Supplement of The Cryosphere, 15, 1731–1750, 2021
https://doi.org/10.5194/tc-15-1731-2021-supplement

The Cryosphere Open Access EGU

Supplement of

Inferring the basal sliding coefficient field for the Stokes ice sheet model under rheological uncertainty

Olalekan Babaniyi et al.

Correspondence to: Noémi Petra (npetra@ucmerced.edu)

In this supplementary material, we provide additional numerical examples with the goal of further demonstrating the robustness and effectiveness of the proposed approach. Specifically, we consider

1. using a true auxiliary parameter, a_{true}, which is a given function rather than a sample from the associated prior,
2. using different (in terms of mean and variance) prior distributions for the basal sliding coefficient, and
3. the use of a reguralization-type approach, where we tune the parameters of the prior and likelihood, to account for the approximation errors.

For ease of exploration, the numerical studies in this supplemental material are applied to the problem outlined in Example 1a of the manuscript, i.e., the case of an uncertain flow rate factor in the two-dimensional linear Stokes ice sheet model. The true basal sliding coefficient, β_{true}, for all examples is also the same as in Example 1a in the manuscript.

S.1 Studying the effect of the true auxiliary parameter on approximating the posterior

In this case we set the true auxiliary parameter used to generate the data to

$$a_{\text{true}} = \sin(2\pi x_1/L) + \sin(2\pi x_2/H),$$

where $L = 10^4$ m and $H = 250$ m, rather than using a sample from the prior. In Fig. S.1 we show the marginal prior and Laplace approximated posterior distributions, as well as three draws from each of the distributions, the corresponding MAP estimates, and the true basal sliding coefficient. As expected, the accurate MAP estimate, $\boldsymbol{\beta}_{\text{REF}}$, closely resembles the true basal sliding MAP coefficient, and the truth is well supported by the posterior distribution. Second, the Laplace posterior found using the conventional error approach is significantly different (compared to the results in the manuscript) for this choice of the auxiliary parameter, though it is still inaccurate and somewhat infeasible. Third, similar to Example 1a in the manuscript, the true basal sliding coefficient lies well within the bulk of the (Laplace approximated) posterior for the BAE approach. This study shows that choosing the true auxiliary parameter field as a function that is not a sample from the prior distribution leads to the same conclusion as reported in the manuscript (see Example 1a).

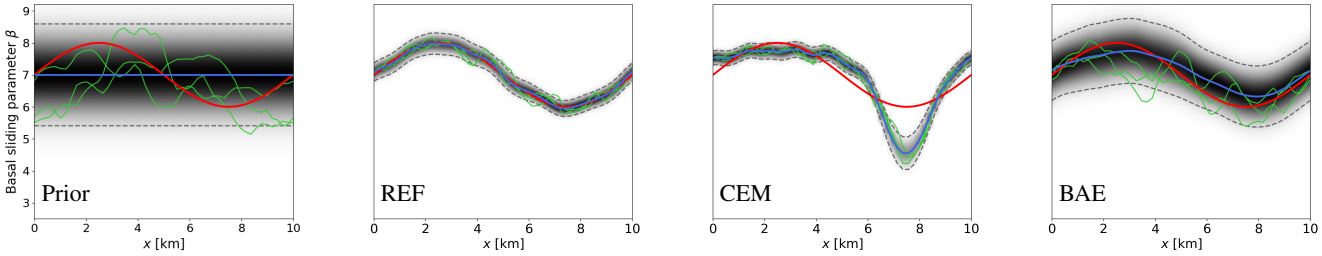

Figure S.1. Prior and MAP estimates of the basal sliding parameter with the true auxiliary parameter (used to generate the data) set to $a_{\text{true}} = \sin(2\pi x_1/L) + \sin(2\pi x_2/H)$. Prior (far left), accurate/reference (REF) case (centre left), conventional error model (CEM) case (center right), and Bayesian approximation error (BAE) case (far right). In each plot, the mean of the distribution (blue line) is shown along with the true basal sliding parameter, β_{true}, (red line), three samples from the respective distributions (green lines), the marginal distribution (shaded) with darker shading indicating higher probability, and the ± 2 (approximate) standard deviation intervals (black dashed line).

S.2 Studying the effect of the prior on approximating the posterior

Here we consider changing the prior mean β_* and variance on β. Specifically, we consider three distinct cases: a) $\beta_* = 6$, b) $\beta_* = 5$, and c) $\beta_* = 5$ and the variance is increased by a factor of 4. In Fig. S.2 we show the marginal prior and Laplace

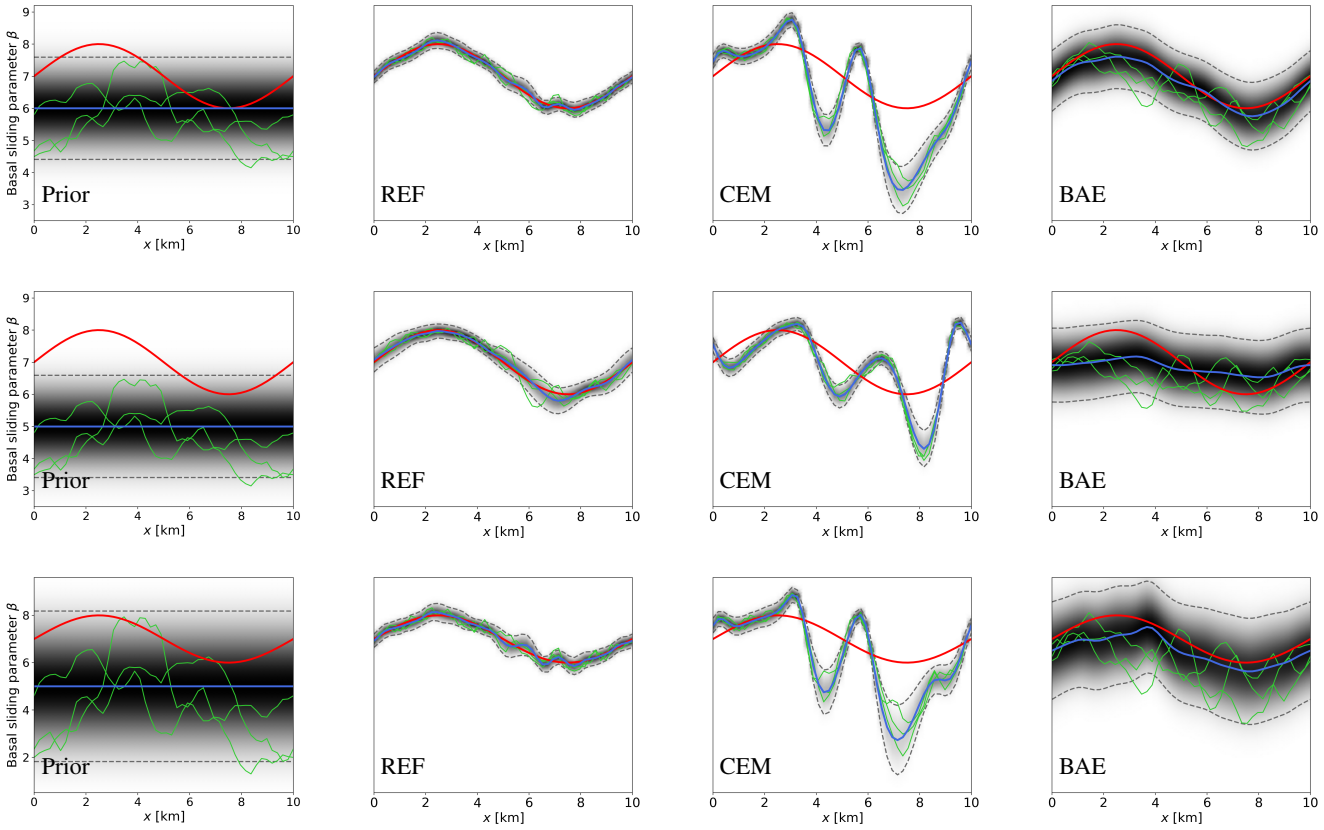

Figure S.2. Prior and MAP estimates of the basal sliding parameter with the use of different priors on $\boldsymbol{\beta}$. Top row: The prior mean is changed to $\beta_* = 6$, with the prior (far left), accurate/reference (REF) case (center left), conventional error model (CEM) case (center right), and Bayesian approximation error (BAE) case (far right). Middle row shows the same plots for changing the prior mean to $\beta_* = 5$, while the bottom row shows the same plots for changing the prior mean to $\beta_* = 5$ and increasing the variance by a factor of 4. In each plot, the mean of the distribution (blue line) is shown along with the true basal sliding parameter, β_{true}, (red line), three samples from the respective distributions (green lines), the marginal distribution (shaded) with darker shading indicating higher probability, and the ± 2 (approximate) standard deviation intervals (black dashed line).

approximated posterior distributions, as well as three draws from each of the distributions, the corresponding MAP estimates, and the true basal sliding coefficient for each case a)–c). The results of this study show the same conclusion as the one reported in the manuscript, though as expected, a more variable prior will generally lead to larger uncertainty (shown on the far right column). In addition, the results show that when the truth is less well supported by the prior, the MAP estimates deteriorate.

S.3 Studying the effect of scaling the prior and likelihood on approximating the posterior

The goal of this section is to study the effect of scaling the prior and likelihood to account for the approximation errors. This study can be seen as varying the regularization parameter (or scaling the variance of the noise) in a deterministic setting. This approach can also be understood as *tempering* the prior or the likelihood in the statistical setting.

We present results for two sets of experiments. In the first set we scale the prior using the (finite-dimensional representation of the) prior in the form $\pi_{\text{prior}}(\boldsymbol{\beta})^{\alpha}$ with $\alpha > 0$, so that the negative log-posterior can be written as

$$-\ln\left(\pi_{\text{post}}(\boldsymbol{\beta}|\boldsymbol{d})\right) \propto -\ln\left(\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{\beta})\right) - \alpha \ln\left(\pi_{\text{prior}}(\boldsymbol{\beta})\right) = \frac{1}{2}\left\|\mathcal{F}(\boldsymbol{\beta}) - \boldsymbol{d}\right\|^2_{\boldsymbol{\Gamma}_{\text{e}}^{-1}} + \frac{\alpha}{2}\left\|\boldsymbol{\beta} - \boldsymbol{\beta}\right\|^2_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}. \tag{1}$$

In the second set of experiments, in which we scale the likelihood, we define the likelihood as $\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{\beta})^{\frac{1}{\alpha}}$, with $\alpha > 0$, so that the negative log-posterior can be written as

$$-\ln\left(\pi_{\text{post}}(\boldsymbol{\beta}|\boldsymbol{d})\right) \propto -\frac{1}{\alpha}\ln\left(\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{\beta})\right) - \ln\left(\pi_{\text{prior}}(\boldsymbol{\beta})\right) = \frac{1}{2\alpha}\left\|\mathcal{F}(\boldsymbol{\beta}) - \boldsymbol{d}\right\|^2_{\boldsymbol{\Gamma}_{\text{e}}^{-1}} + \frac{1}{2}\left\|\boldsymbol{\beta} - \boldsymbol{\beta}\right\|^2_{\boldsymbol{\Gamma}_{\text{pr}}^{-1}}. \tag{2}$$

We note that Equations (1) and (2) are differing only by a factor of α, therefore the MAP estimates for each α for the two experiments are the same. The associated posterior uncertainty on the other hand will in general be different. In addition, when the distribution of the approximation errors is well approximated by $\mathcal{N}(\mathbf{0}, \delta_{\varepsilon}^2\boldsymbol{I})$, i.e., $\boldsymbol{\varepsilon}_* \approx \mathbf{0}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\varepsilon}} \approx \delta_{\varepsilon}^2\boldsymbol{I}$ (i.e., the covariance is a scaled identity), as is the case in Example 1a of the manuscript (see Fig. 6 of the manuscript), tempering the likelihood with $\alpha = \left(1 + \frac{\delta_{\varepsilon}^2}{\delta_e^2}\right)$, where we recall that δ_e^2 denotes the variance of the additive noise in the measurements, the CEM approach can perform similar to the BAE approach. However, in general this is not the case, namely the approximation errors are often highly correlated, see for example Nicholson et al. (2018); Tarvainen et al. (2009); Nissinen et al. (2007); Voutilainen and Kaipio (2009). Furthermore, δ_{ε}^2 is not known a priori and therefore for the CEM approach this needs to be estimated for instance by the L-curve criterion (Vogel, 2002), which requires solving the inverse problem for several α values. The challenge with this approach is that it can be computationally intensive and more importantly, it can fail to provide feasible results even in the case when the covariance of the approximation error is close to diagonal, as we show below.

In Figures S.3 and S.4 we show the marginal Laplace approximated posterior distributions, as well as three draws from each of the distributions, the corresponding MAP estimates, and the true basal sliding coefficient for tempering the prior and likelihood, respectively, with $\alpha = 10^{-2}$, $\alpha = 10^{-1}$, $\alpha = 10^{0}$ (as in the manuscript), $\alpha = 10^{1}$ $\alpha = 10^{2}$, $\alpha = 10^{3}$, and $\alpha = 10^{4}$. For comparison, for both cases, we also show the associated results found using the BAE approach. As expected, for the case of tempering the prior, we see that larger values of α lead to smaller posterior uncertainty, and the estimates approach the prior mean. The case of $\alpha = 10^{3}$ is especially noteworthy, as the MAP estimate appears to be in good agreement with the true basal sliding coefficient. However, the posterior uncertainty is substantially underestimated, with the truth lying more than six posterior standard deviations from the MAP estimate, over most of the domain. That is to say, the resulting posterior is drastically overconfident, and essentially not feasible.

When tempering the likelihood we see that for $\alpha = 10^{2}$ and $\alpha = 10^{3}$, the resulting MAP estimates are in good agreement with the truth, and the (approximate) posterior densities are clearly feasible. However, as stated above, it is not possible to choose α a priori, and obviously the truth is not available. As such, a heuristic approach such as the L-curve criterion would typically be used to find a suitable value for α (usually the point on the curve which maximizes the curvature). To this end, in Fig. S.5 we show the resulting L-curve. Interestingly $\alpha = 10^{0}$ (i.e., the CEM approach in the manuscript) corresponds to the value for which the curvature in the L-curve is largest. As can be seen, however, the resulting (approximate) posterior density is not feasible, with the truth lying far from the bulk of the posterior support, and the MAP estimate being a poor estimate of the truth.

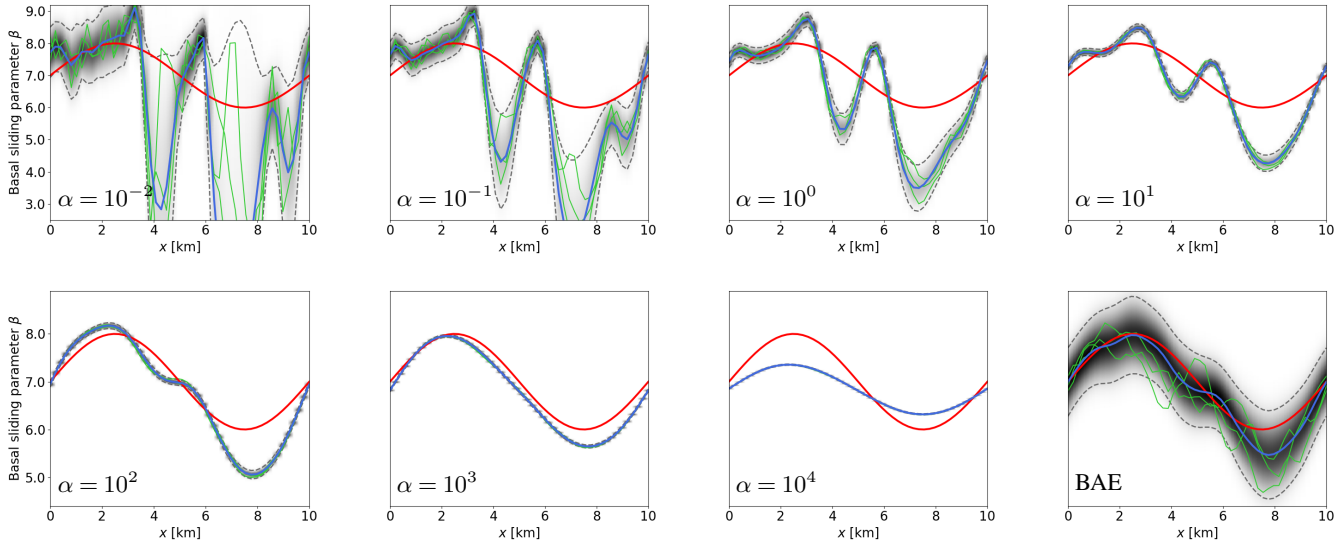

Figure S.3. MAP estimates and approximate posteriors of the basal sliding parameter with tempering of the prior (the corresponding value of α are shown in each plot). For comparison the results using the BAE approach are also shown (bottom far right). In each plot, the mean of the distribution (blue line) is shown along with the true basal sliding parameter, β_{true}, (red line), three samples from the respective distributions (green lines), the marginal distribution (shaded) with darker shading indicating higher probability, and the ± 2 (approximate) standard deviation intervals (black dashed line).

References

Nicholson, R., Petra, N., and Kaipio, J. P.: Estimation of the Robin coefficient field in a Poisson problem with uncertain conductivity field, Inverse Problems, 34, 115 005, 2018.

Nissinen, A., Heikkinen, L., and Kaipio, J.: The Bayesian approximation error approach for electrical impedance tomography—experimental results, Measurement Science and Technology, 19, 015 501, 2007.

Tarvainen, T., Kolehmainen, V., Pulkkinen, A., Vauhkonen, M., Schweiger, M., Arridge, S., and Kaipio, J.: An approximation error approach for compensating for modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography, Inverse Problems, 26, 015 005, 2009.

Vogel, C. R.: Computational methods for inverse problems, SIAM, 2002.

Voutilainen, A. and Kaipio, J. P.: Model reduction and pollution source identification from remote sensing data, Inverse Problems & Imaging, 3, 711, 2009.

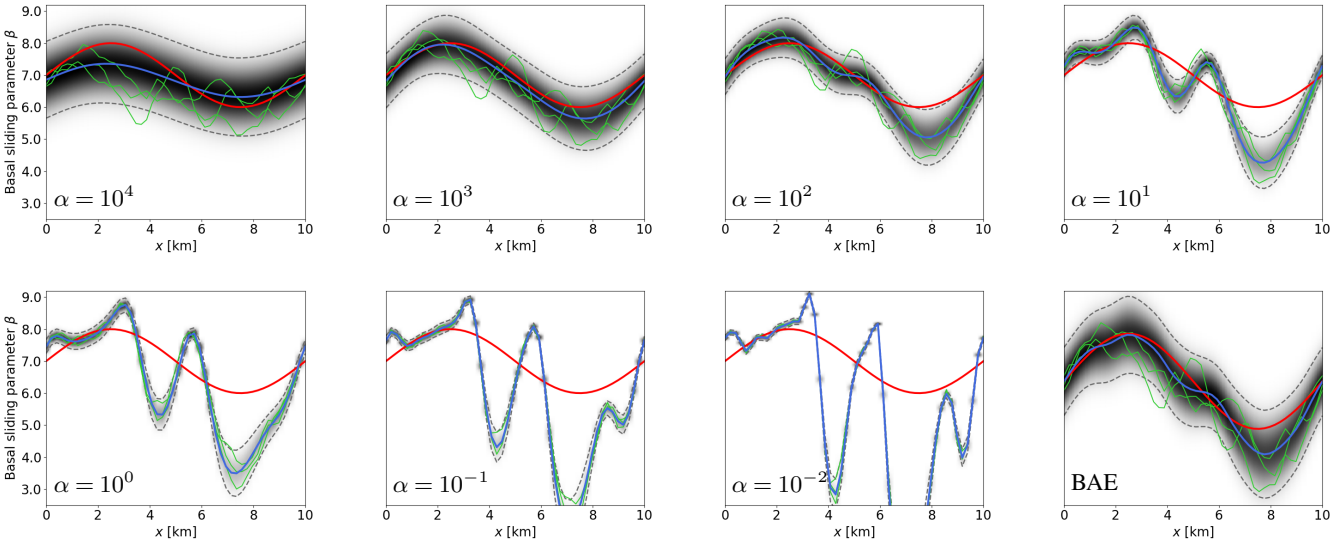

Figure S.4. MAP estimates and approximate posteriors of the basal sliding parameter for Example 1a with tempering of the likelihood (the corresponding value of α are shown in each plot). For comparison the results using the BAE approach are also shown (bottom far right). In each plot, the mean of the distribution (blue line) is shown along with the true basal sliding parameter, β_{true}, (red line), three samples from the respective distributions (green lines), the marginal distribution (shaded) with darker shading indicating higher probability, and the ± 2 (approximate) standard deviation intervals (black dashed line).

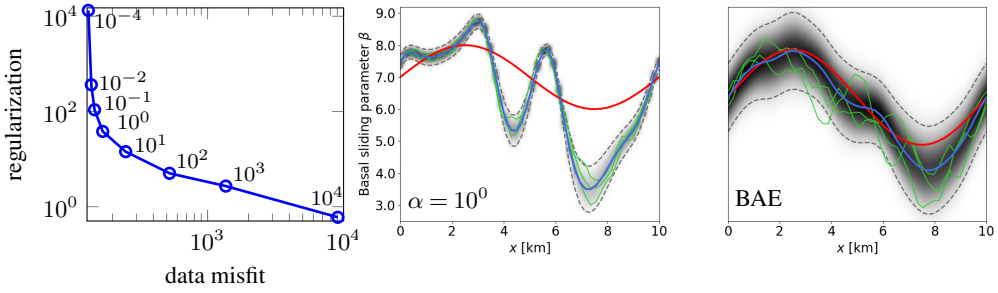

Figure S.5. The L-curve corresponding to tempering the likelihood (far left) and the MAP estimates and approximate posteriors of the basal sliding parameter for tempering of the likelihood with $\alpha = 10^0$ (the points on the L-curve with greatest curvature) (center) along with the results found using the BAE approach (right). In the left and right plots, the mean of the distribution (blue line) is shown along with the true basal sliding parameter, β_{true}, (red line), three samples from the respective distributions (green lines), the marginal distribution (shaded) with darker shading indicating higher probability, and the ± 2 (approximate) standard deviation intervals (black dashed line).