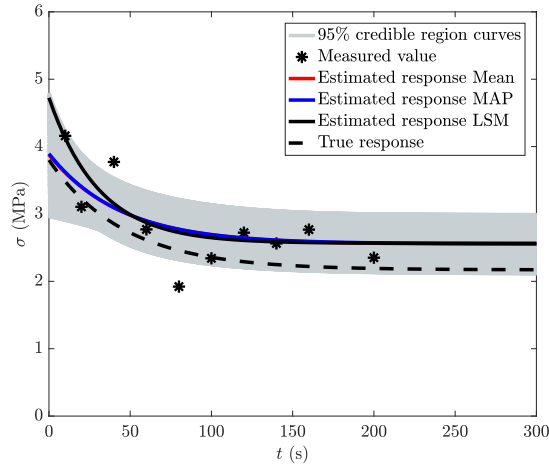


Fig. 22: The stress-time curves for the relaxation experiment with ten measurements. Increasing the number of measurements leads to a narrower region (cf. Figs. 8(b) and 21). This means that the uncertainty decreases as the number of measurements increases. Note that the responses associated with the mean and the MAP point are practically on top of each other.

The estimated values using the LSM are:

$$\mathbf{x}_{\text{LSM}} = \begin{bmatrix} 4.4661 \\ 3.9303 \\ 121.0358 \text{ s} \end{bmatrix} \text{ MPa.} \quad (50)$$

Fig. 22 shows the associated stress-time curves. As the number of measurements is increased compared to the previous examples, the responses associated with the 95% credible region become more localised (i.e. the uncertainty decreases, cf. Figs. 8(b) and 21).

The effect of prior's mean on $\bar{\mathbf{x}}_{\text{post}}$ is shown in Fig. 23. Comparing Fig. 23 with Figs. 9, 10, 11 and 21, it is clear that the prior's mean has again less influence on the posterior's mean. This is most obvious for \bar{E}_0^{post} .

6.4 Identification of the SLS parameters when the measurements are created by the generalised Maxwell model using the relaxation test

In the final example for the relaxation test, we study the results of the Bayesian approach and the LSM when the measurements are generated using the generalised Maxwell model with three spring-dashpots in series (instead of only one as in the SLS model). The model is shown schematically in Fig. 24. We take $E_0 = 4$ MPa, $E_1 = 3$ MPa, $E_2 = 2.5$ MPa, $E_3 = 2$ MPa, $\eta = 140$ MPa.s, $\eta_2 = 110$ MPa.s and

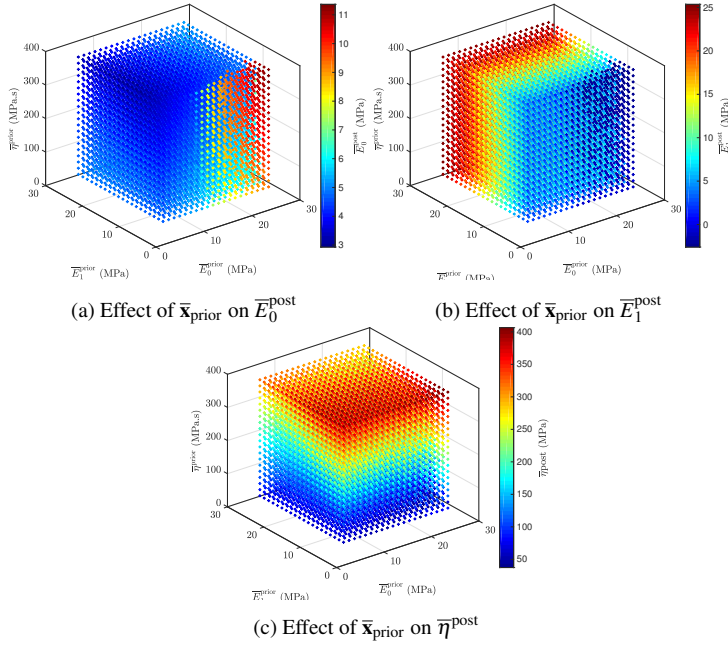


Fig. 23: The effect of the prior's mean (\bar{x}_{prior}) on \bar{E}_0^{post} , \bar{E}_1^{post} and $\bar{\eta}^{\text{post}}$ for ten measurements. The effect of the prior's mean becomes less significant, especially for \bar{E}_0^{post} . For \bar{E}_1^{post} and $\bar{\eta}^{\text{post}}$ one can observe that an increase in the number of measurements decreases the prior's effect.

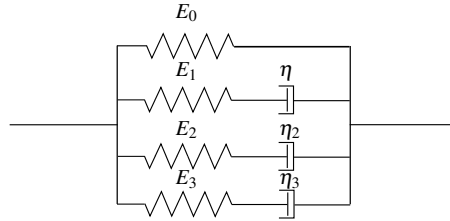


Fig. 24: Schematic representation of the generalised Maxwell model with three spring-dashpots in series.

$\eta_3 = 100$ MPa.s. The same noise distribution is considered as in the previous examples ($S_{\text{noise}} = 0.5$ MPa). The modified normal distribution given in Eq. (15) is selected as the prior distribution with \bar{x}_{prior} and Γ_{prior} given in Eq. (40). Running the MCMC chain for 10^4 sample and burning the first 3000 samples yields:

$$\bar{\mathbf{x}}_{\text{post}} = \begin{bmatrix} 4.5290 \\ 3.4806 \\ 122.3883 \text{ s} \end{bmatrix} \text{ MPa}, \quad \mathbf{\Gamma}_{\text{post}} = \begin{bmatrix} 0.0810 & -0.0175 & -1.7183 \text{ s} \\ -0.0175 & 0.2714 & 1.9272 \text{ s} \\ -1.7183 \text{ s} & 1.9272 \text{ s} & 300.6031 \text{ s}^2 \end{bmatrix} \text{ MPa}^2, \quad (51)$$

and

$$\mathbf{MAP} = \begin{bmatrix} 4.5352 \\ 3.4771 \\ 122.9158 \text{ s} \end{bmatrix} \text{ MPa}. \quad (52)$$

The identified values by the LSM are:

$$\mathbf{x}_{\text{LSM}} = \begin{bmatrix} 3.3702 \\ 9.9743 \\ 306.8281 \text{ s} \end{bmatrix} \text{ MPa}. \quad (53)$$

In Fig. 25 all associated stress-time responses are presented. The LSM response is clearly closer to the true response, than the responses associated with the mean and the MAP of the Bayesian framework. The prior distribution clearly also has a substantial influence for this case.

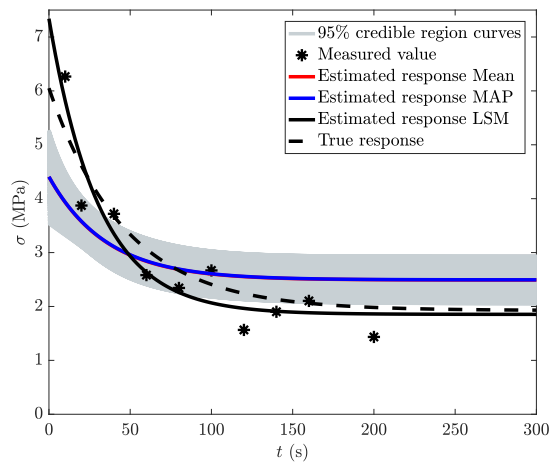


Fig. 25: The stress-time response for the SLS model when ten measurement points are created using the generalised Maxwell model. The response identified by the LSM is significantly closer to the true response than the responses associated with the mean and the MAP point of the Bayesian framework. Hence, also in this case the prior has a significant influence in the Bayesian framework. Note that the responses associated with the mean and the MAP point are on top of each other.

6.5 Comparison with the constant strain-rate test and the creep test

In this subsection we compare the results for the relaxation test to the results for the constant strain-rate test and the creep test. To make our comparison as ‘fair’ as possible, we try to keep every condition as similar to the relaxation test with five measurements. This means amongst others that we also consider five measurements in the constant strain-rate test and the creep test and that we again try to find the same parameter values back ($E_0 = 3.9455$ MPa, $E_1 = 2.9636$ MPa and $\eta = 136.8035$ MPa.s). In the following two subsections we explain how we try to keep the other conditions as similar as possible. The comparison of the results is presented in subsection 6.5.3.

6.5.1 Set-up of the constant strain-rate test

In the constant strain-rate test, strain-rate $\dot{\epsilon}_0$ is chosen such that the maximum stress that will occur is the same as the maximum stress occurring in the relaxation test. Consequently, the following expression to determine the applied strain-rate:

$$\dot{\epsilon}_0 = \frac{(E_0 + E_1)\epsilon_0}{E_0 t_{\text{exp}} + \eta \left(1 - \exp\left(-\frac{E_1 t_{\text{exp}}}{\eta}\right)\right)}, \quad (54)$$

where $t_{\text{exp}} = 300$ s and represents the duration of the relaxation test as well as that of the constant strain-rate test. The strain applied in the relaxation test is represented by ϵ_0 .

Note that the reason to choose $\dot{\epsilon}_0$ according to Eq. (54) is to have the same order of magnitude for the stresses in both tests. In this way, the noise distribution (for which we again use a normal distribution with zero mean and $S_{\text{noise}} = 0.5$ MPa) has approximately the same influence in both tests. Furthermore, we assume that the noise realisation for each measurement at time t_i in the constant strain-rate test is the same as the noise realisation at the same time in the relaxation test. In this way, the noise realisations have approximately the same influence in both tests.

6.5.2 Set-up of the creep test

The noise distribution in the creep test is again chosen to be in the form of Eq. (13). As the quantity measured in a creep test is the strain however, the standard deviation of the noise distribution (i.e. S_{noise} in Eq. (13)) must be re-defined such that a ‘fair’ comparison can be made between the results of the previous two tests and the creep test. For this reason, the standard deviation of the noise distribution in the creep test, $S_{\text{noise}}^{\text{creep}}$ is chosen as follows:

$$S_{\text{noise}}^{\text{creep}} = \frac{S_{\text{noise}}^{\text{relaxation}} \epsilon_{\text{max}}^{\text{creep}}}{\sigma_{\text{max}}^{\text{relaxation}}}, \quad (55)$$

where $S_{\text{noise}}^{\text{relaxation}}$ is the standard deviation of the noise in the relaxation test, $\epsilon_{\text{max}}^{\text{creep}}$ is the maximum strain occurring in the creep test and $\sigma_{\text{max}}^{\text{relaxation}}$ is the maximum stress

that has occurred in the relaxation test. Using Eqs. (2) and (4) $\epsilon_{\max}^{\text{creep}}$ and $\sigma_{\max}^{\text{relaxation}}$ are calculated as follows:

$$\epsilon_{\max}^{\text{creep}} = \frac{\sigma_0}{E_0} + \sigma_0 \left(\frac{1}{E_0 + E_1} - \frac{1}{E_0} \right) \exp\left(-\frac{E_0 E_1 t_{\text{exp}}^{\text{creep}}}{\eta(E_0 + E_1)} \right), \quad (56)$$

and

$$\sigma_{\max}^{\text{relaxation}} = (E_0 + E_1) \epsilon_0, \quad (57)$$

where $t_{\text{exp}}^{\text{creep}}$ represents the duration of the creep test.

Furthermore, a comparison between Eqs. (2), (3) and (4) shows that the coefficient in the exponent of the strain-time expression for the creep test (i.e. $\frac{E_0 E_1}{\eta(E_0 + E_1)}$) is smaller than the one in the strain-time expressions for the other two tests (i.e. $\frac{E_1}{\eta}$). This entails that more time is required in the creep test to achieve the same amount of relaxation as in the previous two tests. Consequently, the experiment duration for the creep test is scaled as $t_{\text{exp}}^{\text{creep}} = \frac{(E_0 + E_1)}{E_0} t_{\text{exp}}^{\text{relaxation}}$. $t_{\text{exp}}^{\text{relaxation}}$ refers here to the duration of the relaxation test (here $t_{\text{exp}}^{\text{relaxation}} = 300$ s and $t_{\text{exp}}^{\text{creep}} = 525.3403$ s).

Because we scale the duration of the creep test, we also scale the time at which each measurement is made as $t_i^{\text{creep}} = \frac{(E_0 + E_1)}{E_0} t_i^{\text{relaxation}}$. t_i^{creep} refers here to the time at which measurement i is made in the creep test and $t_i^{\text{relaxation}}$ refers to the time at which the same measurement is made in the relaxation test.

Using Eq. (55) the standard deviation of the noise distribution in the creep test is determined as $S_{\text{noise}} = 0.0724$. Note that σ_0 in Eq. (4) is chosen such that the final strain in the creep test is the same as the strain applied in the relaxation test (i.e. ϵ_0 in Eq. (2)). Additionally, we assume that noise realisation i in the creep test (ω_i^{creep}) has the same probability to occur as noise realisation i in the relaxation test ($\omega_i^{\text{relaxation}}$). This entails that we scale each noise realisation as follows:

$$\omega_i^{\text{creep}} = \omega_i^{\text{relaxation}} \frac{S_{\text{noise}}^{\text{creep}}}{S_{\text{noise}}^{\text{relaxation}}}. \quad (58)$$

6.5.3 Comparison of results

In this subsection we present the results of both the constant strain-rate test and the creep test and compare them with each other and those of the relaxation test. The prior for the two new tests is the same as we used for the relaxation test with five measurements (using Eq. (15) with the mean and covariance matrix in Eq. (40)). The MCMC chain is again run for 10^4 sample and we have burnt the first 3000 samples. Table 1 shows the estimated values of each parameter and each component of the posterior's covariance matrix ($\mathbf{\Gamma}_{\text{post}}$) for the three tests.

Comparing the values given in Table 1 with each other and the real values (i.e. $E_0 = 3.9455$ MPa, $E_1 = 2.9636$ MPa and $\eta = 136.8035$ MPa.s), one can see that the identified parameter values for the three tests are all approximately of the same accuracy. The relaxation and constant strain-rate test overestimate the value of E_0 , whereas the creep test identifies this value accurately. However, the creep test underestimates the

Table 1: The estimated values of the posterior's mean, MAP and components of the covariance matrix for the three experiments.

Estimated value	Relaxation test	Constant strain-rate test	Creep test
\bar{E}_0^{post} (MPa)	4.8354	4.9115	3.9047
\bar{E}_1^{post} (MPa)	2.4069	2.0435	1.7861
$\bar{\eta}^{\text{post}}$ (MPa.s)	109.7946	109.0151	105.7801
E_0^{MAP} (MPa)	4.8463	4.8448	3.8569
E_1^{MAP} (MPa)	2.4558	1.9383	1.7678
η^{MAP} (MPa.s)	110.6589	107.4682	106.1140
$(\Gamma_{\text{post}})_{11}$ (MPa ²)	0.1453	0.3548	0.0748
$(\Gamma_{\text{post}})_{22}$ (MPa ²)	0.3392	0.4562	0.4173
$(\Gamma_{\text{post}})_{33}$ (MPa ² .s ²)	361.5418	398.6452	420.1033
$(\Gamma_{\text{post}})_{12}$ (MPa ²)	-0.0328	-0.0302	-0.0121
$(\Gamma_{\text{post}})_{13}$ (MPa ² .s)	-1.4935	-1.9472	-0.9538
$(\Gamma_{\text{post}})_{23}$ (MPa ² .s)	0.6262	0.2432	-0.8297

value of E_1 substantially more than the other two tests. This is not very surprising since this results from the fact that the stress is measured in the relaxation and constant strain-rate test and that the strain is measured in the creep test. The posteriors' means and MAPs for the different tests in Table 1 cannot well be compared to begin with, since we only consider five measurements and one prior.

More interesting are the results for the uncertainty of E_0 . After all, the values of $(\Gamma_{\text{post}})_{11}$ vary with a factor of five for the three tests. The difference between the uncertainty of E_0 of the relaxation test and the constant strain-rate test can be explained based on comparing the stress-time relations for both tests (see Eqs. (2) and (3)). If t approaches infinity in the relaxation test ($t \rightarrow +\infty$), an horizontal asymptote occurs ($\sigma = E_0 \varepsilon_0$) which depends only on E_0 . In the the constant strain-rate test however, an oblique asymptote occurs (i.e. $E_0 \dot{\varepsilon}_0 t + \eta \dot{\varepsilon}_0$) which depends on two parameters (E_0 and η). This means if $t \rightarrow \infty$, a measurement in the relaxation test only provides information about E_0 , but in the constant strain-rate test this measurement gives information about E_0 and η . Consequently, E_0 is determined with more certainty in the relaxation test than in the constant strain-rate test. Furthermore, at the beginning of the relaxation test information is directly obtained for E_0 and E_1 , whereas the constant strain-rate does not provide any information at time $t = 0$ s. The same reasoning can be used for the creep test as well. Unlike the stress-time relation for the relaxation test however, the one for the creep test is a combination of homographic and exponential functions. For this reason we obtain a different value for $(\Gamma_{\text{post}})_{11}$ in the creep test.

Figs. 26 and 27 present the stress-time responses and the strain-time responses of the constant strain-rate test and the creep test, respectively. The curves associated with 95% credible region are generated using the points from region which contain 95% of the posterior (approximated by an ellipsoid).

In the remainder of this subsection we investigate the effect of prior's mean $\bar{\mathbf{x}}_{\text{prior}}$ on the results. For this exercise, the prior's covariance matrix is kept constant (i.e. the covariance matrix given in Eq. (40)).

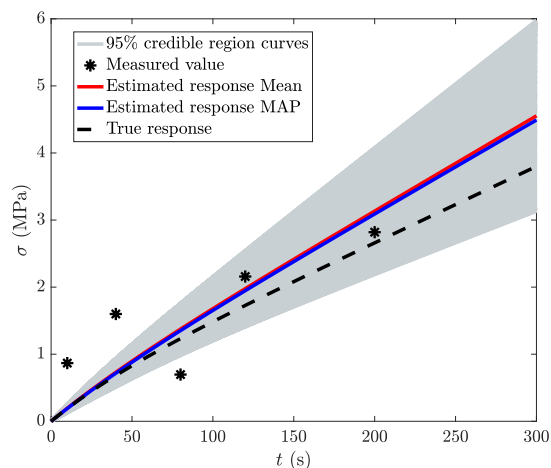


Fig. 26: The stress-time curves for the constant strain-rate test with five measurements. The measurements are made using the same noise distribution as in subsection 6.2. Furthermore, it is assumed that the noise realisation for each measurement at time t_i in the constant strain-rate test is the same as the noise realisation at the same time in the relaxation test. Note that the associated curves with the mean and MAP point are practically on top of each other.

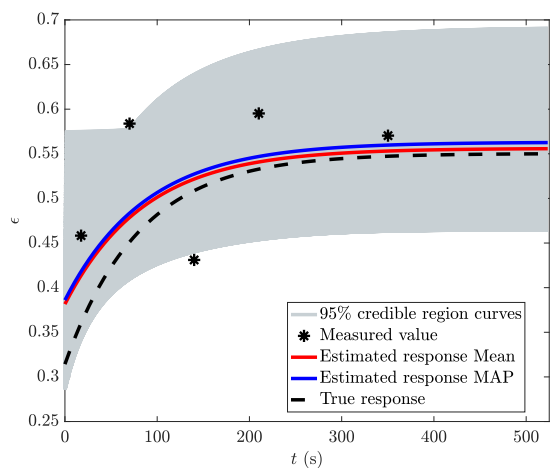


Fig. 27: The strain-time curves for the creep test with five measurements. The measurements are generated using a normal distribution with zero mean and a standard deviation given by Eq. (55). To ensure that the same amount of relaxation occurs as in the other two tests, the time at which each measurement is made is scaled according to $t_i^{\text{creep}} = \frac{(E_0 + E_1)}{E_0} t_i^{\text{relaxation}}$.

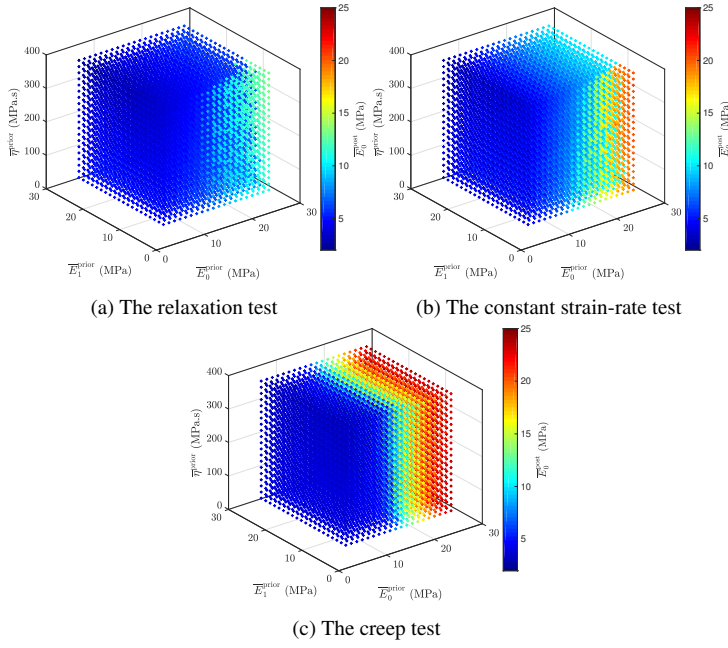


Fig. 28: The effect of \bar{x}_{prior} on \bar{E}_0^{post} for the three tests. \bar{E}_0^{post} depends least on the prior for the relaxation test. \bar{E}_0^{post} depends only on \bar{E}_0^{prior} for the creep test, whereas it depends on \bar{E}_0^{prior} and \bar{E}_1^{prior} in the other two tests.

Fig. 28 shows the effect of \bar{x}_{prior} on \bar{E}_0^{post} for the relaxation, constant strain-rate and creep test. It can be observed that \bar{E}_0^{post} depends least on \bar{x}_{prior} for the relaxation test. It can also be observed that \bar{E}_0^{post} depends only on \bar{E}_0^{prior} for the creep test, whereas it depends on \bar{E}_0^{prior} and \bar{E}_1^{prior} in the other two tests. Note that a possible reason for the different influence of the prior in the creep test is the type of function for its stress-time relation (Eq. (4) is a combination of homographic and exponential function of parameters).

Now we investigate the effect of the prior's mean on the estimated values. The prior's covariance matrix is again kept constant for this. The influence of the prior's mean on the estimated values is very similar to that for the relaxation test, except for $(\Gamma_{\text{post}})_{11}$, $(\Gamma_{\text{post}})_{33}$, $(\Gamma_{\text{post}})_{13}$ and $(\Gamma_{\text{post}})_{23}$. In Fig. 29 the effect of \bar{x}_{prior} on $(\Gamma_{\text{post}})_{11}$ is presented. Only for the creep test an actual trend can be observed.

Fig. 30 indicates that the effect of \bar{x}_{prior} on the estimated value of $(\Gamma_{\text{post}})_{33}$ for the relaxation test, whereas it is the smallest for the constant strain-rate test.

Figs. 31 and 32 show the effect of \bar{x}_{prior} on $(\Gamma_{\text{post}})_{13}$ and $(\Gamma_{\text{post}})_{23}$ for the three tests employed in this work. The estimated value of $(\Gamma_{\text{post}})_{13}$ does not depend on \bar{x}_{prior} for the constant strain-rate test, whereas a dependency can be observed for the other two tests. Furthermore, for $(\Gamma_{\text{post}})_{23}$ we can only see a dependency for the relaxation test.

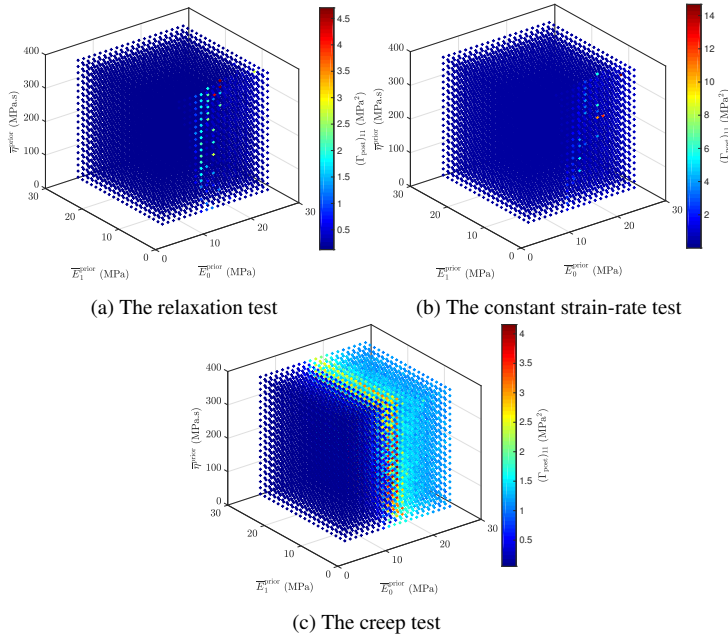


Fig. 29: The effect of $\bar{\mathbf{x}}_{\text{prior}}$ on $(\Gamma_{\text{post}})_{11}$. Only for the creep test a trend can be distinguished.

The values estimated for $(\Gamma_{\text{post}})_{12}$ using the three tests are independent of the mean of the prior, although the value change for each test.

Investigating the effect of the mean of the prior on the estimated values for the parameters and components of the posterior's covariance matrix shows that the constant strain-rate test is less sensitive to the prior's mean compared to the two other tests. This can be due to the existence of the term $\eta \dot{\epsilon}_0$ that does not vanish for $t \rightarrow \infty$ (see Eq. (3)), whereas a constant remains for the other two tests. The nature of the different tests therefore seems to have a considerable influence on the results.

7 Conclusions

In this contribution Bayesian inference is used to identify parameters in viscoelasticity. The model focused on was the standard linear solid model and the experiments of interest were a relaxation test, a constant strain-rate test and a creep test in tension. We have investigated four issues. First, the influence of the number of measurements on the parameter values and their uncertainty was investigated. Second, the identified parameter values were compared to those identified by the least squares method (LSM) for the relaxation test. Thirdly, three different types of experiments were considered in order to study how they affect the results. Finally, the sensitivity on the prior distribution was investigated for all the aforementioned cases.

The obtained results allow us to draw four conclusions:

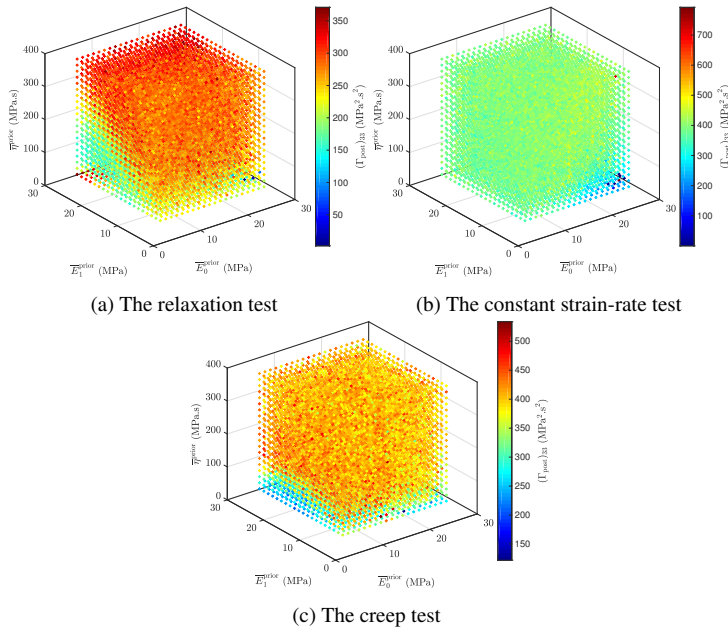


Fig. 30: The effect of $\bar{\mathbf{x}}_{\text{prior}}$ on $(\bar{\Gamma}_{\text{post}})_{33}$. The influence for the relaxation test is most significant, whereas almost no influence can be observed for the constant strain-rate test.

- (1) The prior has a substantially larger effect on the identified values (i.e. the mean and the MAP) in viscoelasticity than in elastoplasticity. The influence on the damping parameter is especially large. The reason for the prior's substantial influence in viscoelasticity compared to elastoplasticity is that no clear domains can be distinguished in viscoelasticity, whereas two domains are present in elastoplasticity (in which an elastic and elastoplastic domain can be distinguished).
- (2) An increase of the number of measurement points reduces the influence of the prior, but the influence on the damping parameter is still substantially present. If only two measurement points are present furthermore, the LSM has no unique solution but the prior in the Bayesian framework regularises this case such that it becomes unique.
- (3) For the relaxation test with more than two measurements, the stress-time responses associated with the LSM results are practically always closer to the true responses than the responses associated with the mean and MAP values identified by the Bayesian framework. It is however questionable if this means that the LSM is 'better' than the Bayesian framework. After all, the Bayesian framework incorporates the assumption that other measurements could have been made as well (incorporated in the prior). Furthermore, the Bayesian framework treats the parameters as random variables and consequently, the parameter values come with a credible region. This is not the case for the LSM. In other words, if one desires to propagate the uncertainty of the parameters in a predictive model, the

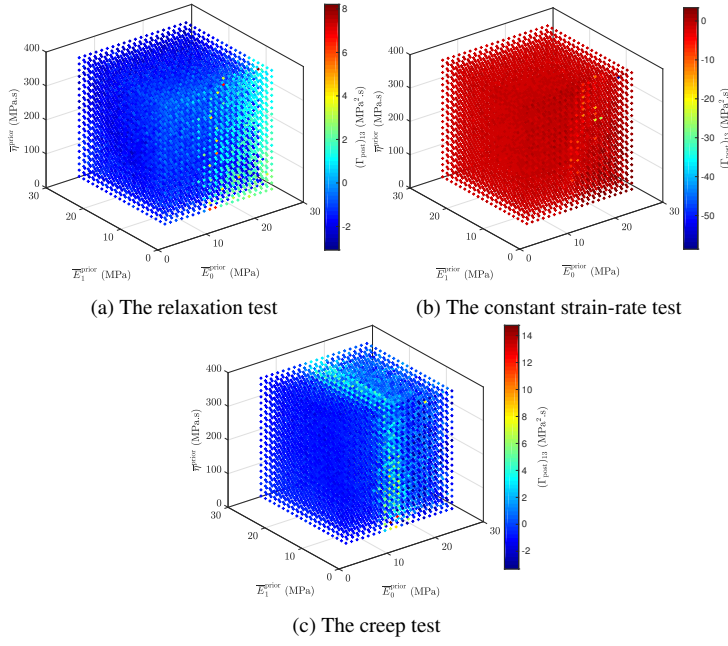


Fig. 31: The effect of \bar{x}_{prior} on $(\Gamma_{\text{post}})_{13}$ for the three employed tests. Unlike the constant strain-rate test, the estimated value for $(\Gamma_{\text{post}})_{13}$ is function of the mean of the prior for the two other tests.

standard LSM gives no estimate for the interval of a particular parameter. The Bayesian framework is however able to do this, including the dependency on the other parameters.

- (4) The use of BI has enabled us to distinguish that the uncertainty of the identified parameter values resulting from the constant strain-rate test is substantially larger than those resulting from the relaxation test and the creep test. This can be explained by the fact that asymptotic behaviour occurs in the relaxation and creep test, for which only one parameter is responsible. Consequently, the uncertainty of this parameter is relatively small and because of that, the uncertainties of the other parameters are also relatively small compared to those resulting from the constant strain-rate experiment. The results however show that this does not mean that the influence of the prior is larger.

The Bayesian identification approach and the LSM both have their advantages and disadvantages to identify material parameters. BI incorporates a regularisation that makes cases with few measurements solvable, which cannot be solved using the standard formulation of the LSM. An result of BI, which cannot be obtained using the standard LSM, is that the parameters come with uncertainties. This is essential for the propagation of uncertainties in mechanical predictions. On the other hand, the prior information in BI can also significantly influence the results negatively. Probably, it will always remain a matter of taste to prefer one method above the other. In case of

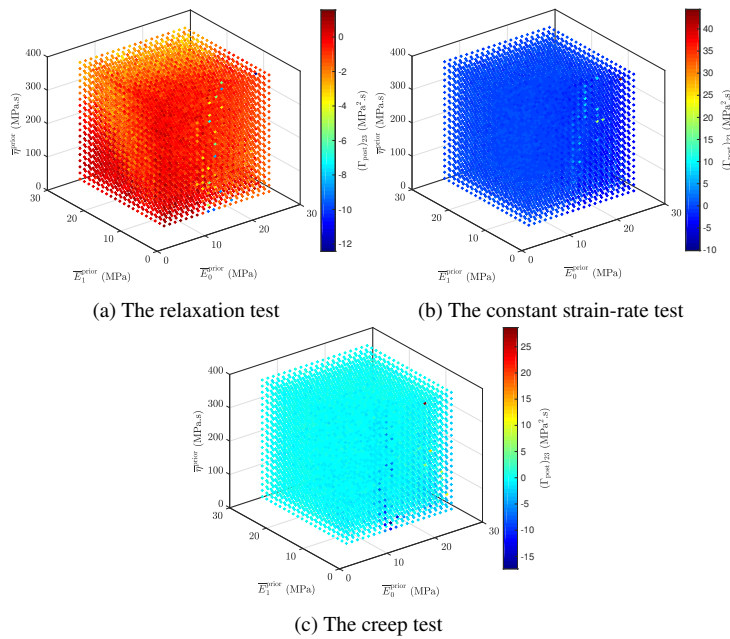


Fig. 32: The effect of $\bar{\mathbf{x}}_{\text{prior}}$ on $(\bar{\mathbf{T}}_{\text{post}})_{23}$ for the three employed tests. Except for the relaxation test, the other tests are independent of the prior's mean.

a small number of measurements and obtaining an uncertainty however, BI seems to be unavoidable.

Acknowledgements The authors would like to acknowledge the financial support from the University of Luxembourg, the European Research Council Starting Independent Research Grant (ERC Stg grant agreement No. 279578) entitled “Towards real time multiscale simulation of cutting in nonlinear materials with applications to surgical simulation and computer guided surgery” and the Luxembourg National Research Fund (project No. INTER/FNRS/15/11019432).

References

- Abhinav, S., Manohar, C.: Bayesian parameter identification in dynamic state space models using modified measurement equations. *INT. J. NONLINEAR. MECH.* **71**, 89-103 (2015)
- Alvin, K.: Finite element model update via Bayesian estimation and minimization of dynamic residuals. *AIAA. J.* **35** (5), 879-886 (1997)
- Andrieu, C., De Freitas, N., Doucet, A., Jordan, M.I.: An introduction to MCMC for machine learning. *MACH. LEARN.* **50** (1-2), 5-43 (2003)
- Babuška, I., Sawlan, Z., Scavino, M., Szabó, B., Tempone, R.: Bayesian inference and model comparison for metallic fatigue data. *COMPUT. METHOD. APPL. M.* **304**, 171-196 (2016)

- Banks, H.T., Hu, S., Kenz, Z.R.: A brief review of elasticity and viscoelasticity for solids. *Adv. Appl. Math Mech.* **3** (1), 1-51 (2011)
- Beck, J.L., Katafygiotis, L.S.: Updating models and their uncertainties. I: Bayesian statistical framework. *J. ENG. MECH-ASCE.* **124** (4), 455-461 (1998)
- Beck, J.L., Au, S.K.: Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *J. ENG. MECH-ASCE.* **128** (4), 380-391 (2002)
- Beex, L.A.A., Verberne, C.W., Peerlings R.H.J.: Experimental identification of a lattice model for woven fabrics: Application to electronic textile. *COMPOS. PART. A-APPL. S.* **48**, 82-92 (2013)
- Brooks, S., Gelman, A., Jones, G., Meng, X.L.: *Handbook of Markov chain Monte Carlo.* CRC press (2011)
- Calvetti, D., Somersalo, E.: *An introduction to Bayesian scientific computing: ten lectures on subjective computing, vol 2.* Springer Science & Business Media (2007)
- Chaparro, B.M., Thuillier, S., Menezes, L.F., Manach, P.Y., Fernandes, J.V.: Material parameters identification: Gradient-based, genetic and hybrid optimization algorithms. *COMP. MATER. SCI.* **44** (2), 339-346 (2008)
- Chiachío, J., Chiachío, M., Saxena, A., Sankararaman, S., Rus, G., Goebel, K.: Bayesian model selection and parameter estimation for fatigue damage progression models in composites. *INT. J. FATIGUE.* **70**, 361-373 (2015)
- Daghia, F., de Miranda, S., Ubertini, F., Viola, E.: Estimation of elastic constants of thick laminated plates within a Bayesian framework. *COMPOS. STRUCT.* **80** (3), 461-473 (2007)
- Elster, C., Wübbeler, G.: Bayesian regression versus application of least squares an example. *METROLOGIA.* **53** (1), S10-S16 (2016)
- Fitzenz, D.D., Jalobeanu, A., Hickman, S.H.: Integrating laboratory creep compaction data with numerical fault models: A Bayesian framework. *J. GEOPHYS. RES-SOL.* **112**, B08410 (2007)
- Gelman, A., Roberts, G.O., Gilks, W.R.: Efficient Metropolis jumping rules. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (eds) *Bayesian statistics 5*, pp. 599-607. Oxford University Press, New York (1996)
- Genovese, K., Lamberti, L., Pappalettere, C.: Improved global-local simulated annealing formulation for solving non-smooth engineering optimization problems. *INT. J. SOLIDS STRUCT.* **42** (1), 203-237 (2005)
- Gogu, C., Haftka, R., Molimard, J., Vautrin, A.: Introduction to the Bayesian approach applied to elastic constants identification. *AIAA. J.* **48** (5), 893-903 (2010)
- Gogu, C., Yin, W., Haftka, R., Ifju, P., Molimard, J., Le Riche, R., Vautrin, A.: Bayesian identification of elastic constants in multi-directional laminate from moiré interferometry displacement fields. *EXP. MECH.* **53** (4), 635-648 (2013)
- Goldberg, D.E.: *Genetic algorithms in search, optimization, and machine learning.* Addison-Wesley Publishing Company (1989)
- Haario, H., Saksman, E., Tamminen, J.: Adaptive proposal distribution for random walk Metropolis algorithm. *Computational Statistics.* **14** (3), 375-396 (1999)
- Hernandez, W.P., Borges, F.C.L., Castello, D.A., Roitman, N., Magluta, C.: Bayesian inference applied on model calibration of fractional derivative viscoelastic model. In: Steffen Jr, V., Rade, D.A., Bessa, W.M. (eds) *DINAME 2015-*

- Proceedings of the XVII International symposium on dynamic problems of mechanics, Natal (2015)
- Higdon, D., Lee, H., Bi, Z.: A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine scale information. *IEEE. T. SIGNAL. PROCES.* **50**, 388-399 (2002)
- Isenberg, J.: Progressing from least squares to Bayesian estimation. Proceedings of the 1979 ASME design engineering technical conference, 1-11 (1979)
- Kaipio, J., Somersalo, E.: Statistical and computational inverse problems, vol **160**. Springer Science & Business Media (2007)
- Kenz, Z.R., Banks, H.T., Smith, R.C.: Comparison of frequentist and Bayesian confidence analysis methods on a viscoelastic stenosis model. *SIAM/ASA Journal on Uncertainty Quantification.* **1** (1), 348-369 (2013)
- Koutsourelakis, P.S.: A novel Bayesian strategy for the identification of spatially varying material properties and model validation: An application to static elastography. *INT. J. NUMER. METH. ENG.* **91** (3), 249-268 (2012)
- Kristensen, J., Zabaras, N.: Bayesian uncertainty quantification in the evaluation of alloy properties with the cluster expansion method. *COMPUT. PHYS. COMMUN.* **185** (11), 2885-2892 (2014)
- Lai, T.C., Ip K.H.: Parameter estimation of orthotropic plates by Bayesian sensitivity analysis. *COMPOS. STRUCT.* **34** (1), 29-42 (1996)
- Liu, P., Au, S.K.: Bayesian parameter identification of hysteretic behavior of composite walls. *PROBABILIST. ENG. MECH.* **34**, 101-109 (2013)
- Liu, X.Y., Chen, X.F., Ren, Y.H., Zhan, Q.Y., Wang, C., Yang, C.: Alveolar type II cells escape stress failure caused by tonic stretch through transient focal adhesion disassembly. *INT. J. BIOL. SCI.* **7** (5), 588-599 (2011)
- Madireddy, S., Sista, B., Vemaganti, K.: A Bayesian approach to selecting hyperelastic constitutive models of soft tissue. *COMPUT. METHOD. APPL. M.* **291**, 102-122 (2015)
- Magorou, L.L., Bos, F., Rouger, F.: Identification of constitutive laws for wood-based panels by means of an inverse method. *COMPOS. SCI TECHNOL.* **62** (4), 591-596 (2002)
- Marwala, T., Sibusiso, S.: Finite element model updating using Bayesian framework and modal properties. *J. AIRCRAFT.* **42** (1), 275-278 (2005)
- Marzouk, Y.M., Najm, H.N., Rahn, L.A.: Stochastic spectral methods for efficient Bayesian solution of inverse problems. *J. COMPUT. PHYS.* **224** (2), 560-586 (2007)
- Mehrez, L., Kassem, E., Masad, E., Little, D.: Stochastic identification of linear-viscoelastic models of aged and unaged asphalt mixtures. *J. MATER. CIVIL. ENG.* **27** (4), 04014149 (2015)
- Miles, P., Hays, M., Smith, R., Oates, W.: Bayesian uncertainty analysis of finite deformation viscoelasticity. *MECH. MATER.* **91** (Part 1), 35-49 (2015)
- Most, T.: Identification of the parameters of complex constitutive models: Least squares minimization vs. Bayesian updating. In: Straub, D. (ed) Reliability and optimization of structural systems, pp. 119-130. CRC press (2010)
- Muto, M., Beck, J.L.: Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *J. VIB. CONTROL.* **14** (1-2), 7-34

- (2008)
- Nichols, J.M., Link, W.A., Murphy, K.D., Olson, C.C.: A Bayesian approach to identifying structural nonlinearity using free-decay response: Application to damage detection in composites. *J. SOUND VIB.* **329** (15), 2995-3007 (2010)
- Oden, J.T., Prudencio E.E., Hawkins-Daarud A.: Selection and assessment of phenomenological models of tumor growth. *MATH MOD. METH. APPL. S.* **23** (7), 1309-1338 (2013)
- Oh, C.K., Beck, J.L., Yamada, M.: Bayesian learning using automatic relevance determination prior with an application to earthquake early warning. *J. ENG. MECH-ASCE.* **134** (12), 1013-1020 (2008)
- Orosz, M., Molnarka, G., Monos, E.: Curve fitting methods and mechanical models for identification of viscoelastic parameters of vascular wall-a comparative study. *MED. SCI. MONITOR.* **3** (4), MT599MT604 (1997)
- Rappel, H., Beex, L.A.A., Hale, J.S., Bordas, S.P.A.: Bayesian inference for the stochastic identification of elastoplastic material parameters: Introduction, misconceptions and additional insight. arXiv preprint (2016). arXiv:160602422
- Rosić, B.V., Kčerová, A., Sýkora, J., Pajonk, O., Litvinenko, A., Matthies, H.G.: Parameter identification in a probabilistic setting. *ENG. STRUCT.* **50**, 179-196 (2013)
- Sarkar, S., Kosson, D.S., Mahadevan, S., Meeussen J.C.L., van der Sloot, H., Arnold J.R., Brown, K.G.: Bayesian calibration of thermodynamic parameters for geochemical speciation modeling of cementitious materials. *CEMENT. CONCRETE. RES.* **42** (7), 889-902 (2012)
- Ulrych T.J., Sacchi, M.D., Woodbury, A.: A Bayes tour of inversion: A tutorial. *GEOPHYSICS.* **66** (1), 55-69 (2001)
- Wang, J., Zabaras, N.: A Bayesian inference approach to the inverse heat conduction problem. *INT. J. HEAT MASS. TRAN.* **47** (17-18), 3927-3941 (2004)
- Zhang, E., Chazot, J.D., Antoni, J., Hamdi, M.: Bayesian characterization of Young's modulus of viscoelastic materials in laminated structures. *J. SOUND VIB.* **332** (16), 3654-3666 (2013)
- Zhao, X., Pelegri, A.A.: A Bayesian approach for characterization of soft tissue viscoelasticity in acoustic radiation force imaging. *INT. J. NUMER. METH. BIOMED. ENGGNG.* **32** (4), e02741 (2016)