Supplement of Hydrol. Earth Syst. Sci., 22, 1175–1192, 2018
https://doi.org/10.5194/hess-22-1175-2018-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Evaluation of statistical methods for quantifying fractal scaling in water-quality time series with irregular sampling

Qian Zhang et al.

Correspondence to: Qian Zhang (qzhang@chesapeakebay.net)

S1. Evaluation of Methods for Simulating Fractal Scaling

S2. Evaluation of Methods for Estimating Fractal Scaling

S3. Concentration Residuals for Real Water-Quality Data

S1. Evaluation of Methods for Simulating Fractal Scaling

To compare the performance of methods for estimating fractal scaling in irregular time series, we need to first generate time series with known spectral slopes. Thus, we compared seven common simulation methods for self-similar series generation (**Table S1**). For each simulation method, we generated two years of 15-minute data for six prescribed spectral slopes. We then obtained sub-samples at four frequencies with equal number of samples, which are hourly, 6-hourly, daily, and weekly (**Figure S1**). For each subset, we then estimated the spectral slope using each of the five proposed estimation methods (**Table S2**).

Table S1. Summary of the seven methods for simulating time series with fractal scaling.

No.	R function	R package or reference	For simulation of
1	*arfima.sim()*	*{arfima}*	ARFIMA time series
2	*fracdiff.sim()*	*{fracdiff}*	ARFIMA time series
3	*SimulateFGN()*	*{FGN}*	Fractional Gaussian noise
4	*fgnSim()*	*{fArma}*	Fractional Gaussian noise
5	*FDSimulate()*	*{fractal}*	Time-varying fractionally differenced process
6	*produce.short.fractional.noises()*	*Witt & Malamud (2013)*	Fractional noises
7	*ss.gen.fourier()*	*Paxson (1997)*	Approximate fractional Gaussian noise

Table S2. Summary of the five methods for estimating spectral slope in regular time series.

No.	R function	R package or reference	For estimation of spectral slope using
1	*hurstSpec()*	*{fractal}*	Spectral regression
2	*FDWhittle()*	*{fractal}*	Whittle's method
3	*DFA()*	*{fractal}*	Detrended fluctuation analysis
4	*FitFGN()*	*{FGN}*	Exact MLE estimation
5	*fracdiff()*	*{fracdiff}*	MLE for ARFIMA(d)

The simulation methods can only generate approximately self-similar time series – see an example in **Figure S1**. Regardless of the methods, the simulated time series generally exhibit smaller spectral slopes at coarser frequencies, particularly for series with strong fractal scaling. Considering all seven simulation methods (**Table S1**) and five estimation methods (**Table S2**), *i.e.*, 35 pairs in total, the departure from self-similarity varies with the selected methods. In general, the "*FDSimulat*e" simulation method appears to produce more "self-similar" series than the other methods and the "*fracdiff*" estimation method appears to provide the most certain estimation of the spectral slope (**Figure S2**).

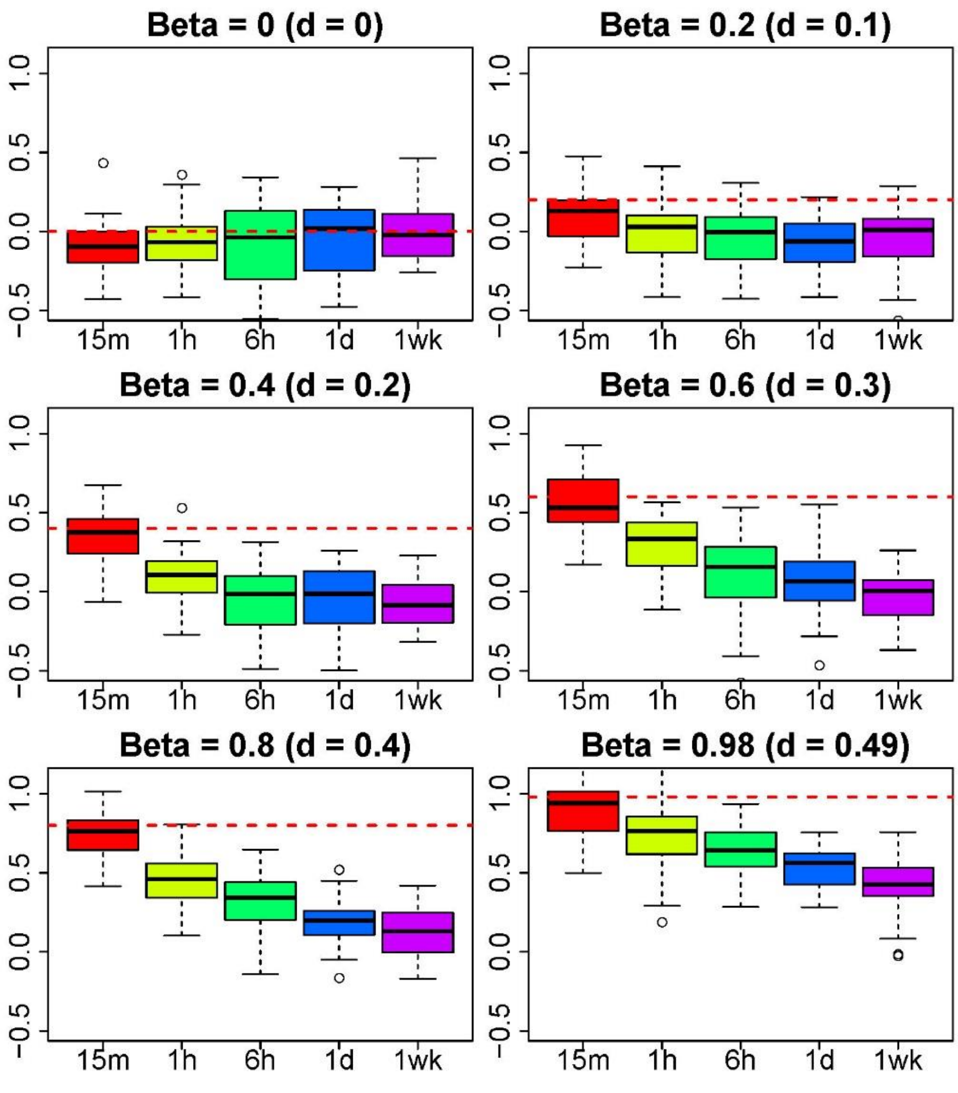

Figure S1. Estimated spectral slope using the "*FDwhittle*" method for time series simulated with the "*fracdiff.sim*" method and for sub-sampled time series (*i.e.*, hourly, 6-hourly, daily, and weekly). The red dashed line indicates the spectral slope used to simulating the time series.

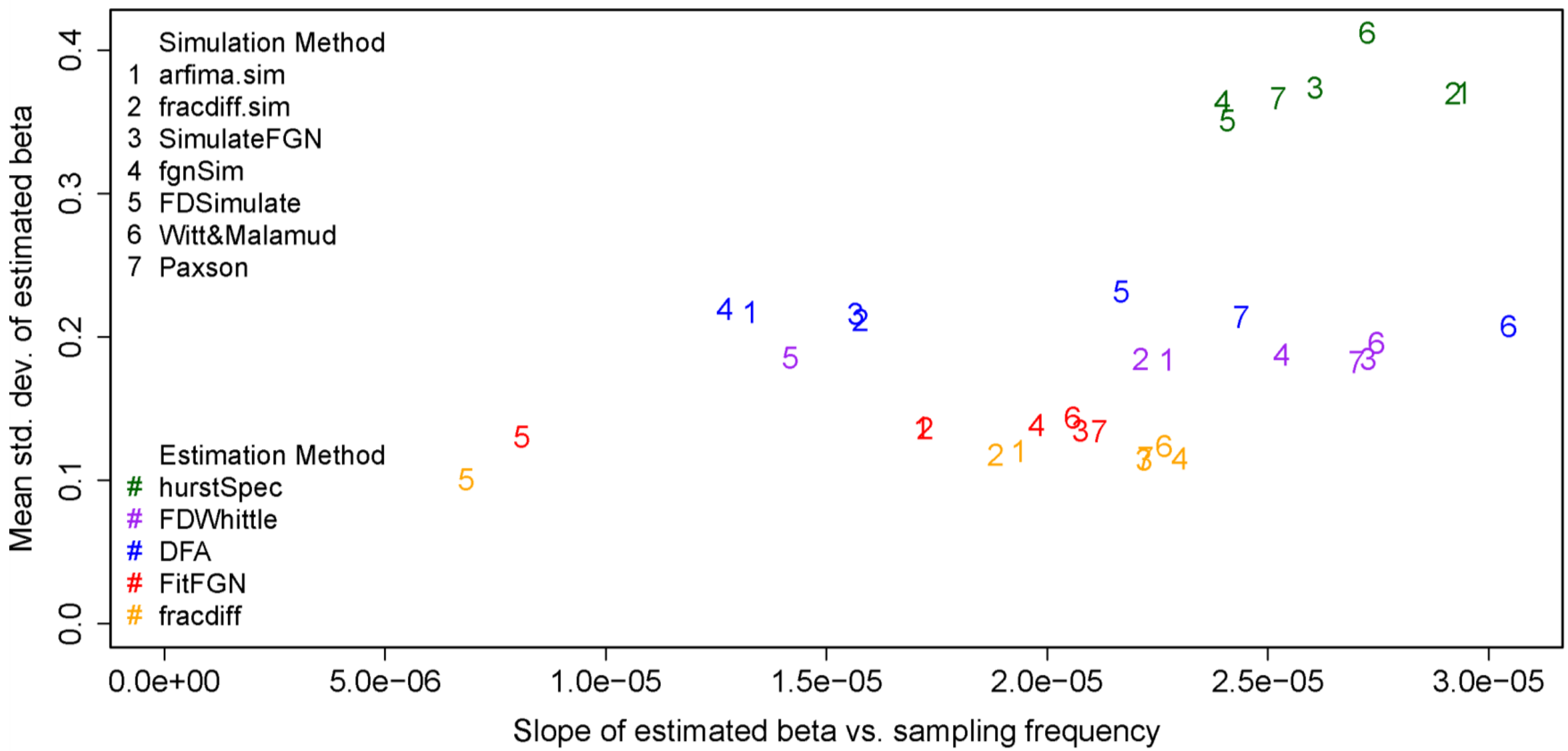

Figure S2. Comparison of the 35 pairs of simulation and estimation methods for self-similar time series. The x-axis shows the slope of the estimated β vs. sampling frequency (*measure of "self-similarity"*). See Figure S1 for an example showing the variation of the estimated β as a function of sampling frequency. The y-axis shows the mean standard deviation of the estimated β (*measure of "uncertainty"*). The numbers shown on plot indicate simulation methods, whereas the colors indicate estimation methods. See Tables 1-2 for method details.

S2. Evaluation of Methods for Estimating Fractal Scaling

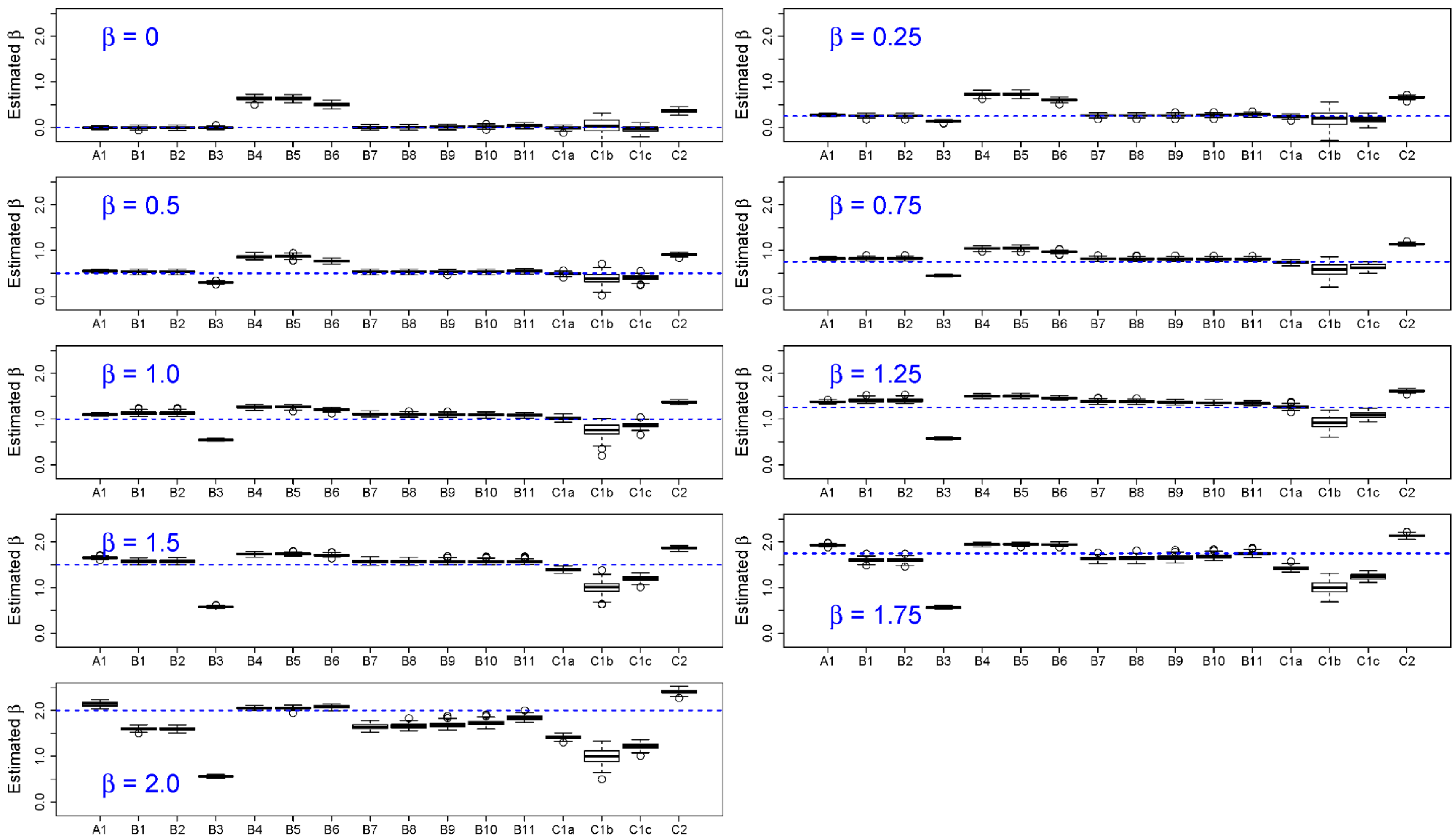

Figure S3. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 0.01, μ = 1) distributed gap intervals. The blue dashed lines indicate the true β values.

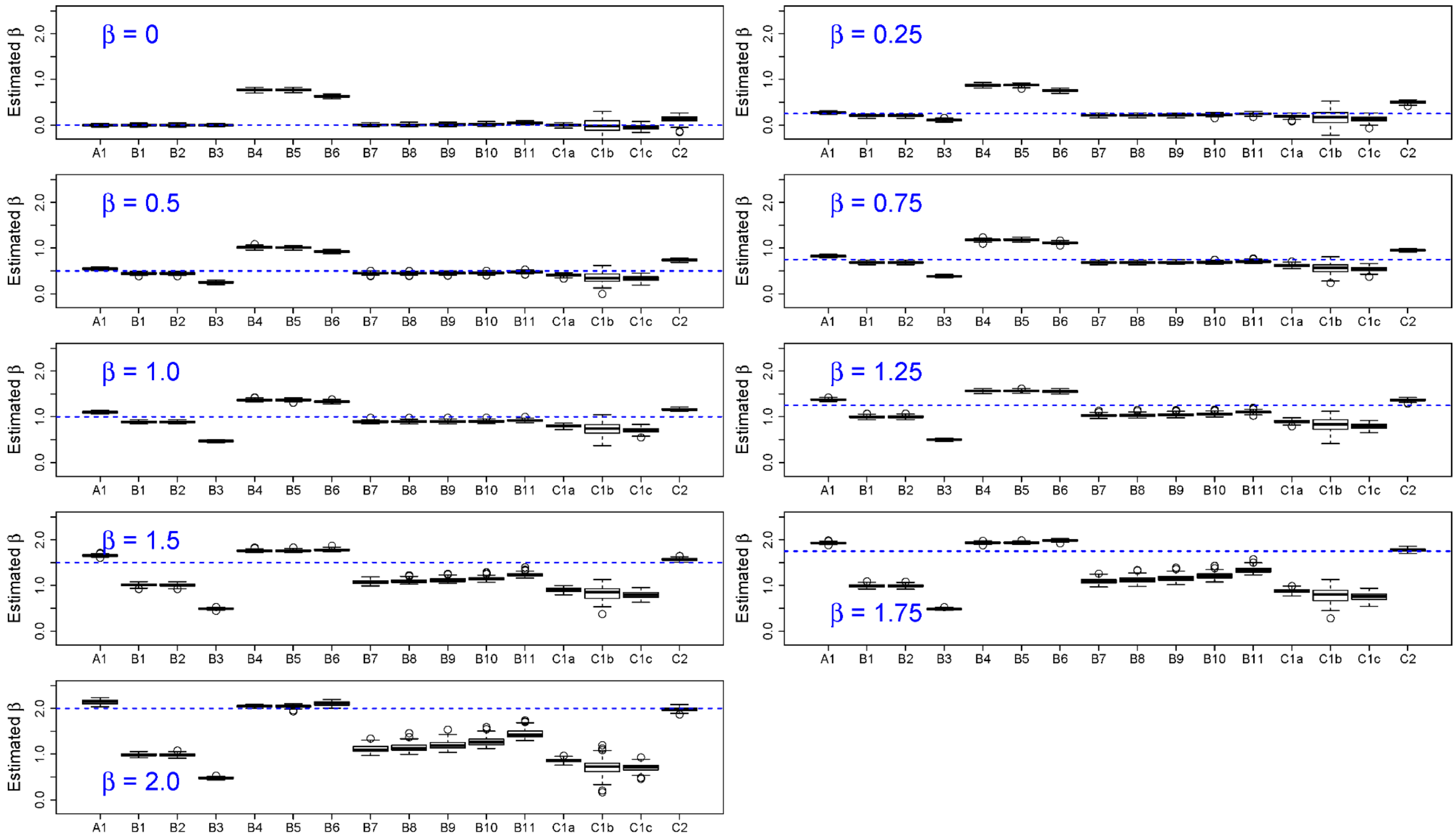

Figure S4. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 0.1, μ = 1) distributed gap intervals. The blue dashed lines indicate the true β values.

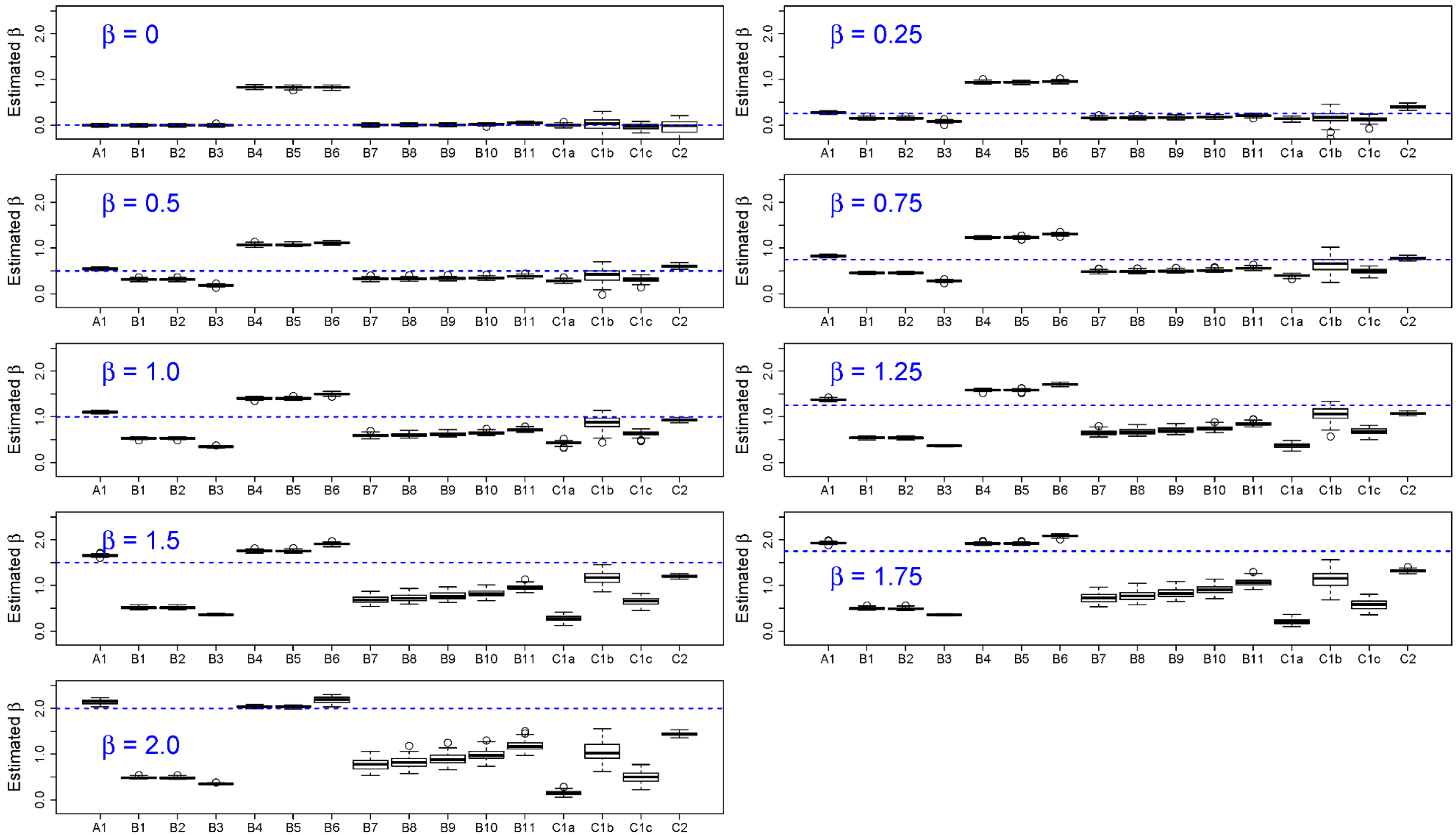

Figure S5. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 1, μ = 1) distributed gap intervals. The blue dashed lines indicate the true β values.

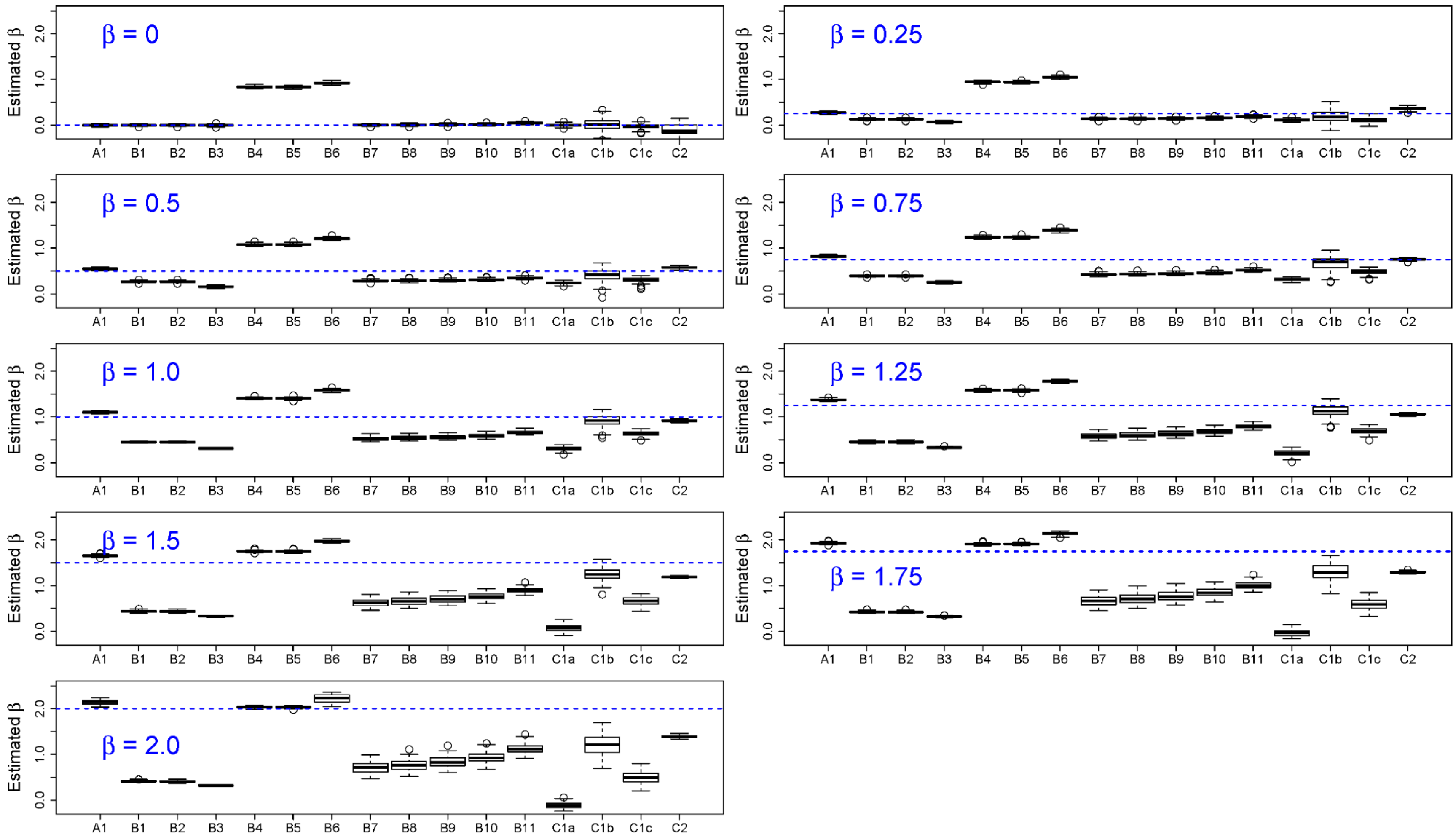

Figure S6. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 10, μ = 1) distributed gap intervals. The blue dashed lines indicate the true β values.

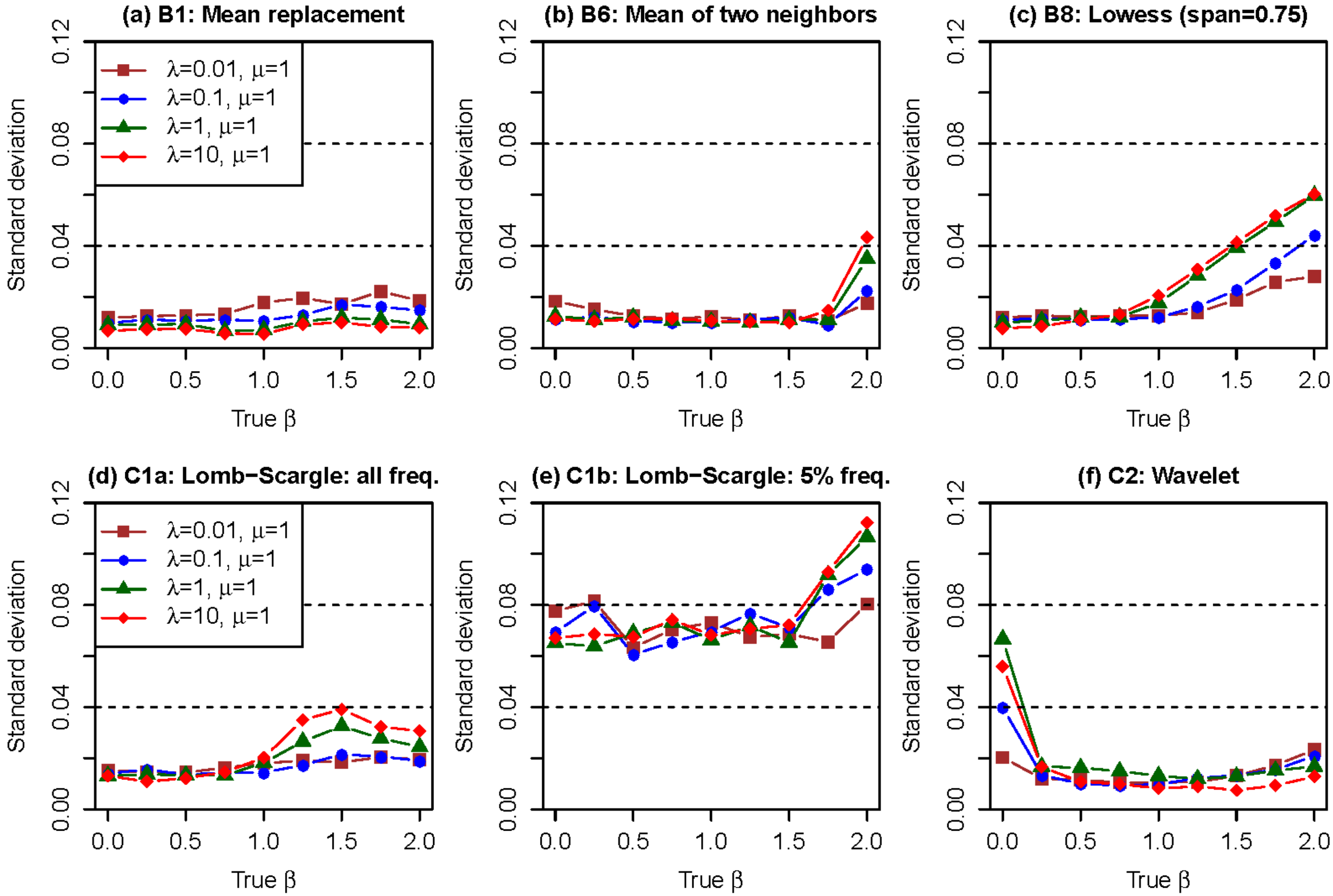

Figure S7. Comparison of standard deviation in estimated spectral slope in irregular data that are simulated with varying prescribed β values (100 replicates), length of 9125, and mean gap interval of 2 (*i.e.*, μ = 1).

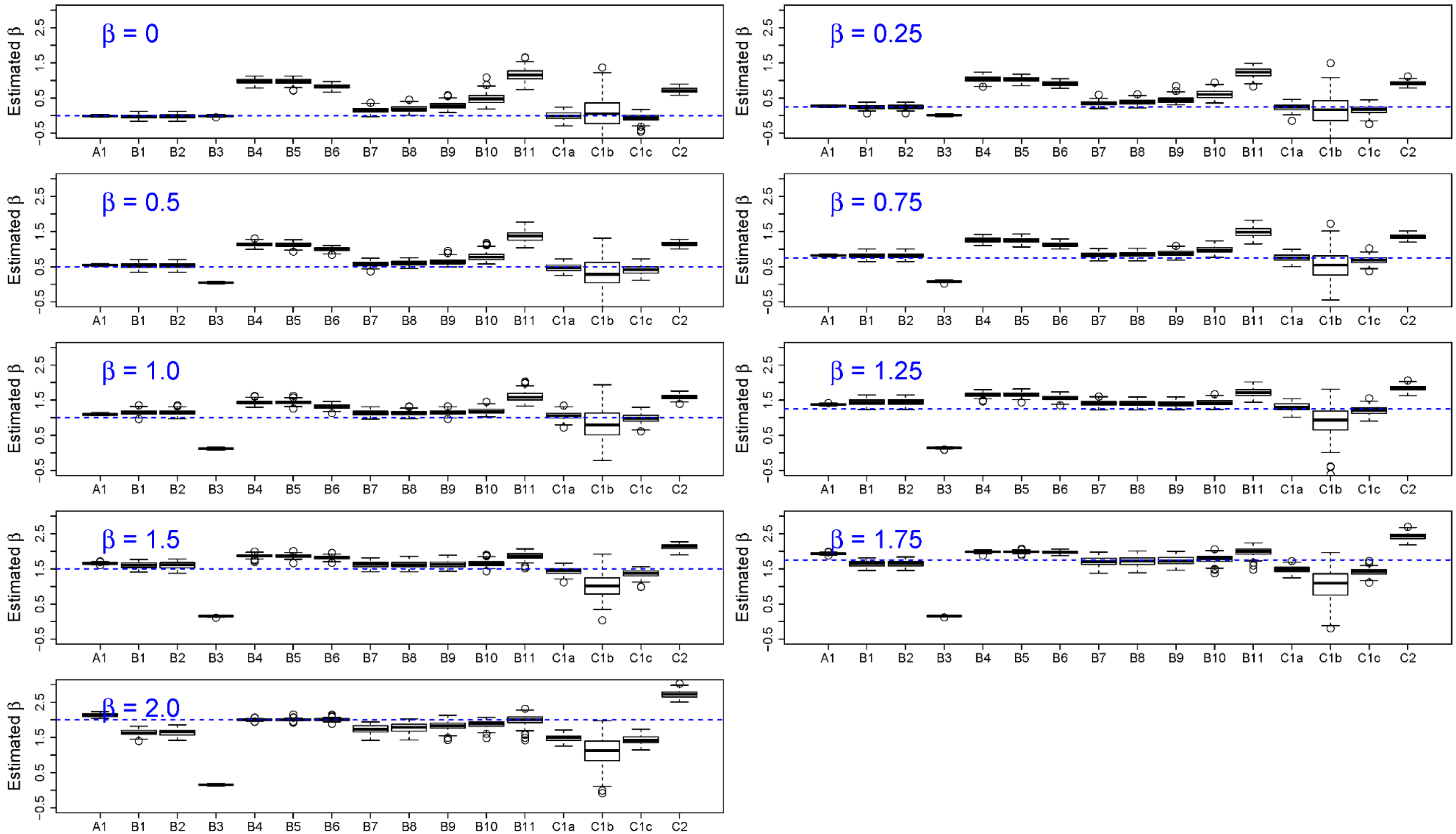

Figure S8. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 0.01, μ = 14) distributed gap intervals. The blue dashed lines indicate the true β values.

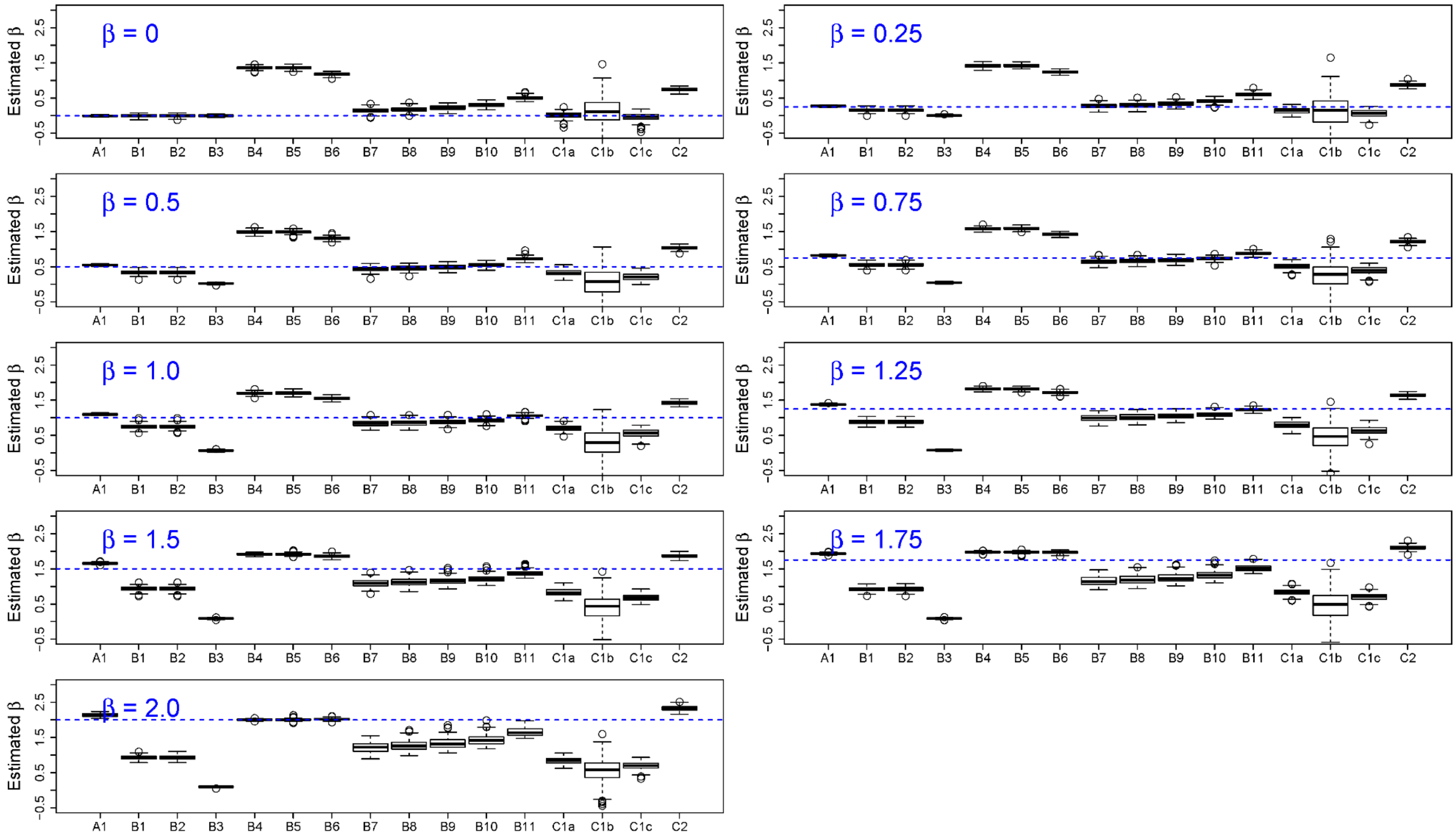

Figure S9. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 0.1, μ = 14) distributed gap intervals. The blue dashed lines indicate the true β values.

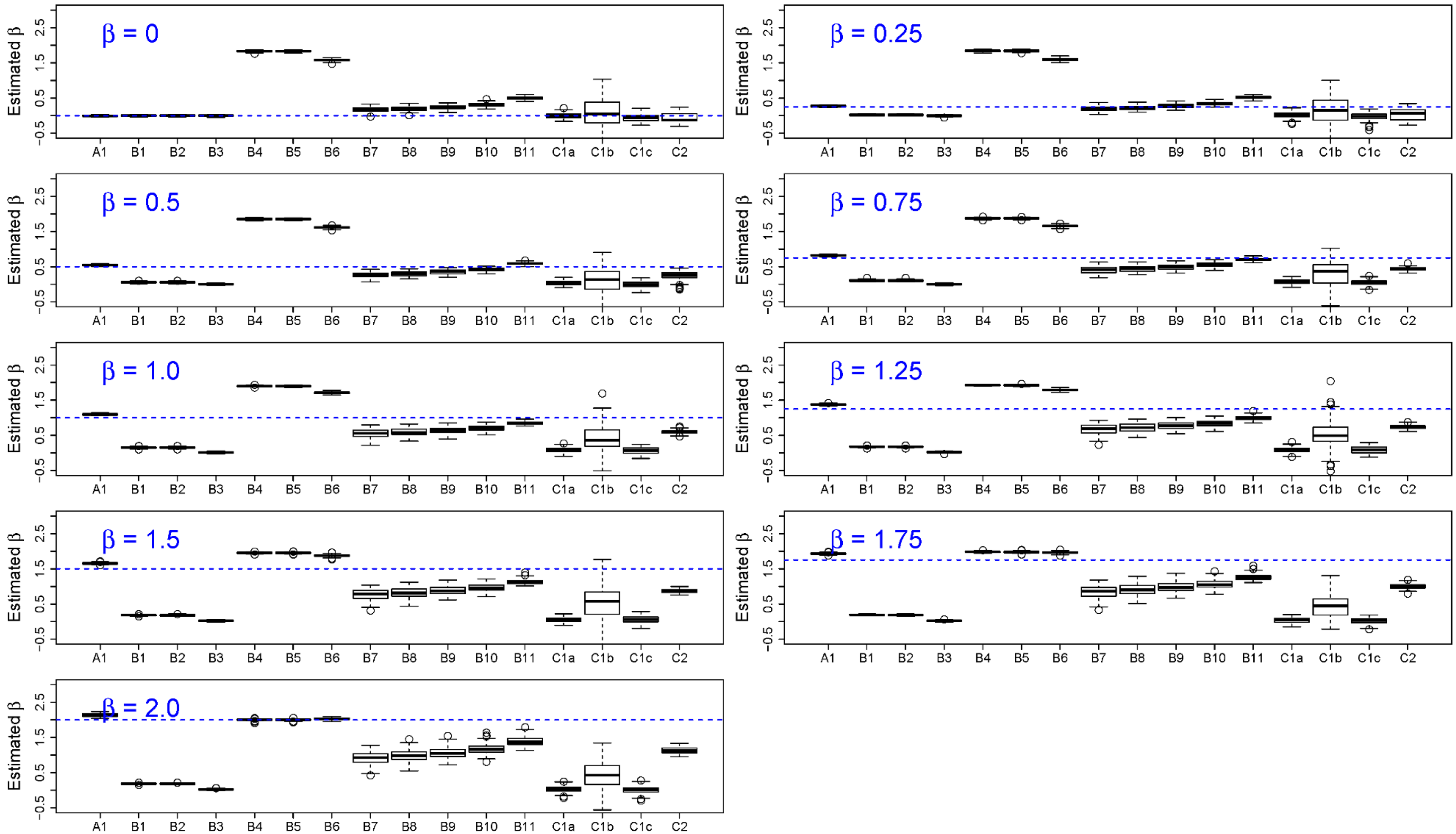

Figure S10. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 1, μ = 14) distributed gap intervals. The blue dashed lines indicate the true β values.

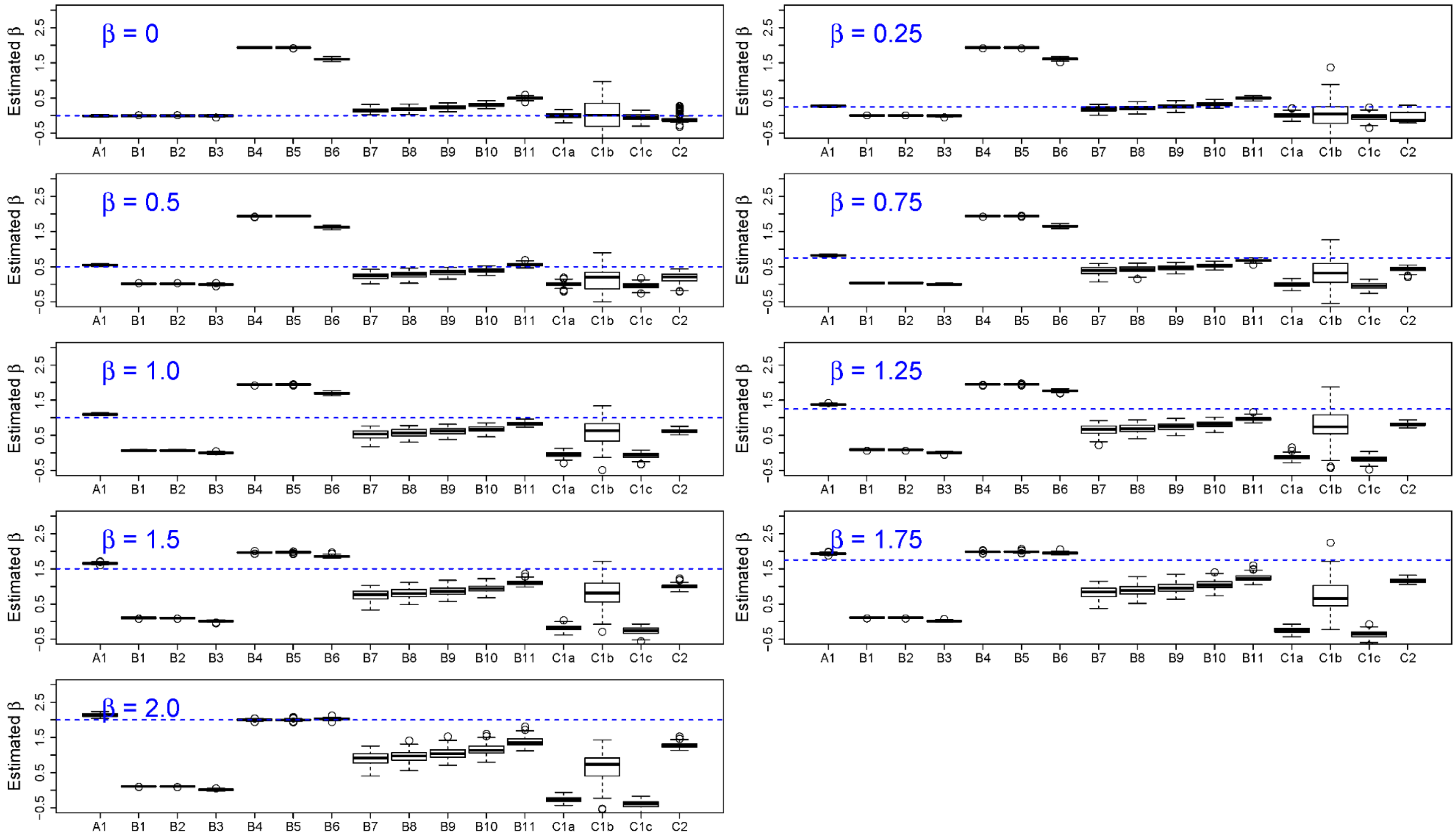

Figure S11. Comparison of methods for estimating spectral slope in irregular data (100 replicates) that are simulated with varying prescribed β values, length of 9125, and NB (λ = 10, μ = 14) distributed gap intervals. The blue dashed lines indicate the true β values.

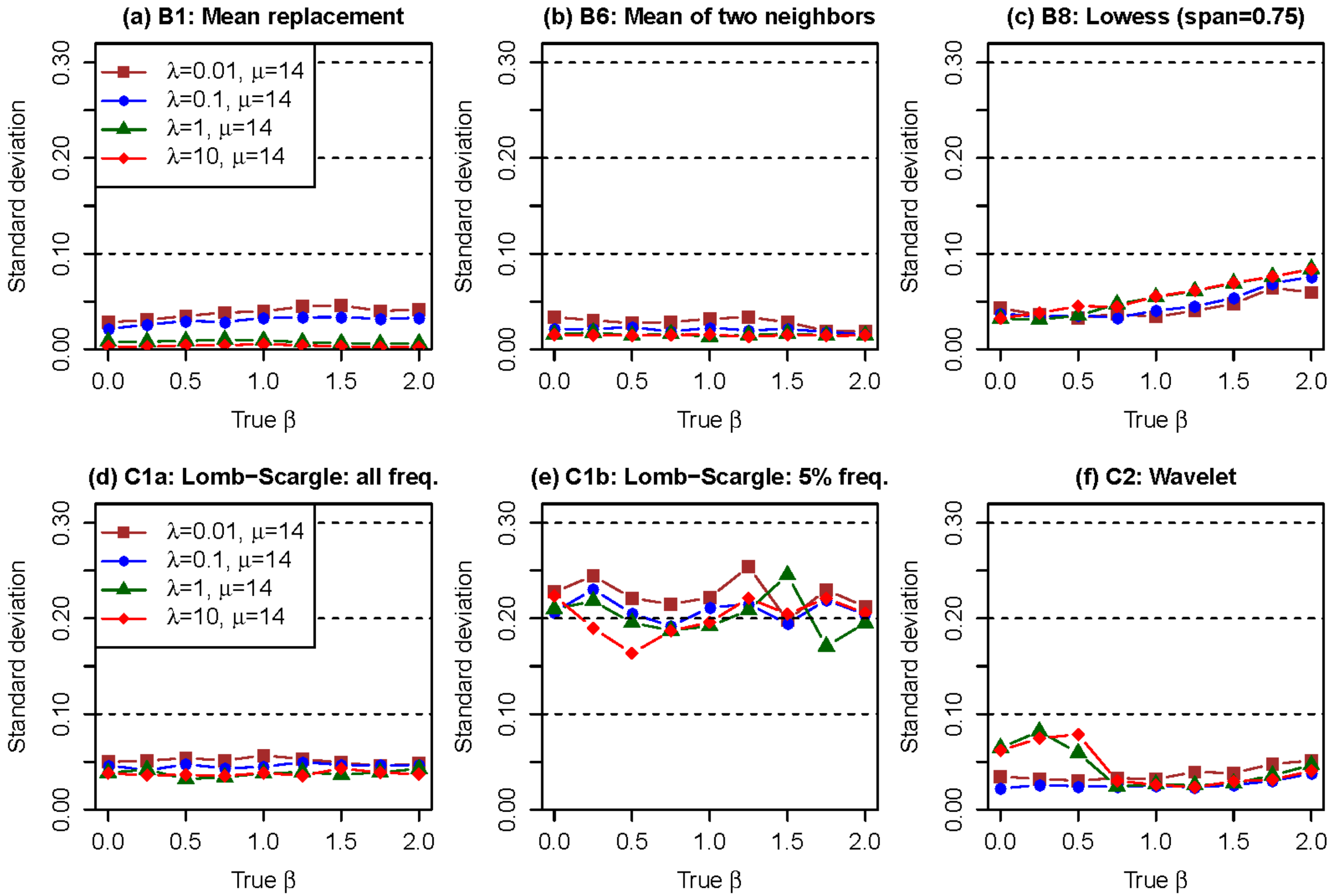

Figure S12. Comparison of standard deviation in estimated spectral slope in irregular data that are simulated with varying prescribed β values (100 replicates), length of 9125, and mean gap interval of 15 (*i.e.*, μ = 14).

S3. Concentration Residuals for Real Water-Quality Data

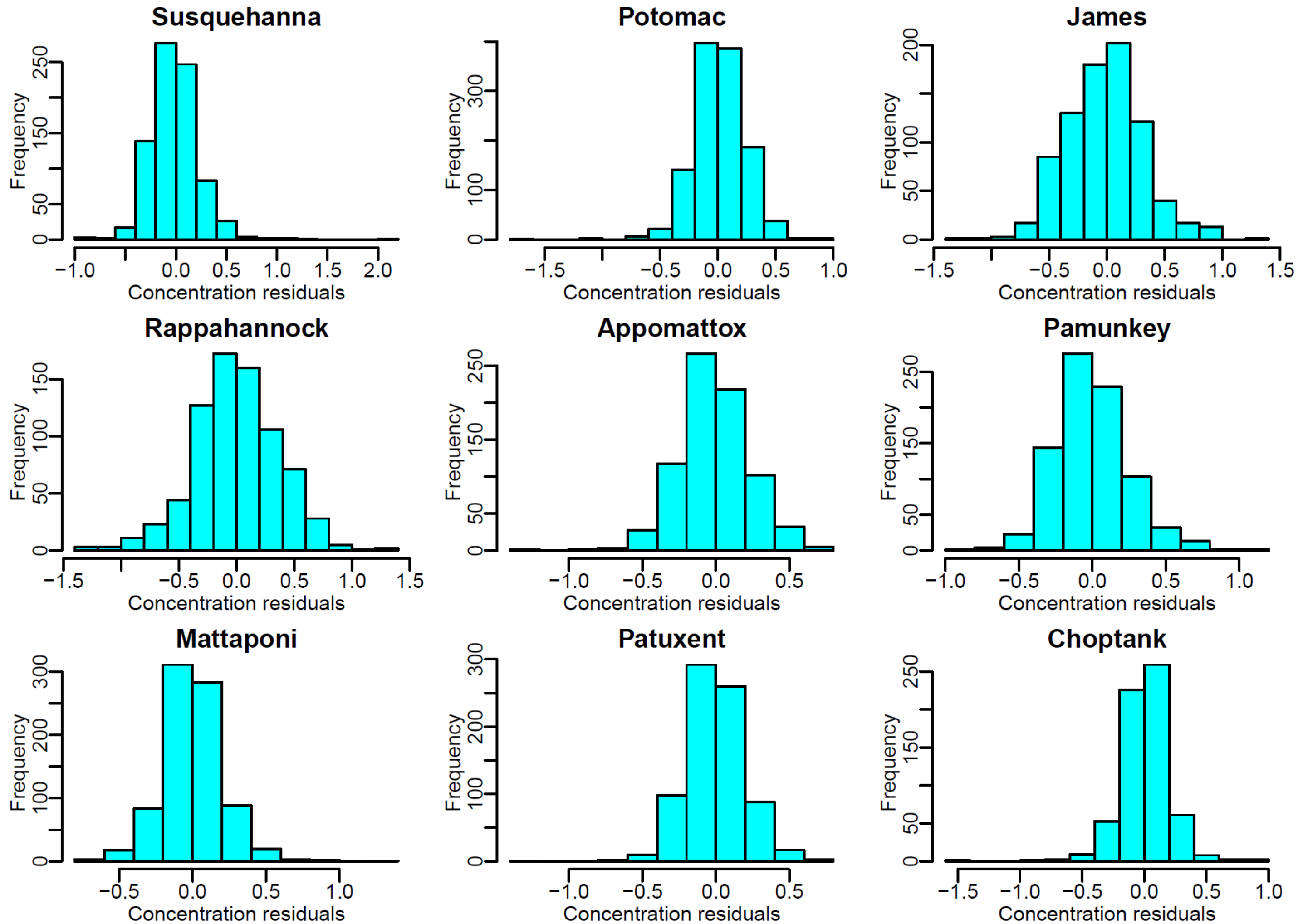

Figure S13. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total nitrogen (TN) at the nine Chesapeake Bay monitoring sites. See Table 1 for site and data details.

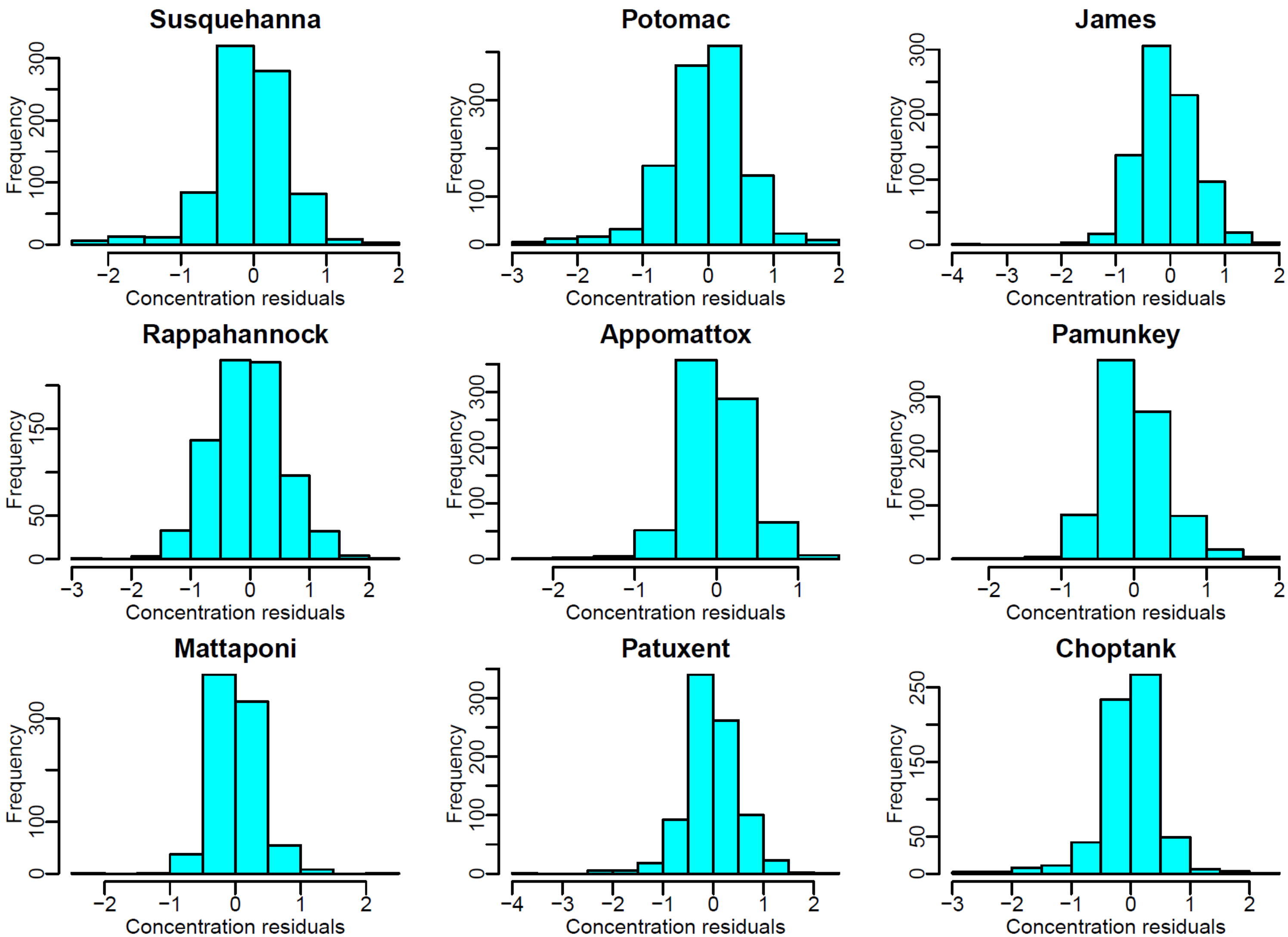

Figure S14. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total phosphorus (TP) at the nine Chesapeake Bay monitoring sites. See Table 1 for site and data details.

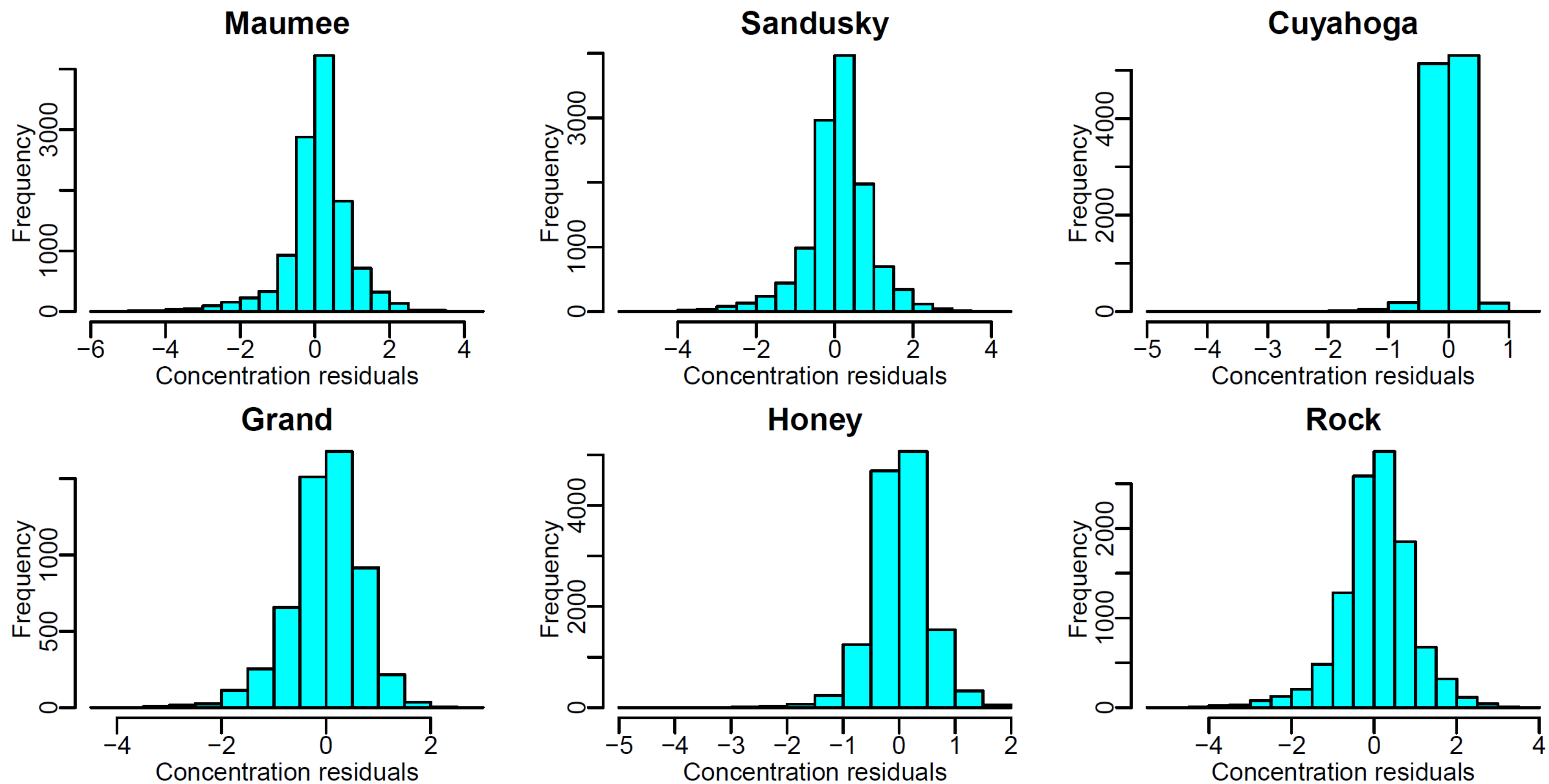

Figure S15. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for nitrate-plus-nitrite (NO_x) at the six Lake Erie and Ohio River monitoring sites. See Table 1 for site and data details.

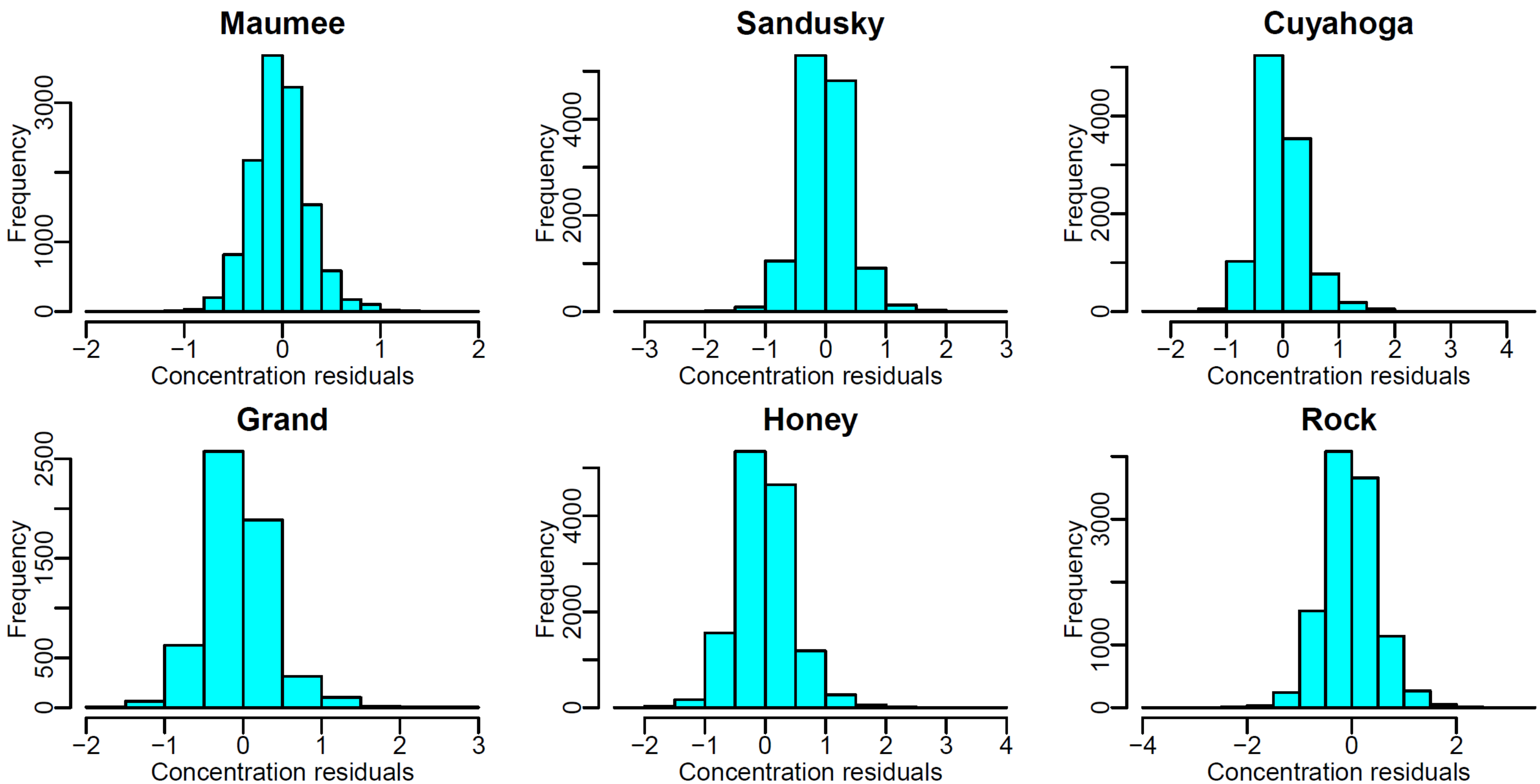

Figure S16. Histogram of concentration residuals from the WRTDS method, expressed in natural log concentration units, for total phosphorus (TP) at the six Lake Erie and Ohio River monitoring sites. See Table 1 for site and data details.