# 乱数シミュレーションによるフィッティングモデル評価

# Applicability of Random Simulation to the Evaluation of Interpolation

Yuji Sode (IEP at North Carolina State Univ.)
Email: ysode@ncsu.edu
Email: yuji.sode@gmail.com

**Quality and quantity of paleontological samples vary depending on materials to investigate (e.g. Wani, 2003; Hammer and Harper, 2006; Mustoe and Smith, 2023). Sample data is occasionally accompanied with discreteness while continuity is a prerequisite for the interpolation in paleontological studies. Therefore, an approach that has tolerance of both discreteness and continuity is required to evaluate the sample-model interpolation.**

## Materials and Method

**I.** Error model $Q$ was expressed using a random variable $u$

$Q(p) = 1+p(2u-1)$, where $u = [0, 1]$. ... (1)

A model with deviation $F$ was simulated as a product of predetermined function $f$ and $Q$

$F(x) = f(x)Q(p)$. ... (2)

**II.** A sample data model was expressed by 20 percent of deviation

$F(x) = f(x)Q(0.2)$, where $f(x) = (3x-2)^2$. ... (3)

A sample data set was generated as 100 sets of discrete samples through the sample data model.

**III.** Four fitting models $f(x)^n$ against the generated sample data set were estimated using the least squares method with gnuplot version 5.2 (Williams, Kelley, *et al*. 1986-1993, 1998, 2004, 2007-2019)

$f(x)^n = (Ax+B)^n$, ... (4)

where $n = 1, 2, 3$, and $-1$.

**IV.** A value $dL/dx$ that could distinguish the given fitting models was selected for the random simulation. This value for the evaluation was defined based on the partial curve length with length $L$ between two points

$dL = \Sigma L - \max(L)$ ... (5)

and the horizontal range

$dx = \max(x) - \min(x)$. ... (6)

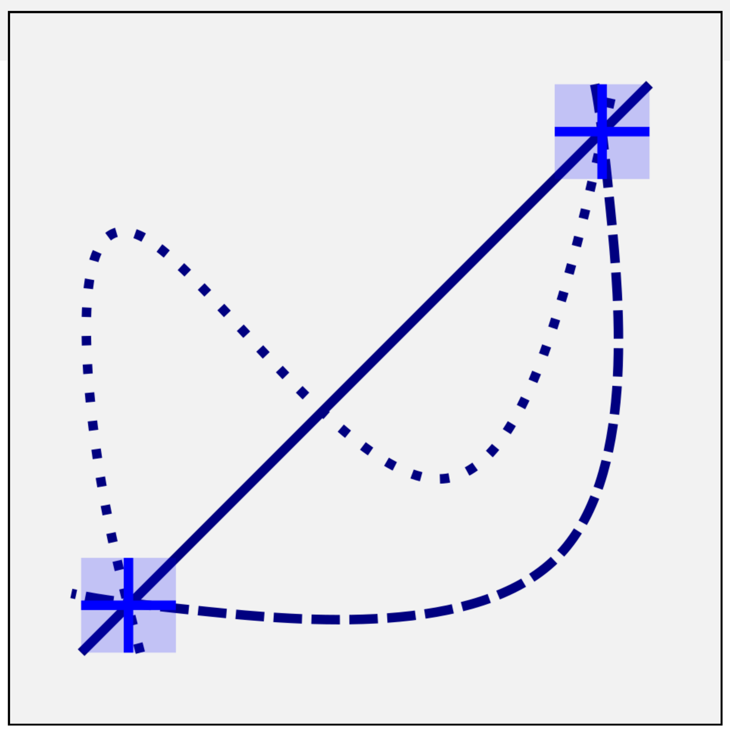

**Figure 1. Schematic diagram showing three interpolations based on different models.**

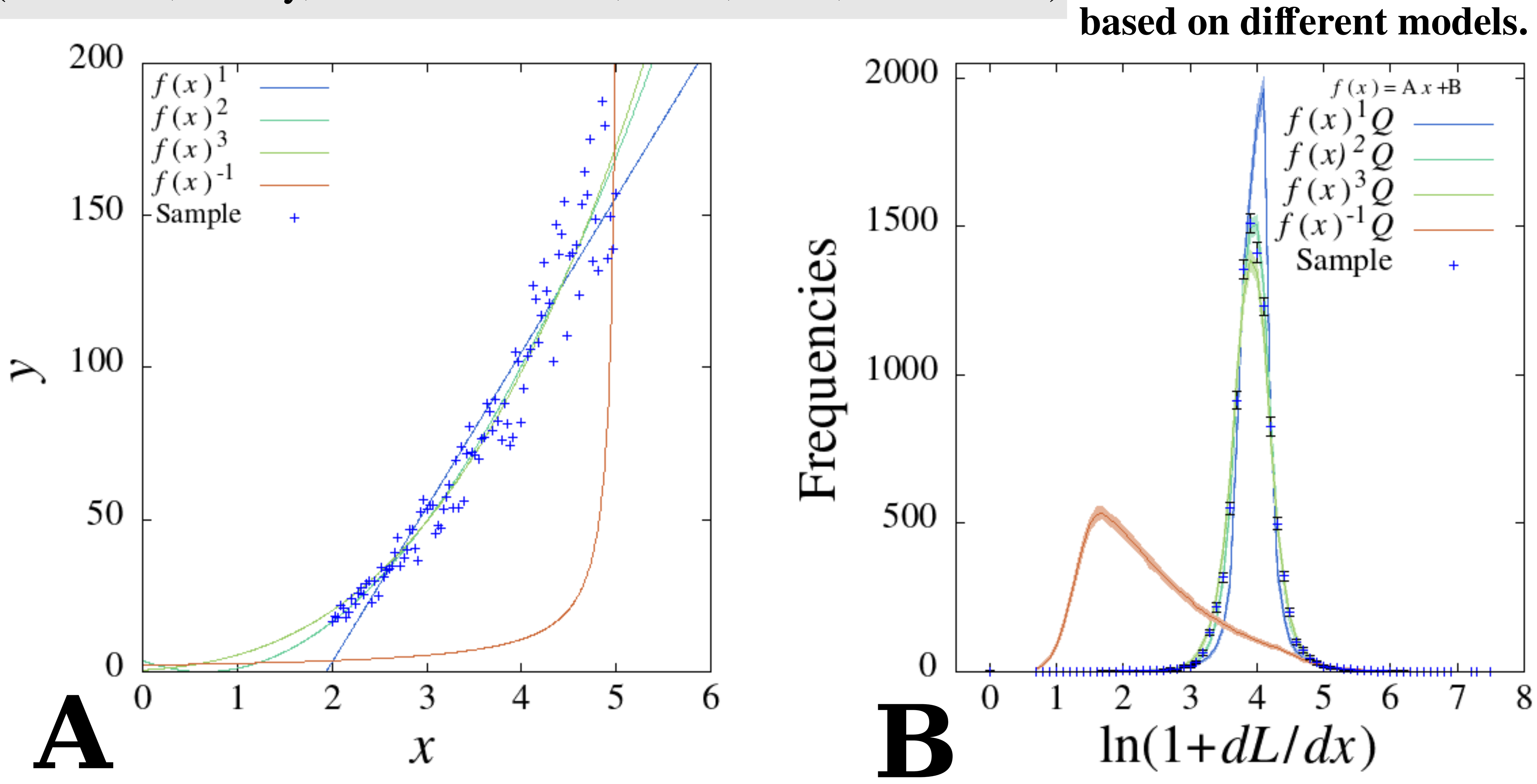


**Figure 2. Graphs showing sample data and result of evaluation.**

**A. Graphs showing sample data and results of curve fitting.**

Every point is plot of the generated data set from a model, which is $(3x-2)^2$ with 20 percent of deviation. Every line is result of curve fitting using the least squares method.

**B. Graphs showing frequency distributions of ln (1+*dL*/*dx*).**

This graph shows frequency distributions of $\ln(1+dL/dx)$, which are defined using 3 random points on a given function $f(x)$ or generated sample data. Every line is result of given function $f(x)$. Every point is result of generated sample data. Each graph represents average of a 100 sets of results from random simulation. Filled areas or error bars represent standard deviations.

**Table 1. Results of the least squares method.**

| Fitting models | A | B | Variance of residuals (reduced chisquares) |
|---|---|---|---|
| $(Ax+B)^{-1}$ | -08.97E-02 | 45.23E-02 | **5612.88** |
| $(Ax+B)^1$ | 50.97E+00 | -98.81E+00 | **156.49** |
| $(Ax+B)^2$ | 02.99E+00 | -01.92E+00 | **119.89** |
| $(Ax+B)^3$ | 95.01E-02 | 81.96E-02 | **125.26** |

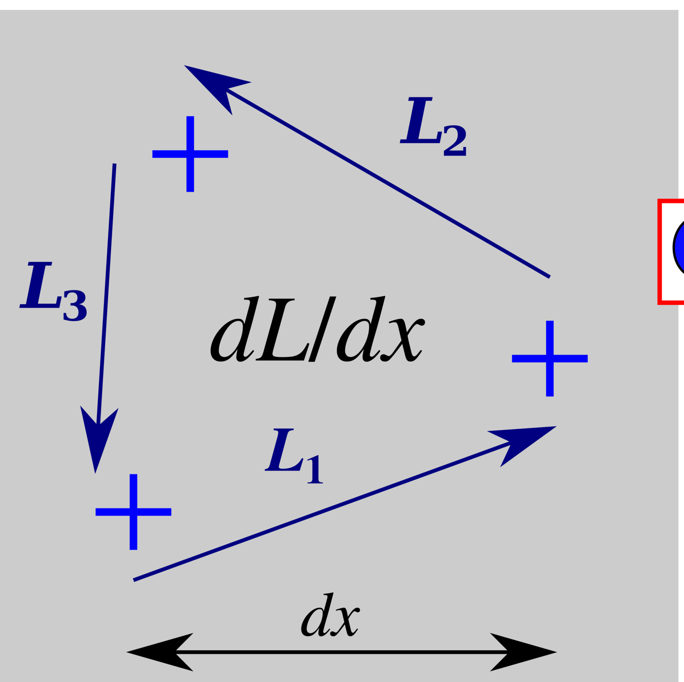


**Figure 3. Schematic diagram of *dL*/*dx*.**

This diagram shows vector relations of $dL/dx$ between three points. A length $L_i$ is expressed as a norm of a vector; $L_i = |\mathbf{L}_i|$, where $i = 1, 2$, and 3.

## Conclusion

**The random simulation could evaluate the sample-model interpolations that were based on the least squares method.**

**References cited**

Hammer, Ø. and Harper, D., A., T., 2006: *Paleontological Data Analysis*. Oxford, Blackwell Publishing.

Mustoe, G., E. and Smith, E., T., 2023: Timing of Opalization at Lightning Ridge, Australia: New Evidence from Opalized Fossils. *Minerals*, vol. 13, (12), 1471. https://doi.org/10.3390/min13121471

Wani, R., 2003: Taphofacies models for Upper Cretaceous Ammonoids from the Kotanbetsu area, northwestern Hokkaido, Japan. *Palaeogeography*, *Palaeoclimatology*, *Palaeoecology*, vol. 199, p. 71-82.

Williams, T., Kelley, C., *et al*., 1986-1993, 1998, 2004, and 2007-2019: *Gnuplot* 5.2: *An Interactive Plotting Program*. http://www.gnuplot.info/ (last accessed on 2023 Dec. 09).