Integrating laboratory compaction data with numerical fault models:

a Bayesian framework

D.D. Fitzenz
USGS, Menlo Park, CA

A. Jalobeanu
LSIIT, CNRS, Strasbourg - France

S.H. Hickman
USGS, Menlo Park, CA

N. Sleep
Stanford Univ., Palo Alto, CA

MaxEnt’05
Aug. 2005, San Jose, CA
How creep acts on fault strength

Simplified structure of the Nojima fault

Porosity reduction → pore pressure increase → strength decrease, hence influence of compaction on the timing and size of earthquakes
A general compaction law

Most theoretical creep laws (e.g., diffusion creep, subcritical crack growth) include a stress exponent and an activation energy. Our choice:

$$\dot{\Phi} = \theta_0 \times \sigma_{\text{eff}}^{\theta_1} \times \exp\left(-\frac{\theta_2}{(RT)}\right) \times \exp(\theta_3 \Phi)$$

This law was inferred for fault gouge-like materials (Rutter & Wanten, 2001). After integration:

$$f(\Phi) = -\frac{1}{\alpha} \log \left(\alpha \gamma + e^{-\alpha \Phi}\right) = \Phi$$

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

\(\sigma_{\text{eff}}\): effective confining pressure

\(\Delta t\): duration of compaction

Chester et al, EPSL
**Optimal parametrization**

This problem is very non-linear, therefore inference can be difficult. We need to **reparametrize** it to make it “more linear”: \((\alpha, \gamma) \rightarrow (\lambda, \mu)\)

\[
\begin{align*}
\lambda &= -\frac{1}{\alpha} \log (\alpha \gamma), \\
\mu &= \frac{1}{\alpha}
\end{align*}
\]

The same is valid for \(\theta\), let us choose a more linear parametric form:

\[
\Theta = F(\theta) = \{\theta_0/\theta_3, \theta_1/\theta_3, \theta_2/\theta_3, 1/\theta_3\}
\]

<table>
<thead>
<tr>
<th>Original parametrization</th>
<th>New parametrization, more linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f(\varphi) = -\frac{1}{\alpha} \log (\alpha \gamma + e^{-\alpha \varphi}))</td>
<td>(f(\varphi) = -\lambda \log (e^{-\mu/\lambda} + e^{\varphi/\lambda}))</td>
</tr>
<tr>
<td>(\gamma = -\Delta t \theta_0 \sigma_{\text{eff}}^{\theta_1} e^{-\theta_2/(RT)})</td>
<td>(\lambda = h(\Theta) = -e_0 - k_1 \Theta_1 + k_2 \Theta_2 - k_0 \Theta_3 + \Theta_3 \log \Theta_3)</td>
</tr>
<tr>
<td>(\alpha \equiv \theta_3)</td>
<td>(\mu \equiv \Theta_3)</td>
</tr>
</tbody>
</table>
The graphical model

- Each node represents a **random variable**
- The arrows encode the **dependence** structure (conditional pdfs)
- The root nodes represent the **prior** densities
- “plates”: sets of nodes (n experiments with i data points each)

We introduce the variables $\nu$ (**value of $\mu$ for each experiment**)

**Graph hierarchy $\rightarrow$ hierarchical inference in 3 steps**
**Step A: for each exp. infer \( \lambda, \nu \)**

For each exp. \( i \), integrate out the nuisance variables \( \varphi, \phi \) to get the posterior
\[
P(\lambda, \nu \mid \{\varphi^i_{\text{obs}}, \phi^i_{\text{obs}}\}) \propto \prod_i \int \int P(\varphi^i) P(\varphi^i_{\text{obs}} \mid \varphi^i) P(\phi^i) P(\phi^i_{\text{obs}} \mid \phi^i) P(\varphi^i, \phi^i, \lambda, \nu) d\varphi^i d\phi^i
\]

We need to compute this integral: (we use a piecewise linear approx. of \( f \))
\[
I(\lambda, \nu, \varphi^i_{\text{obs}}, \phi^i_{\text{obs}}) = \int_0^{100} G_{\omega}(\varphi^i_{\text{obs}}, \sigma^2) G_f(\varphi, \lambda, \nu)(\phi^i_{\text{obs}}, \sigma^2) d\phi^i
\]

Bayesian inference: find an approx. of the energy
\[
U(\lambda, \nu) = -\log P(\lambda, \nu \mid \{\varphi^i_{\text{obs}}, \phi^i_{\text{obs}}\}) = -\sum_i \log I(\lambda, \nu, \varphi^i_{\text{obs}}, \phi^i_{\text{obs}})
\]
Step A - posterior pdf of $\lambda, \nu$

**Approximation step:**
Find a quadratic approx. of $U(\lambda, \nu)$ around its optimum
⇔ Find a Gaussian approx. of the posterior pdf around its mode

1. Optimize $U$ w.r.t. $\lambda, \nu$ (nested line search)
   we need the first derivatives:
   \[
   \frac{\partial U}{\partial u} = - \sum_i \frac{1}{I(\lambda, \nu, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)} \frac{\partial I(\lambda, \nu, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)}{\partial u} \]

2. Compute the second derivatives at the optimum:
   \[
   \frac{\partial^2 U}{\partial u \partial u'} \approx \sum_i \frac{1}{I(\lambda, \nu, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)^2} \frac{\partial I(\lambda, \nu, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)}{\partial u} \frac{\partial I(\lambda, \nu, \phi_{\text{obs}}^i, \phi_{\text{obs}}^i)}{\partial u'}
   \]

3. Write the Gaussian approx. of the posterior pdf:
   \[
   P(\lambda, \nu \mid \{\phi_{\text{obs}}^i, \phi_{\text{obs}}^i\}) \sim \mathcal{N}(\hat{\lambda}, \hat{\nu} \mid A) \quad \text{where} \quad A_{\lambda\nu} = \frac{\partial^2 U}{\partial u \partial \nu}
   \]

2D Gaussian pdf
Step B: infer $\Theta$ from $\{\lambda^n, \nu^n\}$

Simple marginalization step:

$$P(\Theta \mid \{\lambda^n_{obs}, \nu^n_{obs}\}) = P(\Theta) \prod_n G(h(\Theta), \Theta_3)((\lambda^n_{obs}, \nu^n_{obs}), A)$$

Bayesian inference: find an approx. of the energy

$$U'(\Theta) = -\log P(\Theta \mid \{\lambda^n_{obs}, \nu^n_{obs}\})$$

$$U'(\Theta) = -\log P(\Theta) + \frac{1}{2} \sum_n A_{00}^n (h(\Theta) - \lambda^n_{obs})^2 + A_{11}^n (\Theta_3 - \nu^n_{obs})^2$$

1. Optimize $U'$ w.r.t. $\Theta$
2. Second derivatives at the optimum
Final step: revert from $\Theta$ to $\theta$

→ Gaussian approx. of the posterior pdf of $\Theta$:

$$P(\Theta \mid \{\lambda_{\text{obs}}^n, \nu_{\text{obs}}^n\}) \sim \mathcal{N}(\tilde{\Theta}, \tilde{\Sigma}) \quad \text{where} \quad B_{uv} = \frac{\partial^2 U'}{\partial u \partial v}$$

We need the Gaussian approx. of the posterior pdf of $\theta$:

$P(\theta \mid \{\phi_{\text{obs}}^i, \phi_{\text{obs}}^i \}^n) \sim P(\theta \mid \lambda_{\text{obs}}^n, \nu_{\text{obs}}^n) \sim \mathcal{N}(\tilde{\theta}, \Sigma)$

**Hierarchical inference**

Use the Jacobian $J$ of the transform $F$

$$J_{uv} = \frac{\partial \Theta_i}{\partial \Theta_l} \quad \Theta = F(\theta) = \{\theta_0/\theta_3, \theta_1/\theta_3, \theta_2/\theta_3, 1/\theta_3\}$$

Thus we get the optimal $\theta$ and the corresponding covariance matrix $\Sigma$:

$$\hat{\theta} = F^{-1}(\tilde{\Theta}) \quad \text{and} \quad \Sigma = (J^TBJ)^{-1}$$
Simulated compaction data

- $\phi = f(\varphi)$, with parameters as in *Rutter and Wanten, 2001*
- Gaussian noise added to both $\varphi$ and $\phi$ to simulate measurement errors.

Example of construction of the couples of measurements for the special case of **time series**:
Compaction experiments: 3 series

In each series: 2 or 4 values of T, and 3 values of $\sigma_{\text{eff}}$

<table>
<thead>
<tr>
<th>Exp#</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>380</td>
<td>380</td>
<td>380</td>
<td>420</td>
<td>420</td>
<td>420</td>
</tr>
<tr>
<td>$\sigma_{\text{eff}}$</td>
<td>35</td>
<td>70</td>
<td>105</td>
<td>15</td>
<td>35</td>
<td>70</td>
<td>15</td>
<td>35</td>
<td>70</td>
<td>15</td>
<td>35</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exp#</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>340</td>
<td>340</td>
<td>340</td>
<td>380</td>
<td>380</td>
<td>380</td>
<td>420</td>
<td>420</td>
<td>420</td>
</tr>
<tr>
<td>$\sigma_{\text{eff}}$</td>
<td>35</td>
<td>70</td>
<td>105</td>
<td>35</td>
<td>70</td>
<td>105</td>
<td>15</td>
<td>35</td>
<td>70</td>
<td>15</td>
<td>35</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exp#</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>440</td>
<td>440</td>
<td>440</td>
<td>580</td>
<td>580</td>
<td>580</td>
<td>720</td>
<td>720</td>
<td>720</td>
</tr>
<tr>
<td>$\sigma_{\text{eff}}$</td>
<td>15</td>
<td>100</td>
<td>200</td>
<td>15</td>
<td>100</td>
<td>200</td>
<td>15</td>
<td>100</td>
<td>200</td>
<td>15</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>

T : K
$\sigma_{\text{eff}}$ : MPa
Results: 95% confidence regions for $P(\theta_1, \theta_2 | \text{data})$

- 6 experiments
- 12 experiments, same T and $\sigma_{\text{eff}}$ range, $\theta_1 \theta_2$ correlation = 0.98
- 12 experiments, larger T and $\sigma_{\text{eff}}$ range, $\theta_1 \theta_2$ correlation = 0.52

- More data, less uncertainty
- More experiments, less uncertainty
- Larger T and $\sigma_{\text{eff}}$ range, less correlation
Results: marginals

<table>
<thead>
<tr>
<th></th>
<th>$\theta_0$</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6 exp.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimum</td>
<td>4 $\times$ 10^{-07}</td>
<td>3.0</td>
<td>113000</td>
<td>0.70</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.5 $\times$ 10^{-07}</td>
<td>0.12</td>
<td>1300</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>12 exp.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimum</td>
<td>4 $\times$ 10^{-07}</td>
<td>3.06</td>
<td>114400</td>
<td>0.72</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.2 $\times$ 10^{-07}</td>
<td>0.04</td>
<td>140</td>
<td>0.009</td>
</tr>
<tr>
<td><strong>True values</strong></td>
<td>2.3 $\times$ 10^{-07}</td>
<td>3.13</td>
<td>114500</td>
<td>0.73</td>
</tr>
</tbody>
</table>

- Stress exponent
- Activation energy (kJ/mol)
Conclusion

• We built a general Bayesian framework to analyse laboratory creep compaction experiments;

• To better condition the inverse problem, we chose to a) reparametrize, b) use a hierarchical approach;

• We successfully tested the model on a set of compaction simulations derived from real experiments. Despite the heavy non-linearity, we can retrieve accurate estimates of both the stress exponent and the activation energy.

• To reduce uncertainty: more data, and/or more experiments
  To reduce covariance: larger $T$ and $\sigma_{\text{eff}}$ range

• We can make use of all the available data ≠ common methods use only exp. with same $T$ and varying $\sigma_{\text{eff}}$, or same $\sigma_{\text{eff}}$ and varying $T$, and only the observation points corresponding to the same state of the sample.
Perspectives

- Recent boreholes drilled through active fault zones (e.g., Taiwan, San Andreas) will provide repeated measurements of fault-zone properties and collection of cores for further studies in the lab;

- Other geophysical in situ experiments (e.g., trapped waves) might allow to infer the evolution of fault-zone porosity;

- Our general inference method, because it uses couples of measurements (no need for a time of reference), will just need another hierarchical level to link the measured properties (e.g., pore pressures, wave velocities) to porosity.

- This undrained case, once applied to real lab data, can already be used in forward fault models.
Long-term goal: hazard assessment

- Design **physics-based seismic hazard** assessment tools (*including interseismic compaction*),
- Capable of **integrating lab and field data** and their uncertainties
- Complete with a **measure of the robustness** and quality of the computed earthquake scenarios

State of the art:

- Mostly based on statistics on past seismicity
- Not time-dependent

Probabilities (shown in boxes) of one or more major (M>=6.7) earthquakes on faults in the San Francisco Bay Region during the coming 30 years.

http://quake.wr.usgs.gov