ID	name	RMSE	MAE	ME	$\mathbb{R}^2$	m	τ	ES
EXT09	clogP (Biobyte)	0.23 [0.16, 0.29]	0.17 [0.12, 0.23]	0.01 [-0.07, 0.10]	0.94 [0.86, 0.98]	0.94 [0.80, 1.02]	0.85 [0.75, 0.93]	0.27 [0.07, 0.42]
EXT12	MoKa_logP	0.28 [0.19, 0.35]	0.22 [0.15, 0.29]	0.03 [-0.07, 0.14]	0.91 [0.81, 0.97]	0.93 [0.76, 1.05]	0.83 [0.72, 0.91]	0.23 [0.07, 0.40]
EXT11	logP(o/w) (MOE)	0.32 [0.22, 0.41]	0.24 [0.17, 0.32]	0.11 [-0.01, 0.22]	0.90 [0.81, 0.95]	0.85 [0.75, 1.02]	0.80 [0.66, 0.91]	0.28 [0.11, 0.49]
EXT10	h_logP (MOE)	0.43 [0.33, 0.51]	0.35 [0.26, 0.45]	0.17 [0.02, 0.32]	0.83 [0.62, 0.93]	0.91 [0.67, 1.06]	0.74 [0.55, 0.90]	0.09 [0.01, 0.26]
EXT13	SlogP (MOE)	0.59 [0.48, 0.69]	0.49 [0.36, 0.62]	0.28 [0.08, 0.48]	0.71 [0.41, 0.87]	0.84 [0.60, 1.04]	0.55 [0.33, 0.73]	0.13 [0.01, 0.26]
EXT08	YANK-SMIRNOFF-tip3p-dry-oct	1.26 [0.88, 1.61]	0.97 [0.69, 1.29]	-0.85 [-1.21, -0.51]	0.56 [0.18, 0.83]	1.12 [0.63, 1.47]	0.50 [0.23, 0.74]	1.15 [0.99, 1.28]
EXT07	YANK-GAFF-tip3p-dry-oct	1.27 [0.68, 1.76]	0.88 [0.56, 1.26]	-0.74 [-1.16, -0.37]	0.55 [0.19, 0.88]	1.21 [0.79, 1.62]	0.60 [0.34, 0.81]	1.22 [1.02, 1.37]
EXT02	YANK-GAFF-tip3p-wet-oct	1.83 [1.39, 2.24]	1.53 [1.16, 1.92]	0.24 [-0.45, 0.91]	0.27 [0.01, 0.59]	1.17 [0.19, 1.85]	0.31 [0.04, 0.57]	0.88 [0.72, 1.07]
EXT05	YANK-SMIRNOFF-tip3p-wet-oct	1.99 [1.38, 2.62]	1.56 [1.13, 2.07]	0.27 [-0.46, 1.02]	0.26 [0.01, 0.57]	1.24 [0.23, 1.92]	0.30 [0.02, 0.57]	0.90 [0.71, 1.08]
EXT01	YANK-GAFF-TIP3P-FB-wet-oct	2.05 [1.36, 2.66]	1.57 [1.11, 2.11]	-0.60 [-1.35, 0.11]	0.22 [0.00, 0.58]	1.11 [-0.15, 1.79]	0.24 [-0.07, 0.53]	0.93 [0.75, 1.11]
EXT04	YANK-SMIRNOFF-TIP3P-FB-wet-oct	2.07 [1.52, 2.59]	1.69 [1.26, 2.17]	-0.55 [-1.33, 0.18]	0.22 [0.00, 0.58]	1.14 [-0.05, 1.90]	0.26 [-0.06, 0.56]	0.84 [0.66, 1.02]
EXT06	YANK-SMIRNOFF-OPC-wet-oct	2.54 [1.90, 3.18]	2.11 [1.62, 2.66]	-2.11 [-2.66, -1.62]	0.48 [0.12, 0.83]	1.40 [0.85, 1.82]	0.49 [0.22, 0.73]	0.69 [0.49, 0.89]
EXT03	YANK-GAFF-opc-wet-oct	2.85 [2.06, 3.65]	2.37 [1.84, 3.03]	-2.37 [-3.02, -1.84]	$0.43 \ [0.07, \ 0.85]$	1.41 [0.68, 1.93]	$0.56 \ [0.29, \ 0.78]$	0.61 [0.45, 0.78]

## Notes

- RMSE: Root mean square error

- MAE: Mean absolute error

- ME: Mean error

- R2: R-squared, square of Pearson correlation coefficient

- m: slope of the line fit to predicted vs experimental logP values

-  $\tau$ : Kendall rank correlation coefficient

- ES: error slope calculated from the QQ Plots of model uncertainty predictions

- Mean and 95% confidence intervals of RMSE, MAE, ME, R2, and m were calculated by bootstrapping with 10000 samples.

- 95% confidence intervals of ES were calculated by bootstrapping with 1000 samples.