

Bayesian regression quantifies uncertainty of binding parameters from isothermal titration calorimetry more accurately than error propagation

Van N. T. La and David D. L. Minh

E-mail:

Abstract

We compare several different methods to quantify the uncertainty of binding parameters estimated from isothermal titration calorimetry data: the asymptotic standard error from maximum likelihood estimation, error propagation based on a first-order Taylor series expansion, and the Bayesian credible interval. When the methods are applied to simulated experiments and to measurements of Mg(II) binding to EDTA, the asymptotic standard error underestimates the uncertainty in the free energy and enthalpy of binding. Error propagation overestimates the uncertainty for both quantities, except in the simulations, where it underestimates the uncertainty of enthalpy for confidence intervals less than 70%. In both data sets, Bayesian credible intervals are much closer to observed confidence intervals.

Introduction

Isothermal titration calorimetry (ITC) is widely used to characterize binding processes involving biomolecules including proteins,¹ small organic molecules,² DNA/RNA,^{3,4} and lipids.⁵ ITC data is routinely analyzed to estimate thermodynamic parameters - the Gibbs free energy ΔG and the enthalpy ΔH - for simple binding processes. Based on the relation $\Delta G = \Delta H - T\Delta S$ that includes the temperature T , the entropy ΔS may also be obtained. These parameters have often been estimated using what we will refer to as the standard procedure: a nonlinear least squares regression method implemented in the Origin software package that is distributed with the MicroCal VP-ITC instrument and its successors. The software yields a maximum likelihood estimate (MLE) of the parameters and asymptotic standard error (ASE). Unfortunately, the ASE underestimates the uncertainty by as much as an order of magnitude⁶!

The severe underestimation of uncertainty is mainly a consequence of ignoring the error in the titrant concentration. In the standard procedure, the error in the titrand concentration (in the sample cell) is handled by assigning the stoichiometry n as a free parameter. On the other hand, the titrant concentration (in the syringe) is treated as a constant. This assumption is made because ITC data can only be used to estimate *the ratio* of the titrant:titrand concentrations as opposed to individual values.⁷ However, it is a poor assumption because large errors (10–20%) in titrant concentration between laboratories have been observed.⁶

In 2015, Boyce *et. al.* suggested that the uncertainty of the standard procedure could be adjusted based on error propagation.⁸ Specifically, based on the Taylor expansion, errors in K_a , ΔH , and the site parameter n may be corrected by the relative error of the titrant concentration. However, they did not show that the resulting uncertainty estimate leads to accurate confidence intervals. Confidence intervals are accurate when the X% confidence interval includes the true value X% of the time, where X is a confidence level. In 2018, Nguyen *et. al.* described the analysis of ITC data with Bayesian regression. They found that *Bayesian credible intervals* (BCIs) - regions which contain a specified percentage of the

Bayesian posterior - are more accurate *confidence intervals* than those based on the ASE.⁹ There was no comparison to confidence intervals based on the ASE augmented with error propagation. The purpose of this short manuscript is to address this oversight.

Materials and Methods

Integrated heat data

Our data are integrated heats, $\mathcal{D} \in \{q_1, q_2, \dots, q_N\}$, where q_n is the integrated heat of injection n . We analyzed simulations as well as ITC experiments that were previously described.⁹

Simulations are useful because it is inexpensive to collect large amounts of data and because thermodynamic parameters are known exactly. Simulations of simple 1:1 binding were performed in a similar way as in Nguyen *et. al.*⁹ A total of 1000 integrated heat curves with 24 injections each were modeled depending on the free energy of binding ΔG , the enthalpy of binding ΔH , the enthalpy of dilution and stirring per injection ΔH_0 , the concentration of receptor (titrand) $[R]_0$, the concentration of ligand (titrant) $[L]_s$, and the standard deviation of the measurement error σ . The thermodynamic parameters and enthalpy of injection were fixed at $\Delta G = -10$ kcal/mol, $\Delta H = -5$ kcal/mol, and $\Delta H_0 = 0.5$ μ cal. $[R]_0$ and $[L]_s$ were sampled from lognormal distributions with mean values of 0.1 and 1.0 mM, respectively. Based on the uncertainty of 10% observed by Myszkowski *et. al.*,⁶ the variance was set at either small (5% of the mean) or large (10% of the mean). Measurement error was modeled as normally distributed with a zero mean and standard deviation of $\sigma = 1$ μ cal.

We also analyzed 14 integrated heat curves from previously performed experiments in which MgCl_2 was titrated into a sample cell containing EDTA in a MicroCal VP-ITC calorimeter.⁹

Regression

Data were analyzed via Bayesian regression and maximum likelihood estimation. In both procedures, integrated heat curves for simple 1:1 binding were modeled as previously described.⁹ They are functions of the aforementioned parameters,

$$\boldsymbol{\theta} \equiv (\Delta G, \Delta H, \Delta H_0, [R]_0, [L]_s, \sigma). \quad (1)$$

Observed injection heats q_n were treated as normally distributed about the true heats $q_n^*(\boldsymbol{\theta})$,

$$q_n \sim \mathcal{N}(q_n^*(\boldsymbol{\theta}), \sigma^2). \quad (2)$$

Thus, the likelihood function of an integrated heat curve $\mathcal{D} \in \{q_1, q_2, \dots, q_N\}$ is,

$$p(\mathcal{D}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\sigma^N} \exp \left[-\frac{1}{2\sigma^2} \sum_{n=1}^N (q_n - q_n^*(\boldsymbol{\theta}))^2 \right]. \quad (3)$$

Bayesian regression

For the Bayesian regression, the prior of parameters was independent, such that $p(\boldsymbol{\theta}) = \prod_i p(\theta_i)$. As in Nguyen *et. al.*,⁹ uniform priors were used for ΔG , ΔH , ΔH_0 . Lognormal priors were used for the concentrations of the ligand and the receptor,

$$\ln[\mathcal{X}]_0 \sim \mathcal{LN}([\mathcal{X}]_0, (\delta[\mathcal{X}]_0)^2), \quad (4)$$

where $[\mathcal{X}]_0 \in \{[R]_0, [L]_s\}$ is the stated value of each quantity. δ was assumed to be either 5% or 10%. The uninformative Jeffreys prior was used for σ :¹⁰

$$p(\sigma) \propto \frac{\sigma_0}{\sigma}, \quad (5)$$

where $\sigma_0 = 1$ cal. For this model, the posterior probability density is,

$$p(\boldsymbol{\theta}|\mathcal{D}) \propto p(\mathcal{D}|\boldsymbol{\theta}) p(\boldsymbol{\theta}), \quad (6)$$

Sampling from the Bayesian posterior was performed using a Markov Chain Monte Carlo method, as in Nguyen *et. al.*⁹ but with a few small adjustments. Instead of using pymc3, the regression was implemented in numpyro.^{11,12} After 2000 warm-up moves, 10000 (opposed to 5000⁹) samples from 4 chains were stored. The X% BCI of each parameter was calculated based on the smallest interval that contains X% of the posterior samples. Additionally, the uncertainty δ was set at either 5% and 10%, opposed to only 10%.

Maximum likelihood estimation

For the MLE, parameter estimates were obtained as,

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \log p(\mathcal{D}|\boldsymbol{\theta}). \quad (7)$$

The covariance matrix of the asymptotic standard error (ASE) was estimated based on the inverse Fisher information matrix,

$$\text{cov}(\hat{\boldsymbol{\theta}}) \approx -\frac{1}{N} \left[\frac{\partial \log \mathcal{L}_{\mathcal{N}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\top}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \right]^{-1}. \quad (8)$$

We used `scipy.optimize.minimize` function from the python package `scipy`¹³ to implement this MLE model, estimate the parameters, and automatically calculate the covariance matrix for parameter uncertainty. The X% CI of each parameter was defined by an interval in which the lower bound was the $1 - X/2$ percentile, and the upper bound was the $1 + X/2$ percentile of the normal distribution with a mean as the estimated value and standard deviation as the ASE.

Maximum likelihood estimation with error propagation

We performed error propagation to augment the ASE of MLE parameters based on the formula provided by Boyce *et. al.*,⁸

$$\left(\frac{s_\theta}{\theta}\right)^2 = \left(\frac{s_{\theta,ASE}}{\theta}\right)^2 + \left(\frac{s_{[L]_s}}{[L]_s}\right)^2 \quad (9)$$

In this equation, s are standard errors and $\theta \in \{\Delta G, \Delta H\}$ are the parameters affected by the uncertainty of ligand concentration $[L]_s$. $s_{\theta,ASE}$ is the ASE, $s_{[L]_s}$ is the standard error in the ligand concentration, and s_θ is the error estimate of the parameter θ that incorporates ligand concentration error. The uncertainty of the ligand concentration $s_{[L]_s}$ can be estimated by another experiment or based on previous estimates. Considering uncertainty in both protein and ligand concentrations, we used either 5% or 10% for both $s_{[R]_0}/[R]_0$ and $s_{[L]_s}/[L]_s$. CIs of this procedure were estimated similarly to the MLE procedure.

Results

Bayesian posteriors are converged

As in Nguyen *et. al.*,⁹ the Markov chain Monte Carlo protocol leads to converged BCIs. In a representative run for one of the 1000 simulated integrated heat curves (Fig. S1), fewer than 10% of samples are required before estimated percentiles of the posterior density are stable (Fig. S2). Comparable convergence behavior is observed when sampling the Bayesian posterior for the ITC experiments (Fig. S3).

Error propagation expands confidence intervals to be larger than Bayesian credible intervals

For each of the 14 experiments, 95% CIs of ΔG and ΔH are shown in Fig. 1. Because the true value of parameters is unknown, the median estimate is shown as a proxy. Panels (a) and (b) reproduce Fig. 6 from Nguyen *et. al.*⁹ The 95% CIs of ΔG encompass the median value in nearly every experiment. For CIs based on the ASE, 95% CIs of ΔH are too small. Panel (c) shows that error propagation increases CIs to encompass the median, but the CIs appear to be larger than necessary.

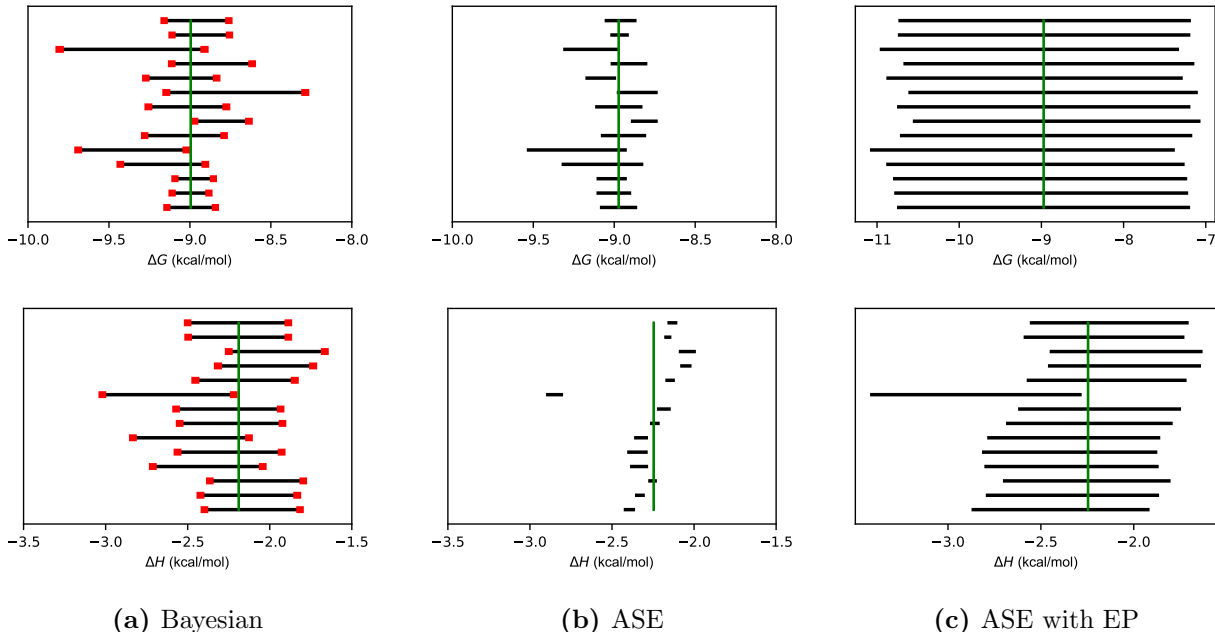


Figure 1: Uncertainty estimates of Mg(II):EDTA dataset. 95% credible intervals estimated from the Bayesian posterior (left), confidence intervals calculated by ASE from nonlinear least squares (middle), and confidence intervals calculated by ASE with EP (right) for parameters specifying magnesium binding to EDTA. The median MCMC samples are shown by the vertical green lines. The standard deviations of the lower and upper bounds are denoted as red bars and estimated by bootstrapping.

Even with error propagation, BCIs provide more accurate CIs than the ASE

The accuracy of CIs was more carefully assessed by coverage plots, in which predicted confidence intervals are plotted against the percentage of BCIs and CIs that contain the true values of ΔG and ΔH . For accurate CIs, points should lie along the diagonal. If points are below the diagonal, CIs are underestimated. If points are above the diagonal, CIs are overestimated.

Coverage plots were generated for 1000 simulations with high error (Fig. 2) and low error (Fig. S4) and for 14 experiments (Fig. 3). In all of the coverage plots, Bayesian credible intervals are closest to the diagonal. As expected, the ASE consistently underestimates confidence intervals. For ΔG , the ASE with error propagation overestimates confidence intervals for nearly all confidence levels. For ΔH , the story is more subtle. In the simulations at both high and low error, confidence intervals are underestimated for confidence levels less than 70% but somewhat overestimated for higher confidence levels. In the experiments, confidence intervals are overestimated for confidence intervals greater than 30%.

Conclusions

In both ITC simulations and experiments, BCIs provide more accurate uncertainty estimates for thermodynamic binding parameters than the ASE, without or with error propagation. The ASE underestimates the uncertainties of all datasets. Error propagation overestimates the uncertainties in the experimental dataset, but in simulations it underestimates the uncertainty of enthalpy for confidence intervals less than 70%.

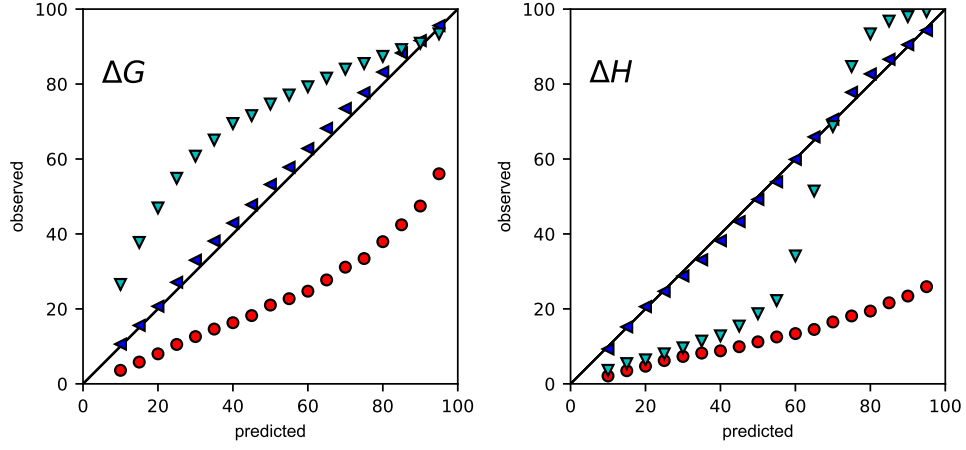


Figure 2: Uncertainty validation of the simulation dataset at high error 10%. The predicted rate (%) of CIs containing the true values were plotted against the observed rate (%) for Bayesian credible intervals (blue leftward triangles), nonlinear least squares confidence intervals (red circles), and nonlinear least squares confidence intervals with error propagation (cyan downward triangles). Error bars of Bayesian procedure, which were standard deviations based on 100 bootstrapping samples, were too small to be visible.

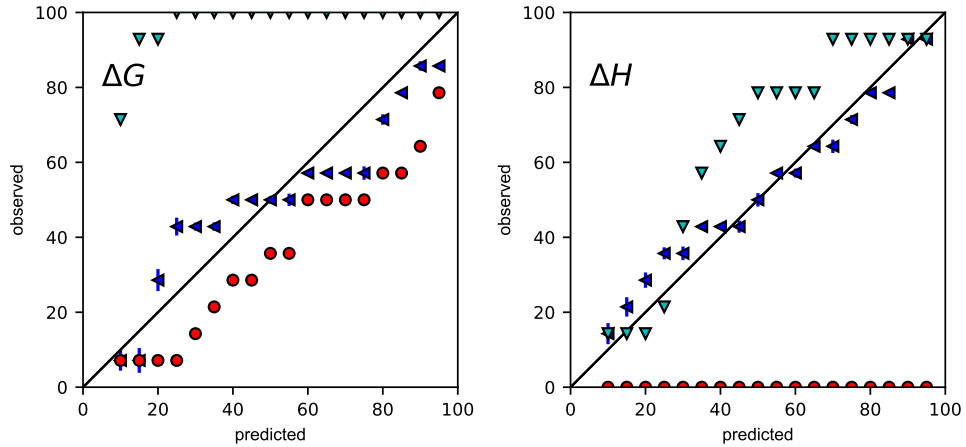


Figure 3: Uncertainty validation of Mg(II)-EDTA dataset. The predicted rate (%) of CIs containing the true values were plotted against the observed rate (%) for Bayesian credible intervals (blue leftward triangles), nonlinear least squares confidence intervals (red circles), and nonlinear least squares confidence intervals with error propagation (cyan downward triangles). Error bars of Bayesian procedure were standard deviations based on 100 bootstrapping samples.

Acknowledgment

We thank Joel Tellinghuisen for suggesting this comparison.

This work was supported by an award entitled “Collaborative Research: CDS&E: Elucidating Binding using Bayesian Inference to Integrate Multiple Data Sources” (#1905324) from the Chemical Measurement and Imaging Program in the Division of Chemistry of the National Science Foundation (<https://www.nsf.gov/>), for which DDLM is the principal investigator.

Supplementary

S1 Figure. 1000 simulated curves of 1:1 binding ITC data. Parameters for the curves are in the Materials and Methods section of the main text.

S2 Figure. Convergence analysis of Bayesian sampling for a representative simulated dataset 10,000 samples were generated from the Bayesian posterior. All six parameters are shown. The 5-th, 25-th, 50-th, 75-th, and 95-th percentiles are represented as lines with blue circle, green square, red diamond, cyan upward triangle, and magenta downward triangle, respectively. The error bars estimated by 100 bootstrapping samples are too small to be visible.

S3 Figure. Convergence analysis of the Bayesian posterior for Mg(II)-EDTA datasets 10,000 samples were generated from the Bayesian posterior. All six parameters are shown. The 5-th, 25-th, 50-th, 75-th, and 95-th percentiles are represented as lines with blue circle, green square, red diamond, cyan upward triangle, and magenta downward triangle, respectively. The error bars estimated by 100 bootstrapping samples are too small to be visible.

S4 Figure. Uncertainty validation of the simulation dataset at low error 5%. The predicted rate (%) of CIs containing the true values were plotted against the observed rate (%) for Bayesian credible intervals (blue leftward triangles), nonlinear least squares confidence intervals (red circles), and nonlinear least squares confidence intervals with error propagation (cyan downward triangles). Error bars of Bayesian procedure, which were standard deviations based on 100 bootstrapping samples, were too small to be visible.

References

- (1) Leavitt, S.; Freire, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current Opinion in Structural Biology* **2001**, *11*, 560–566. [https://doi.org/10.1016/s0959-440x\(00\)00248-7](https://doi.org/10.1016/s0959-440x(00)00248-7).
- (2) Duff, M.R.; Grubbs, J.; Howell, E.E. Isothermal Titration Calorimetry for Measuring Macromolecule-Ligand Affinity. *Journal of Visualized Experiments* **2011**, *55*, e2796. <https://doi.org/10.3791/2796>.
- (3) Feig, A.L. Chapter 19 - Studying RNA–RNA and RNA–Protein Interactions by Isothermal Titration Calorimetry. In *Methods in Enzymology*; Academic Press, 2009; Vol. 468, *Biophysical, Chemical, and Functional Probes of RNA Structure, Interactions and Folding: Part A*, pp. 409–422. [https://doi.org/10.1016/S0076-6879\(09\)68019-8](https://doi.org/10.1016/S0076-6879(09)68019-8).
- (4) Malecek, K.; Ruthenburg, A. Chapter Nine - Validation of Histone-Binding Partners by Peptide Pull-Downs and Isothermal Titration Calorimetry. In *Methods in Enzymology*; Wu, C.; Allis, C.D., Eds.; Academic Press, 2012; Vol. 512, *Nucleosomes, Histones & Chromatin Part A*, pp. 187–220. <https://doi.org/10.1016/B978-0-12-391940-3.00009-3>.
- (5) Swamy, M.J.; Sankhala, R.S.; Singh, B.P. Thermodynamic Analysis of Protein-Lipid Interactions by Isothermal Titration Calorimetry. *Methods in Molecular Biology (Clifton, N.J.)* **2019**, *2003*, 71–89. https://doi.org/10.1007/978-1-4939-9512-7_4.
- (6) Myszka, D.G.; Abdiche, Y.N.; Arisaka, F.; Byron, O.; Eisenstein, E.; Hensley, P.; Thomson, J.A.; Lombardo, C.R.; Schwarz, F.; Stafford, W.; et al. The ABRF-MIRG'02 Study: Assembly State, Thermodynamic, and Kinetic Analysis of an Enzyme/Inhibitor Interaction. *Journal of Biomolecular Techniques* **2003**, *14*, 247–269.

- (7) Estelle, A.B.; George, A.; Barbar, E.J.; Zuckerman, D.M. Quantifying cooperative multisite binding in the hub protein LC8 through Bayesian inference. *PLOS Computational Biology* **2023**, *19*, e1011059. <https://doi.org/10.1371/journal.pcbi.1011059>.
- (8) Boyce, S.E.; Tellinghuisen, J.; Chodera, J.D. Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry. preprint, Biochemistry, 2015. <https://doi.org/10.1101/023796>.
- (9) Nguyen, T.H.; Rustenburg, A.S.; Krimmer, S.G.; Zhang, H.; Clark, J.D.; Novick, P.A.; Branson, K.; Pande, V.S.; Chodera, J.D.; Minh, D.D.L. Bayesian analysis of isothermal titration calorimetry for binding thermodynamics. *PLOS ONE* **2018**, *13*, e0203224. <https://doi.org/10.1371/journal.pone.0203224>.
- (10) Jeffreys, H. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1946**, *186*, 453–461. <https://doi.org/10.1098/rspa.1946.0056>.
- (11) Bingham, E.; Chen, J.P.; Jankowiak, M.; Obermeyer, F.; Pradhan, N.; Karaletsos, T.; Singh, R.; Szerlip, P.; Horsfall, P.; Goodman, N.D. Pyro: deep universal probabilistic programming. *The Journal of Machine Learning Research* **2019**, *20*, 973–978. <https://doi.org/10.48550/arXiv.1810.09538>.
- (12) Phan, D.; Pradhan, N.; Jankowiak, M. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro, 2019. <https://doi.org/10.48550/arXiv.1912.11554>.
- (13) Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **2020**, *17*, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.