

1. Laboratory Measurement Bias and Repeatability Model (LBM)

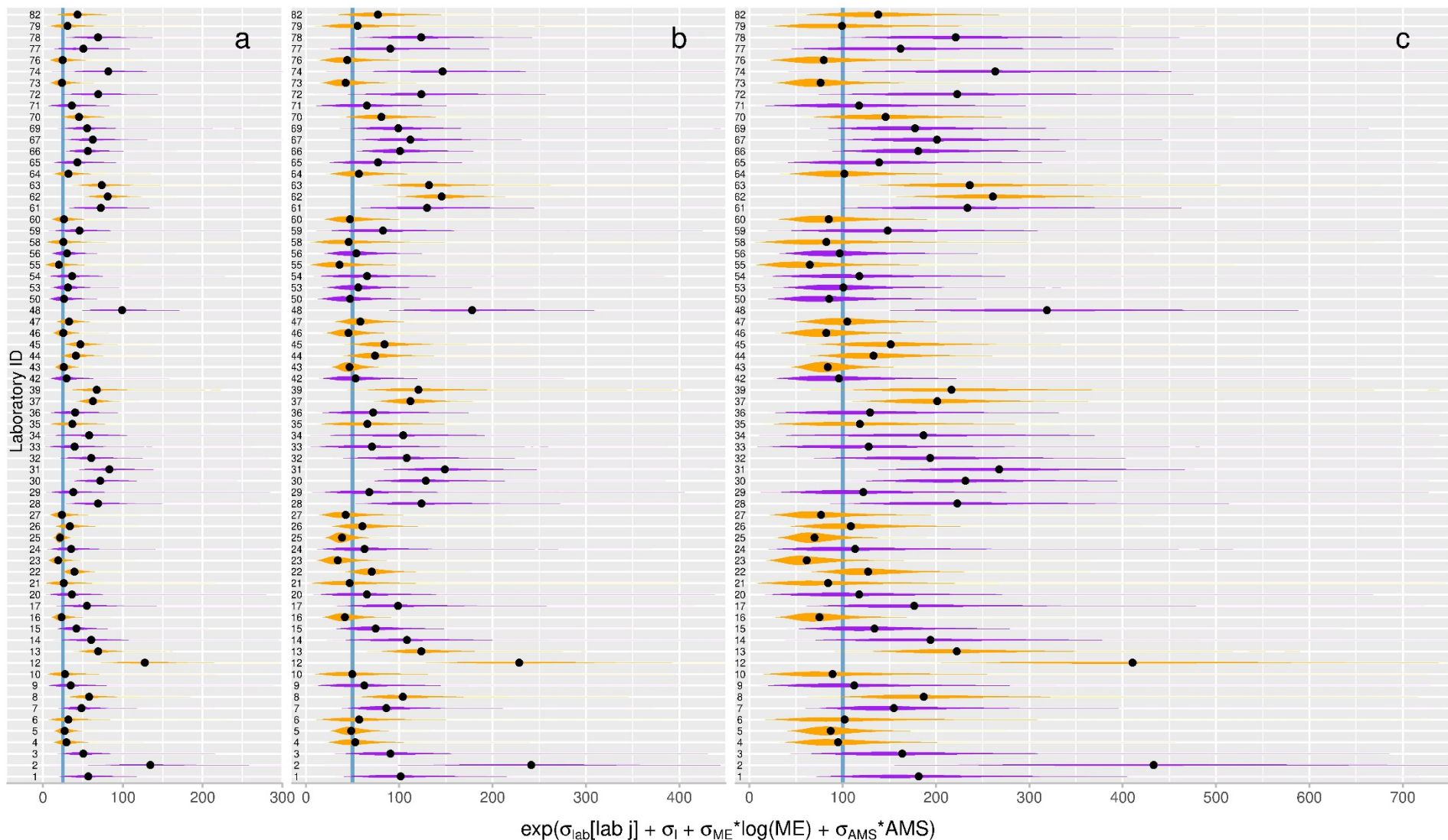


Figure S3.5. Three examples showing how reported measurement error (ME) affects within-laboratory standard deviations: (a) ME = 25 ^{14}C years, (b) ME = 50 ^{14}C years, and (c) ME = 100 ^{14}C years. Vertical blue lines mark the reported error values. Gold posterior densities are AMS laboratories and purple posterior densities are GPC/LSC laboratories. Black dots mark median values in each posterior density. GPC/LSC effects on within-laboratory standard deviations are included in these posterior distributions. X-axes are on the ^{14}C year scale.

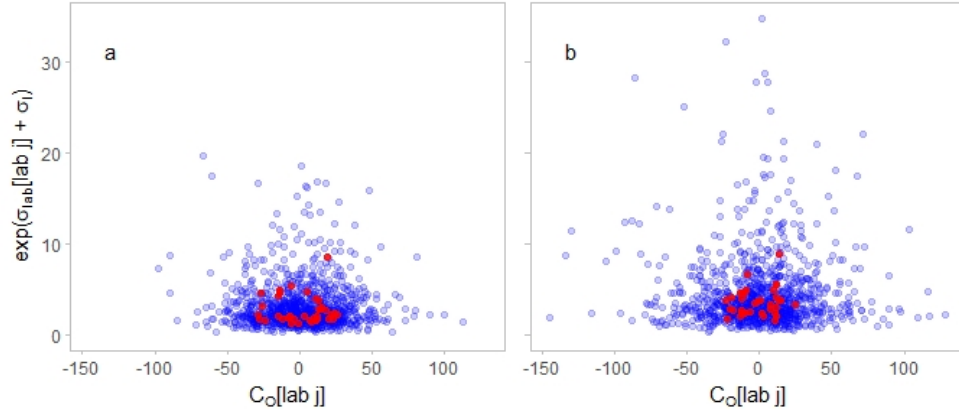


Figure S3.6. Laboratory parameters for (a) AMS and (b) GPC/LSC laboratories. Blue dots show 1000 simulated laboratories and red dots show mean posterior values for laboratories in the VIRI dataset. Both axes are displayed on the ^{14}C year scale.

Table S3.1. Posterior means and 95% highest posterior density intervals (HPDI) for VIRI laboratory parameters. The effects of GPC/LSC measurement methods are excluded here. In other words, all laboratories are treated here as AMS laboratories to express variability that is due to laboratory identity exclusive of ^{14}C measurement method.

Lab. ID [j]	Laboratory mean offset (^{14}C years) $C_0[\text{lab } j]$		Intra-laboratory standard deviation (^{14}C years) $\exp(\sigma_{\text{lab}}[\text{lab } j] + \sigma_l)$	
1	1.27	(-48.45 – 47.78)	3.54	(0.41 – 8.77)
2	14.87	(-38.29 – 71.87)	7.89	(1.44 – 17.99)
3	21.31	(-8.09 – 56.94)	2.88	(0.65 – 6.11)
4	-14.20	(-42.09 – 13.91)	2.32	(0.47 – 5.08)
5	-7.52	(-29.32 – 14.78)	2.09	(0.47 – 4.42)
6	11.97	(-18.54 – 41.52)	2.62	(0.41 – 6.40)
7	7.09	(-31.19 – 49.82)	2.85	(0.48 – 6.37)
8	11.62	(-25.75 – 47.68)	4.54	(0.94 – 9.87)
9	8.95	(-30.30 – 49.86)	2.18	(0.24 – 5.52)
10	9.47	(-27.11 – 45.79)	2.32	(0.25 – 5.59)
12	21.16	(-27.98 – 77.48)	9.92	(2.11 – 21.35)
13	5.61	(-31.05 – 41.74)	5.30	(1.25 – 11.18)
14	-12.61	(-65.43 – 37.58)	3.70	(0.53 – 8.91)
15	-16.92	(-52.35 – 17.87)	2.44	(0.42 – 5.30)
16	7.74	(-10.95 – 26.83)	1.82	(0.35 – 3.92)
17	-1.70	(-51.05 – 46.33)	3.49	(0.41 – 8.72)
20	-8.79	(-44.27 – 26.74)	2.24	(0.19 – 5.50)
21	11.63	(-24.49 – 43.44)	2.19	(0.18 – 5.57)
22	18.04	(-6.20 – 42.05)	3.03	(0.76 – 6.40)
23	-0.60	(-23.73 – 22.72)	1.53	(0.24 – 3.42)
24	2.00	(-38.73 – 45.11)	2.21	(0.20 – 5.35)
25	7.93	(-4.37 – 20.12)	1.60	(0.49 – 3.10)
26	21.80	(-8.74 – 50.63)	2.65	(0.58 – 5.81)
27	10.90	(-17.98 – 41.21)	1.92	(0.37 – 4.29)

28	-9.63	(-58.45 – 39.77)	4.14	(0.78 – 9.69)
29	-6.62	(-55.05 – 43.62)	2.42	(0.22 – 6.21)
30	9.20	(-35.61 – 55.32)	4.28	(0.69 – 9.66)
31	11.28	(-33.59 – 57.48)	4.85	(0.84 – 10.67)
32	-18.65	(-73.86 – 29.13)	3.62	(0.47 – 8.45)
33	-15.23	(-59.45 – 28.70)	2.42	(0.27 – 6.03)
34	-21.21	(-73.84 – 29.64)	3.47	(0.30 – 8.31)
35	23.47	(-17.02 – 63.13)	3.03	(0.33 – 7.25)
36	-10.87	(-54.42 – 29.74)	2.47	(0.33 – 6.01)
37	-15.38	(-45.41 – 15.47)	4.73	(1.16 – 9.54)
39	-29.02	(-84.32 – 20.78)	5.36	(0.81 – 11.78)
42	-11.70	(-46.12 – 24.47)	1.80	(0.26 – 4.24)
43	-11.68	(-32.10 – 9.40)	2.02	(0.44 – 4.22)
44	13.21	(-20.36 – 44.39)	3.24	(0.59 – 7.06)
45	-25.20	(-52.19 – 2.26)	3.59	(0.82 – 7.41)
46	19.72	(-0.37 – 39.30)	1.95	(0.45 – 4.03)
47	26.32	(-1.93 – 52.93)	2.53	(0.53 – 5.38)
48	-8.16	(-59.22 – 40.86)	5.93	(0.94 – 13.58)
50	-18.49	(-51.68 – 12.90)	1.58	(0.20 – 3.61)
53	2.49	(-21.70 – 24.94)	1.82	(0.34 – 3.93)
54	-8.95	(-48.71 – 30.48)	2.26	(0.25 – 5.58)
55	-5.29	(-34.24 – 22.08)	1.67	(0.14 – 4.15)
56	9.84	(-21.37 – 39.56)	1.79	(0.30 – 4.04)
58	-3.11	(-32.11 – 28.43)	2.14	(0.17 – 5.16)
59	11.95	(-35.21 – 61.90)	2.84	(0.29 – 6.99)
60	-26.96	(-53.13 – 0.31)	2.10	(0.37 – 4.65)
61	-9.55	(-58.01 – 39.37)	4.40	(0.62 – 10.24)
62	-6.03	(-38.43 – 26.03)	6.26	(1.31 – 13.27)
63	-14.42	(-60.64 – 27.33)	5.83	(1.07 – 12.76)
64	-28.35	(-60.61 – 6.38)	2.55	(0.44 – 5.71)
65	6.83	(-29.71 – 45.51)	2.62	(0.35 – 6.23)
66	13.15	(-21.49 – 52.30)	3.26	(0.54 – 7.00)
67	10.71	(-34.82 – 57.40)	3.77	(0.62 – 8.79)
69	13.29	(-30.11 – 60.55)	3.29	(0.50 – 7.46)
70	15.81	(-17.36 – 51.36)	3.53	(0.74 – 7.60)
71	-11.08	(-55.76 – 34.09)	2.24	(0.22 – 5.64)
72	-12.08	(-65.29 – 41.20)	4.21	(0.56 – 10.05)
73	-22.75	(-50.96 – 6.54)	1.92	(0.28 – 4.38)
74	12.59	(-38.32 – 67.89)	4.98	(0.68 – 11.78)
76	23.24	(0.83 – 44.15)	1.98	(0.34 – 4.50)
77	-3.51	(-51.84 – 47.22)	3.24	(0.34 – 8.18)
78	-13.89	(-60.91 – 29.29)	4.11	(0.60 – 9.37)
79	-6.77	(-42.76 – 29.40)	2.61	(0.32 – 6.37)
82	15.92	(-33.75 – 69.33)	3.67	(0.40 – 9.18)

1.4. HMC Diagnostics and Posterior Predictive Checks

To ensure convergence of HMC chains, we examined trace plots, R-hat values, and effective sample sizes. Trace plots suggest convergence of HMC chains for all parameters. All R-hat values are below 1.01. Effective sample sizes close to 10,000 indicate efficient sampling of posterior parameter spaces. Of the 148 parameters, 101 have effective sample sizes of 10,000. The median effective sample size is also 10,000, the mean is 8362, and the minimum is 1056.4 (Figure S3.7).

We also performed a posterior predictive check by simulating data from the model and plotting these simulated data against the VIRI ^{14}C measurements (Figure S3.8). For simplicity, where a laboratory contributed multiple ^{14}C measurements for a given sample material, we used the average reported measurement error for that laboratory ID and sample (for example, laboratories 5 and 22 in the panel for Sample G). In all but nine of the 361 VIRI ^{14}C measurements (97.5%), the 95% prediction intervals overlap the observed values. This indicates that the LBM does a reasonable job recovering the observed values.

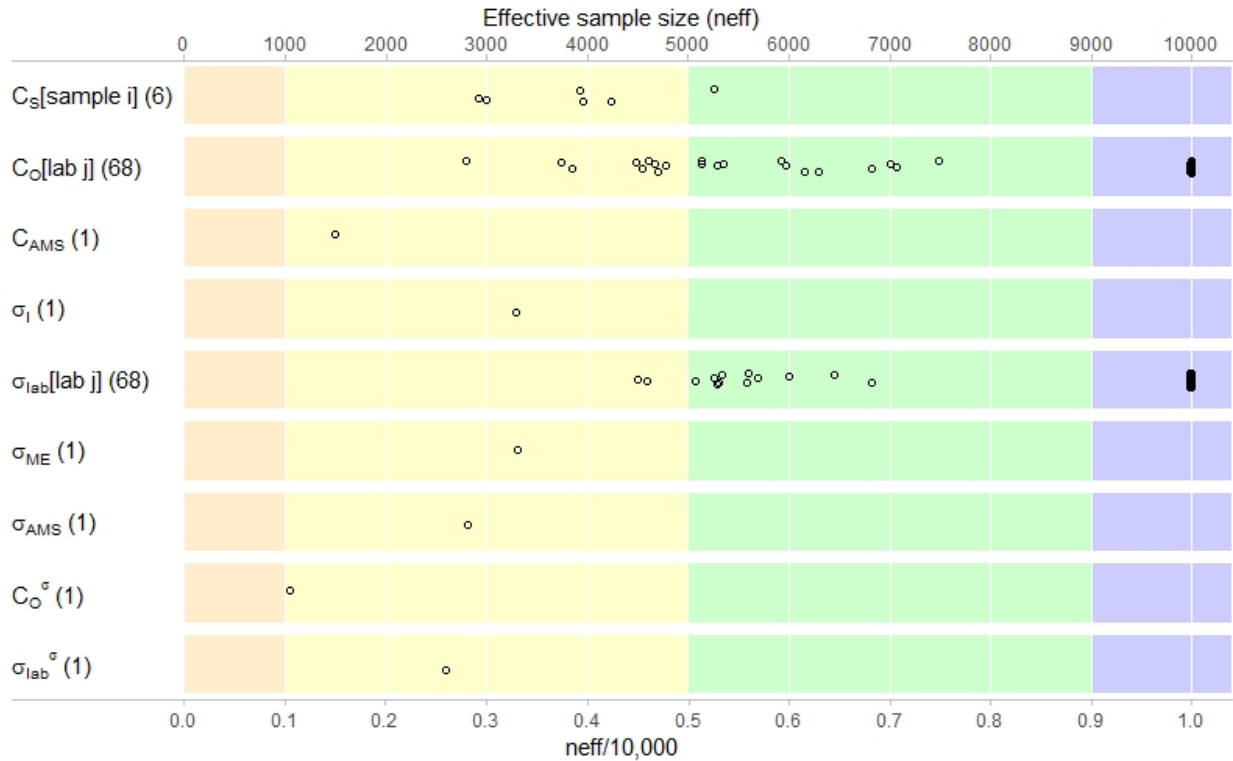


Figure S3.7. Distribution of effective sample sizes across parameters. Parenthetical values show the number of parameter distributions that were sampled for each parameter type.

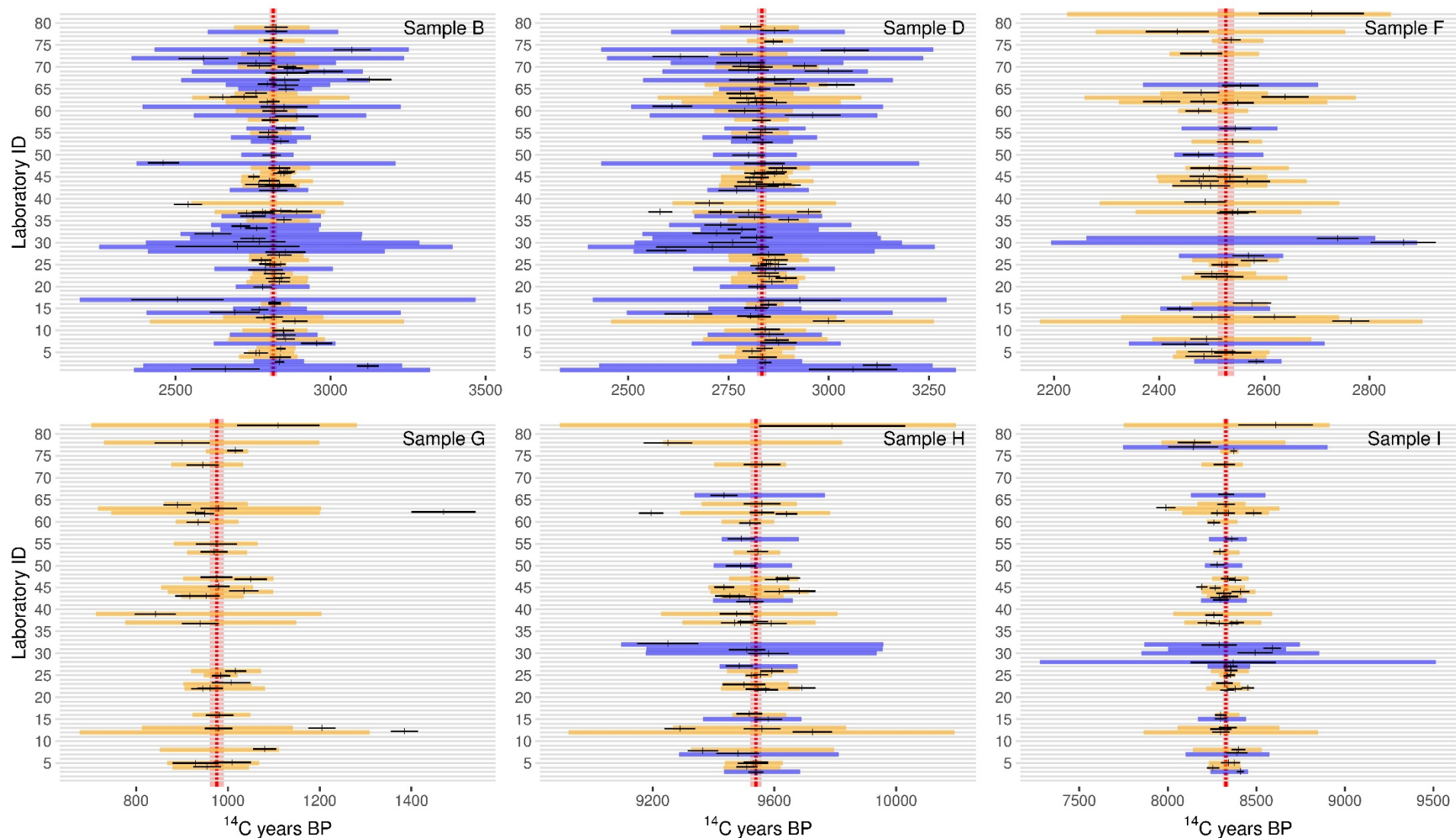


Figure S3.8. Posterior predictive check for the LBM. Each panel is a sample material reported in the VIRI study (Scott et al. 2007, 2010a, 2010b). Laboratory IDs are listed along the y-axes. Reported ^{14}C measurements and their associated errors (1σ) are indicated by vertical and horizontal black segments, respectively. Horizontal gold and purple bars show the 95% posterior prediction intervals for AMS and GPC/LSC laboratories, respectively. The dashed red lines and bands indicate the mean and 95% highest posterior density intervals for the ^{14}C value of each sample material.

Citations

1. Scott EM, Cook GT, Naysmith P, Bryant C, O'Donnell D (2007) A Report on Phase 1 of the 5th International Radiocarbon Intercomparison (VIRI). *Radiocarbon* 49(02):409–426.
2. Scott EM, Cook GT, Naysmith P (2010) The Fifth International Radiocarbon Intercomparison (VIRI): An Assessment of Laboratory Performance in Stage 3. *Radiocarbon* 52(03):859–865.
3. Scott EM, Cook GT, Naysmith P (2010) A Report on Phase 2 of the Fifth International Radiocarbon Intercomparison (VIRI). *Radiocarbon* 52(03):846–858.
4. Scott EM, Cook GT, Naysmith P (2007) Error and Uncertainty in Radiocarbon Measurements. *Radiocarbon* 49(02):427–440.
5. McElreath R (2017) *rethinking v1.59: Statistical Rethinking book package* Available at: <https://www.github.com/user/rmcelreath/rethinking>.
6. Stan Development Team (2018) *RStan 2.17.2: The R interface for Stan* Available at: <http://mc-stan.org>.