

Integrating Laboratory Compaction Data With Numerical Fault Models: a Bayesian Framework

D. D. Fitzenz*, A. Jalobeanu[†], S. H. Hickman* and N. Sleep**

*USGS, EHZ team, Menlo Park, CA, USA

[†]LSIIT (UMR 7005 CNRS-ULP) @ ENSPS, Illkirch, France

**Geophysics Department, Stanford University, CA, USA

Abstract. When analyzing rock deformation experimental data, one deals with both uncertainty and complexity. Though each part of the problem might be simple, the relationships between them can form a complex system. This often leads to partial or only qualitative data analyses from the experimental rock mechanics community, which limits the impact of these studies in other communities (e.g., modelling). However, it is a perfect case study for graphical models.

We present here a Bayesian framework that can be used both to infer the parameters of a constitutive model from rock compaction data, and to simulate porosity reduction within direct fault models from a known (e.g. lab-derived) constitutive relationship, while keeping track of all the uncertainties. This latter step is crucial if we are to go toward process-based seismic hazard assessment. Indeed, the rate of effective stress build-up¹ (namely due to fault compaction) as well as the recovery of fault strength determine how long it will take for different parts of the previously ruptured fault to reach failure again, thus controlling both the timing and the size of the next rupture. But deterministic models need to rigorously incorporate uncertainties if they are to be useful in creating probabilistic assessments of seismic hazard. It is therefore important to work within a framework able to assess model validity as well as use data uncertainties.

Our approach involves a hierarchical inference scheme using several steps of marginalization. Existing experimental data are rarely adequate to completely define a single constitutive relationship for given physical fault material parameters over temperature and effective confining pressures of relevance to actual fault zones. We therefore focus on one rather general, though experimentally derived, compaction law to illustrate how applying the proposed inference scheme on simulated data can help design compaction experiments to provide better constraints on creep parameters.

Keywords: Bayesian inference, marginalization, rock physics, fault modeling

INTRODUCTION

Long-term goals of this study. To facilitate utilization of laboratory-derived compaction creep laws in theoretical models of the earthquake cycle, it is important that methods be developed which allow uncertainties in these experimentally derived compaction laws to be fully accounted for. Traditionally, however, this has proven to be quite difficult, in part because laboratory data are typically obtained at pressure, temperature and strain rate conditions much different than those that exist in natural fault zones in the earth. To help develop a more robust technique for extrapolating lab fault compaction data to models for time-dependant pore pressure evolution in natural fault zones, we developed a Bayesian inversion method specifically to achieve the following goals:

¹ Effective stress = normal stress minus pore pressure. Its variations can come from tectonic stress transfer projected perpendicular to the fault strike, or from pore pressure changes.

1. Analyze compaction experiments performed on natural and synthetic fault-gouge samples in the lab or from borehole observations of fluid pressure changes in faults.
2. Help design lab experiments to best constrain constitutive parameters for a given type of creep law.
3. Identify changes in mechanical behavior and rate-controlling process during an experiment.
4. Facilitate quantitative extrapolation of constitutive relationships derived from lab experiments to forward models of fault-zone compaction and fluid pressure evolution, as opposed to made-up compaction laws (see [1]).

To do so, we chose a simple expression for a general creep law.

Mathematically, the problem then consists of inverting a set of couples of measurements, seen as input and output of a given transformation law (here, creep), for the parameters of that law. As shown below, this approach is more general than fitting an assumed creep law to a set of experimentally derived porosity time series, in that it allows the processing of other data types (e.g., from boreholes or geophysical field measurements) into the inversion.

When translating the problem into directed graphical models, its hierarchical structure becomes apparent. However, the natural parameters exhibit a nonlinear parametric form. The proposed method consists of finding an optimal parametrization, having a behavior as linear as possible, then perform whatever inference is required, and finally go back to the original parametrization to obtain a physically meaningful constitutive relationship.

CREEP LAWS: THEORY & EXPERIMENTAL DETERMINATION

The creep behavior of rocks depends on the temperature (T), the presence or absence of fluids, the effective confining pressure (the confining pressure minus the pore pressure), the mineralogy and the grain-size distribution of the rock assemblage. A number of theoretical works have been conducted on creep [2, 3], many of which within the broad Material Sciences community rather than the Earth Sciences community. Experimental works were also carried out on Earth materials [4]. Many processes were shown to be active at these conditions [5, 6, 7]. Most of the time, the interpretation of the experimental data was made difficult because the deformation mechanisms controlling compaction were too complex to be fit by a single rate-law, or the range of experimental T,P conditions used were too narrow to allow for a well-determined inversion [8]. Besides, in most cases, the data were analysed “independently” for the different parameters of the laws, and the uncertainties were not evaluated at each step and propagated to the next, making it difficult to assess the robustness of the results. Therefore, and despite the wealth of data obtained for rock creep, creep laws deriving from quantitative analyses of time-dependent compaction of porous rock or fault gouge in the lab at hydrothermal conditions are scarce.

Our strategy: a general creep law, and a unified inversion scheme

Most theoretical laws (dislocation creep, solution transfer, cataclastic flow, stress corrosion) are characterized by 1) a stress exponent, 2) an activation energy, 3) a poros-

ity term, 4) another term, usually dependent on the grain-size [9, 10]. However, for phyllosilicate-bearing gouges, the grain-size is hard to assess. In addition, in natural rocks, more than one process can be occurring at a given time (maybe over different time scales). We therefore choose to invert lab time-dependant compaction data for a general creep law of the form [11]:

$$\frac{\partial \varphi}{\partial t} = \theta_0 \sigma_{\text{eff}}^{\theta_1} e^{-\theta_2/(RT)} e^{\theta_3 \varphi} \quad (1)$$

giving the porosity reduction rate as a function of the effective stress σ_{eff} , temperature T , and instantaneous porosity φ , where θ_0 is a multiplying factor (accounting for e.g. grain size effects), θ_1 is the stress exponent, θ_2 is the activation energy, θ_3 is a porosity term. In such experiments, each porosity measurement (in %) is determined to an accuracy of ± 1 to 2%. We will assume that this interval corresponds to a 95% confidence interval. There are multiple sources of error, depending on the measurement method. We assume that the total error is Gaussian, due to the large number of sources, and that each measurement (or observation) is a realization of a Gaussian random variable of standard deviation 0.5%. In the following, we will evaluate the probability density function (pdf) of the parameters of the general creep law (i.e., equation 1, which we will call *model*) given experimental data for T , σ_{eff} and changes in porosity over time. To be able to compute the pdf, we will first study the relationships between all the elements of our problem.

Although we will be using only one 'model', one could test different creep laws. In particular, once θ_1 and θ_2 in Eq. (1) are determined, one could propose a probable dominant mechanism and try the same inversion scheme for more appropriate creep laws. The Bayesian framework is a natural framework for model selection and assessment, allowing for the determination of the best creep law given a particular dataset.

DIRECT MODEL AND BAYESIAN INFERENCE

Direct model. We will be using directed graphical models to visualize the relationships between observations, parameters, and compaction models². Our model is displayed in Figure 1. To mimic realistic circumstances, we show the relationships between the variables for a set of i observations of pre- and post- compaction porosities for each one of the n temperature and effective stress conditions. The factorial structure of the graph is due to the experimental design: sets of observations and experiments are symbolized by rectangles or plates [12].

Because we want to determine a creep law of the form of equation 1, the compaction model m consists of the function f , which is a solution to this differential equation:

$$f(\varphi) = -\frac{1}{\alpha} \log(\alpha \gamma + e^{-\alpha \varphi}) \quad (2)$$

$$\text{where } \gamma = -\Delta t \theta_0 \sigma_{\text{eff}}^{\theta_1} e^{-\theta_2/(RT)} \quad \text{and} \quad \alpha \equiv \theta_3 \quad (3)$$

² Each node represents a set of random variables. Arrows converging to a node denote a conditional pdf. The terminal nodes without incoming arrows are prior pdfs. The joint pdf of all variables is then given by the product of all prior and conditional pdfs according to the graph structure.

First of all, we need to reparametrize the problem, since it is heavily nonlinear. Our numerical experiments have shown how difficult it is to estimate the parameters when they exhibit a nonlinear parametric form. The proposed solution consists of finding an optimal parametrization, having a behavior as linear as possible. We then perform whatever inference is required, and finally go back to the original parametrization, to obtain a physically meaningful solution. Here is the proposed reparametrization:

$$\lambda = -\frac{1}{\alpha} \log(\alpha \gamma), \mu = \frac{1}{\alpha} \text{ and } \Theta = F(\theta) = \{\log(-\theta_0)/\theta_3, \theta_1/\theta_3, \theta_2/\theta_3, 1/\theta_3\} \quad (4)$$

hence the new expression, with the new constants $k = \{\log \Delta t, \log \sigma_{\text{eff}}, 1/(RT)\}$:

$$f(\varphi) = -\mu \log(e^{-\lambda/\mu} + e^{-\varphi/\mu}) \quad (5)$$

$$\text{where } \lambda = h(\Theta) = -\Theta_0 - k_1 \Theta_1 + k_2 \Theta_2 - k_0 \Theta_3 + \Theta_3 \log \Theta_3 \quad \text{and} \quad \mu \equiv \Theta_3 \quad (6)$$

Because of the complexity of the graph structure, we use a hierarchical approach to the inference. Each one of the n experiments is treated individually (step **A**) before proceeding to the final estimation of the parameters. Indeed, in such a '1 to n ' tree structure, we can perform an independent inference on the tree branches. This is why we introduced the n parameters v , that are the estimates of μ for each experiment. Thus, after Step A is completed, we have n joint posterior pdfs of (λ, v) . Finally, Step **B** consists of inferring the parameters Θ from these pdfs, then reverting to the original θ .

The direct model allows to compute the joint probability of all parameters, observations, and models [13]. We can therefore express *for each experiment n* the joint probability in the following simplified way:

$$P(\{\varphi^i, \phi^i, \varphi_{\text{obs}}^i, \phi_{\text{obs}}^i\}, \lambda, v) = \prod_i P(\varphi^i) P(\varphi_{\text{obs}}^i | \varphi^i) P(\phi^i) P(\phi_{\text{obs}}^i | \phi^i) P(\phi^i | \varphi^i, \lambda, v) \quad (7)$$

where the upper index i spans the number of observations for each experiment.

Inverse problem. Now that the direct problem is clearly stated, it is easy to formulate the inverse problem, or *Bayesian inference*. It consists of finding the probability density for all the parameters of interest knowing all the observations (i.e. the posterior pdf).

However, there are more nodes in the graph than just the parameters and the observations. We therefore have to integrate the complete joint pdf (given by the complete graphical model) with respect to all the intermediate nodes, a.k.a. nuisance variables. This is the *marginalization* step.

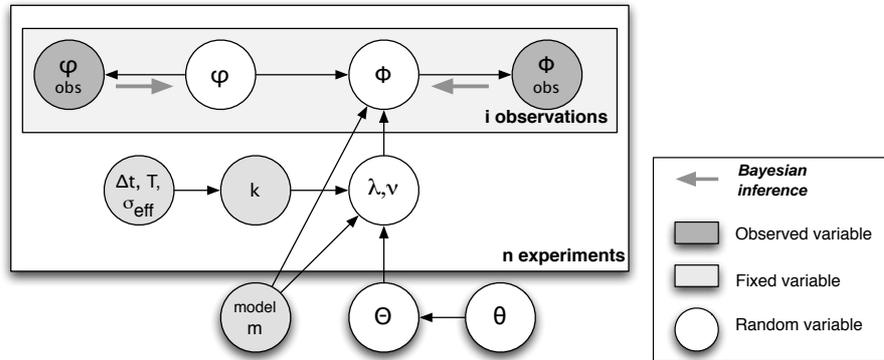


FIGURE 1. Graphical model representing the direct (compaction) model.

To perform the inference, we then only have to use Bayes' theorem, which simply states that the posterior pdf is proportional to the (restricted) joint pdf. Our inference method consists of approximating the posterior pdf by a Gaussian around its mode. Therefore the last step is one of *optimization*. The fewer the dimensions, the easier the optimization, hence the advantage of the marginalization. We can compute the covariance matrix related to the Gaussian pdf by taking the inverse of the matrix of second derivatives of the log posterior at the optimum.

Message passing. Because of the tree structure of the graph, we can use the message passing algorithm and compute in a recursive fashion the posterior pdf of any level of the tree as a function of the conditional pdfs of the levels above it.

At each step, we will use as observations a Gaussian approximation of the posterior pdf found during the previous step [14]. In our simple case, we are using a hierarchical (i.e., recursive) approach in which we first evaluate the joint pdf of λ and ν knowing the observations for each experiment (step A). We then use the set of Gaussian approximation of the resulting posterior pdf of (λ^n, ν^n) as conditional pdfs in the last step, with a mean and covariance matrix respectively denoted by $(\lambda_{\text{obs}}^n, \nu_{\text{obs}}^n)$ and A^n . This is equivalent to the message passing algorithm. We can write the corresponding restricted joint pdf (Eq. 8), and once again, with a step of marginalization and the use of Bayes' theorem, we will derive the pdf of the parameters Θ , then transform this pdf to get the distribution of the creep law parameters θ (step B).

$$P(\Theta, \{\lambda^n, \nu^n, \lambda_{\text{obs}}^n, \nu_{\text{obs}}^n\}) = P(\Theta) \prod_n P(\lambda^n, \nu^n | \Theta) P(\lambda_{\text{obs}}^n, \nu_{\text{obs}}^n | \lambda^n, \nu^n) \quad (8)$$

Inversion method for step A

Step A consists in computing the posterior marginals of (λ, ν) for each experiment, knowing the observations $\{\varphi_{\text{obs}}^i\}$ and $\{\phi_{\text{obs}}^i\}$.

Marginalization. We know that $P(\lambda, \nu | \{\varphi_{\text{obs}}^i, \phi_{\text{obs}}^i\})$ is proportional to the restricted joint probability $P(\lambda, \nu, \{\varphi_{\text{obs}}^i, \phi_{\text{obs}}^i\})$, which can be obtained from the complete joint distribution (equation 7) by integrating out the intermediate parameters φ and ϕ . We can therefore write the product over all the observations i gathered for a given experiment:

$$P(\lambda, \nu | \{\varphi_{\text{obs}}^i, \phi_{\text{obs}}^i\}) \propto \prod_i \int \int P(\varphi^i) P(\varphi_{\text{obs}}^i | \varphi^i) P(\phi^i) P(\phi_{\text{obs}}^i | \phi^i) P(\phi^i | \varphi^i, \lambda, \nu) d\varphi^i d\phi^i$$

We can readily simplify this expression by noticing that we have a deterministic relationship f between ϕ^i and $(\varphi^i, \lambda, \nu)$. The pdf $P(\phi^i | \varphi^i, \lambda, \nu)$ is nothing but the Dirac distribution $\delta(f(\varphi^i, \lambda, \nu) - \phi^i)$. We can also use here the Gaussian assumption that we made, namely that all the porosity distributions: φ_{obs}^i follow a Gaussian distribution around the mean φ^i with a variance σ^2 ; the same holds for ϕ_{obs}^i and ϕ^i , respectively.

After simplification, each double integral can be written as:

$$I(\lambda, \nu, \varphi_{\text{obs}}^i, \phi_{\text{obs}}^i) = \int_0^{100} G_{\varphi}(\varphi_{\text{obs}}^i, \sigma^2) G_{f(\varphi, \lambda, \nu)}(\phi_{\text{obs}}^i, \sigma^2) d\varphi^i \quad (9)$$

where $G_x(\mu, \sigma^2)$ is the Gaussian distribution of x , with mean μ and standard deviation σ . There is no analytic solution to this integral. We chose to approximate f by a piecewise linear function in order to express I and its derivatives using Gaussian integrals.

The inference amounts to finding a Gaussian approximation of the posterior pdf – i.e., a quadratic approximation of the energy:

$$U(\lambda, \mathbf{v}) = -\log P(\lambda, \mathbf{v} | \{\phi_{\text{obs}}^i, \phi_{\text{obs}}^i\}) = -\sum_i \log I(\lambda, \mathbf{v}, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i) \quad (10)$$

The minimum of U (related to the posterior mode) provides the optimal parameters and the second derivative of U at the optimum gives the uncertainty.

The energy $U(\lambda, \mathbf{v})$ exhibits an almost quadratic behavior near its optimum, whereas $U(\alpha, \gamma)$ computed with the original parametrization presents a very curved and asymmetric narrow valley, not suitable for a good determination of the optimum. Thus, choosing a good parametrization not only enables us to perform an efficient optimization, but also allows for an accurate Gaussian approximation of the posterior pdf, which is an important requirement for the two-step inference method proposed here.

Posterior pdf computation. Let us first focus on finding the posterior mode. Various methods could be used. However, the energy landscape is nearly quadratic only near the optimum. With initial parameter values far from the optimum, it proved difficult to apply steepest descent, or even conjugate gradient algorithms. Because of the low dimensionality of this problem, a rather simple but efficient solution consists of performing two nested optimizations – first on λ , then on \mathbf{v} . Thus, for each \mathbf{v} , a line minimization is performed with respect to (w.r.t.) λ . The inner minimization loop is achieved through dichotomy, followed by a secant method to refine the parameter λ , involving the derivative of U w.r.t. λ . The outer loop has the same structure, and only uses the derivative of U w.r.t. \mathbf{v} .

Once the optimization is done, we need to take the second derivatives. We approximate the integral I by a linear function of the parameters near the optimum, hence:

$$\frac{\partial^2 U}{\partial u \partial u'} \simeq \sum_i \frac{1}{I(\lambda, \mathbf{v}, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)^2} \frac{\partial I(\lambda, \mathbf{v}, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)}{\partial u} \frac{\partial I(\lambda, \mathbf{v}, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)}{\partial u'} \quad (11)$$

These 3 derivatives computed at the optimum $(\hat{\lambda}, \hat{\mathbf{v}})$ provide the entries of the symmetric inverse covariance matrix defining the shape of the bivariate Gaussian distribution, which is a good approximation of the posterior pdf:

$$P(\lambda, \mathbf{v} | \{\phi_{\text{obs}}^i, \phi_{\text{obs}}^i\}) \simeq G_{(\lambda, \mathbf{v})}((\hat{\lambda}, \hat{\mathbf{v}}), A) \quad \text{where} \quad A_{uv} = \frac{\partial^2 U}{\partial u \partial v} \quad (12)$$

Inversion method for step B

Computing the posterior pdf of Θ . We refer to the joint pdf given in Eq. (8). From Step A, we consider each optimum previously denoted by $(\hat{\lambda}, \hat{\mathbf{v}})$ as a noisy observation $(\lambda_{\text{obs}}^n, \mathbf{v}_{\text{obs}}^n)$ of the parameters $(\lambda^n, \mathbf{v}^n)$. Thus, we have $P(\lambda_{\text{obs}}^n, \mathbf{v}_{\text{obs}}^n | \lambda^n, \mathbf{v}^n) = G((\lambda^n, \mathbf{v}^n), A)$, i.e. a Gaussian pdf with mean $(\lambda^n, \mathbf{v}^n)$ and inverse covariance matrix A .

Let us call $h(\Theta)$ the theoretical deterministic relationship between the vector Θ and λ , as given in Eq. (6). Then we get $P(\lambda^n, \mathbf{v}^n | \Theta, m^n) = \delta(\lambda^n - h(\Theta)) \delta(\mathbf{v}^n - \Theta_3)$.

What we need to compute is $P(\Theta | \{\lambda_{\text{obs}}^n, \mathbf{v}_{\text{obs}}^n\})$. We therefore need to integrate the joint probability (8) with respect to λ^n and \mathbf{v}^n , which is simple because of the Dirac function. We get the product:

$$P(\Theta | \{\lambda_{\text{obs}}^n, \mathbf{v}_{\text{obs}}^n\}) = P(\Theta) \prod_n G_{(h(\Theta), \Theta_3)}((\lambda_{\text{obs}}^n, \mathbf{v}_{\text{obs}}^n), A) \quad (13)$$

We again use the energy function $U'(\Theta) = -\log P(\Theta | \{\lambda_{obs}^n, v_{obs}^n\})$ instead:

$$U'(\Theta) = -\log P(\Theta) + \frac{1}{2} \sum_n A_{00}^n (h(\Theta) - \lambda_{obs}^n)^2 + A_{11}^n (\Theta_3 - v_{obs}^n)^2 + 2A_{10}^n (h(\Theta) - \lambda_{obs}^n) (\Theta_3 - v_{obs}^n) \quad (14)$$

To find the optimal value of each Θ_j , we need to minimize this energy. We assume a noninformative, uniform distribution of each Θ_j , to simplify the problem. The function h is linear w.r.t. all parameters but Θ_3 , so that U is quadratic w.r.t. these parameters. We perform a 1D minimization w.r.t. Θ_3 ; at each step (evaluation of U' and first derivative w.r.t. Θ_3) all other parameters are optimized by solving a 3×3 linear system.

The second derivatives are computed using a similar approximation as in Step A: we assume a linear behavior for h around the optimum.

$$\frac{\partial^2 U'}{\partial \Theta_j \partial \Theta_l} = \sum_n \frac{\partial h}{\partial \Theta_j} \left(A_{0\varepsilon}^n \frac{\partial h}{\partial \Theta_l} + A_{1\varepsilon}^n \varepsilon' \right) \quad \text{where } \varepsilon = \delta_{j,3}, \varepsilon' = \delta_{l,3} \quad (15)$$

Once we have minimized U' to find the optimal $\hat{\Theta}$, the second derivatives above provide the coefficients B_{jl} of the inverse covariance matrix B of the Gaussian approximation of the posterior pdf of Θ :

$$P(\Theta | \{\lambda_{obs}^n, v_{obs}^n\}) \simeq G_{\Theta}(\hat{\Theta}, B) \quad \text{where } B_{uv} = \frac{\partial^2 U'}{\partial u \partial v} \quad (16)$$

Converting Θ into the creep law parameters θ . Let us denote by F the transform such that $\Theta = F(\theta)$, as specified by Eq. (4). Let J denote the Jacobian matrix of the transform, i.e. the matrix of derivatives of Θ_j w.r.t. θ_l . We make a Laplace approximation, consisting of considering the distribution of θ as a Gaussian also. Then we can write the final result, where J is evaluated at the optimum, as:

$$P(\theta | \{\{\phi_{obs}^i, \phi_{obs}^i\}^n\}) \simeq P(\theta | \{\lambda_{obs}^n, v_{obs}^n\}) \simeq G_{\theta}(F^{-1}(\hat{\Theta}), (J^T B J)^{-1}) \quad (17)$$

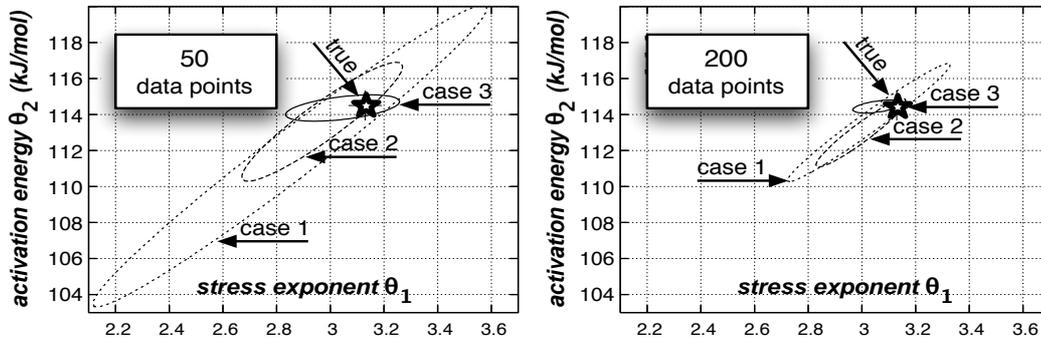
If one is only interested in a subset of the parameters θ , the related distribution is simply obtained by keeping the corresponding entries of Σ ; this will be used in the result section.

SIMULATED COMPACTION DATA: RESULTS & CONCLUSIONS

We model compaction numerically using the solution f to the creep law from [11]. We use equation 2 with the parameters just listed to compute the porosity time series for 3 different series of experiments (settings): **(1)** 6 exp., T from 300 K to 420 K and σ_{eff} from 15 MPa to 105 MPa; **(2)** 12 exp., same range; **(3)** 12 exp., T from 300 K to 720 K and σ_{eff} from 15 MPa to 200 MPa. For each simulated experiment, we then separate the data in terms of pre- and post-compaction porosities. This is done by choosing a compaction duration Δt and selecting couples $(\varphi^i = \phi(t_i), \phi^i = \phi(t_i + \Delta t))$ for a series of times t_i . Finally, we add a Gaussian noise to the porosity with a standard deviation of 0.25% to simulate the acquisition process.

The results are given in the following table for settings 1 and 3. The respective correlations between θ_1 and θ_2 are 0.98 and 0.52, which clearly favors setting 3. The figure below shows the 95% confidence regions for these two parameters for the 3 series, with 50 (left) and 200 (right) data points: the larger the dataset, the lower the uncertainty.

	Exp. Setting #	$-\theta_0$	θ_1	θ_2 (kJ/mol)	θ_3
True value	all	$2.6 \cdot 10^{-7}$	3.13	114.5	0.73
Optimum	1	$4 \cdot 10^{-7}$	3.0	113	0.70
Marginal std. dev.	1	$1.2 \cdot 10^{-7}$	0.12	1.3	0.03
Optimum	3	$4 \cdot 10^{-7}$	3.06	114.4	0.72
Marginal std. dev.	3	$1.2 \cdot 10^{-7}$	0.04	0.14	0.01



Conclusion. We built a general Bayesian framework to analyze laboratory creep compaction experiments. To better condition the inverse problem, we chose to reparametrize and use a hierarchical approach. We successfully tested the model on a set of compaction simulations derived from real experiments. *We show that, despite the heavy non-linearity of the problem, we can retrieve accurate estimates of both the stress exponent and the activation energy, even when the porosity data are noisy.*

We also showed that whereas adding observation points and/or experiments reduces the uncertainty on all parameters, enlarging the range of temperature or effective stress reduces the covariance between stress exponent and activation energy significantly. This type of tests, performed on simulated data, could help better plan future experiments to optimize the determination of constitutive parameters for creep compaction under hydrothermal conditions.

REFERENCES

1. D. D. Fitzenz, and S. A. Miller, *Geophys. J. Int.*, **155**, 111–125 (2003).
2. E. H. Rutter, *J. Geol. Soc., London*, **140**, 725–740(16) (1983).
3. M. Paterson, *Reviews Geophys. Space Physics*, **11**, 355–389 (1973).
4. S. Hickman, and B. Evans, in *Fault Mechanics and Transport Properties of Rocks, Int. Geophys. Ser.*, pp. vol.51, edited by B. Evans and T.-f. Wong, pp. 253–280 (Academic, San Diego, Calif., 1992).
5. B. Bos, C. J. Peach, and C. J. Spiers, *Tectonophysics*, **327**, 173–194 (2000).
6. T. Dewers, and A. Hajash, *J. Geophys. Res.*, **100**, 13,093–13,112 (1995).
7. P. M. Schutjens, *J. Geol. Soc., London*, **148**, 527–539 (1991).
8. J. Chester, S. Lenz, F. M. Chester, and R. Lang, *Earth and Planetary Sciences Letters*, **220**, 435–451 (2004).
9. W. F. Brace, and D. L. Kohlstedt, *J. Geophys. Res.*, **85**, 6248–6252 (1980).
10. B. K. Atkinson, *Fracture mechanics of rocks*, Academic Press, London, 1987.
11. E. H. Rutter, and P. Wanten, *J. Sedimen. Res.*, **70**, 107–116 (2000).
12. M. I. Jordan, *Statistical Science*, **19** (Special Issue on Bayesian Statistics), 140–155 (2004).
13. J. Bernardo, and A. Smith, *Bayesian Theory*, Ed. John Wiley, 1994.
14. M. I. Jordan, and Y. Weiss, *Handbook of Neural Networks and Brain Theory*, 2nd edition, MIT Press, 2002, chap. Graphical models: probabilistic inference.
15. M. I. Jordan, editor, *Learning in Graphical Models*, MIT Press, 1998.