

Supporting Information for:
Bayesian Multilevel Modeling of Retention Data
Informed by Structural Similarity of Analytes

Paweł Wiczling¹

¹Department of Biopharmaceutics and Pharmacodynamics, Medical University of
Gdańsk, Gen. J. Hallera 107, 80-416 Gdańsk, Poland

2025-07-31

Table of contents

1	Table S1. Summary of the MCMC simulations of the marginal posterior distributions of population-level parameters	2
2	Figure S1. Raw data	3
3	Figure S2. Functional group effects	4
4	Figure S3. Goodness of fit plots (GOF)	5
5	Figure S4. Individual parameters, R_i	6
6	Figure S5. Individual parameters, $\eta_{R,i}$	7
	Licenses	8
	Original Computing Environment	8

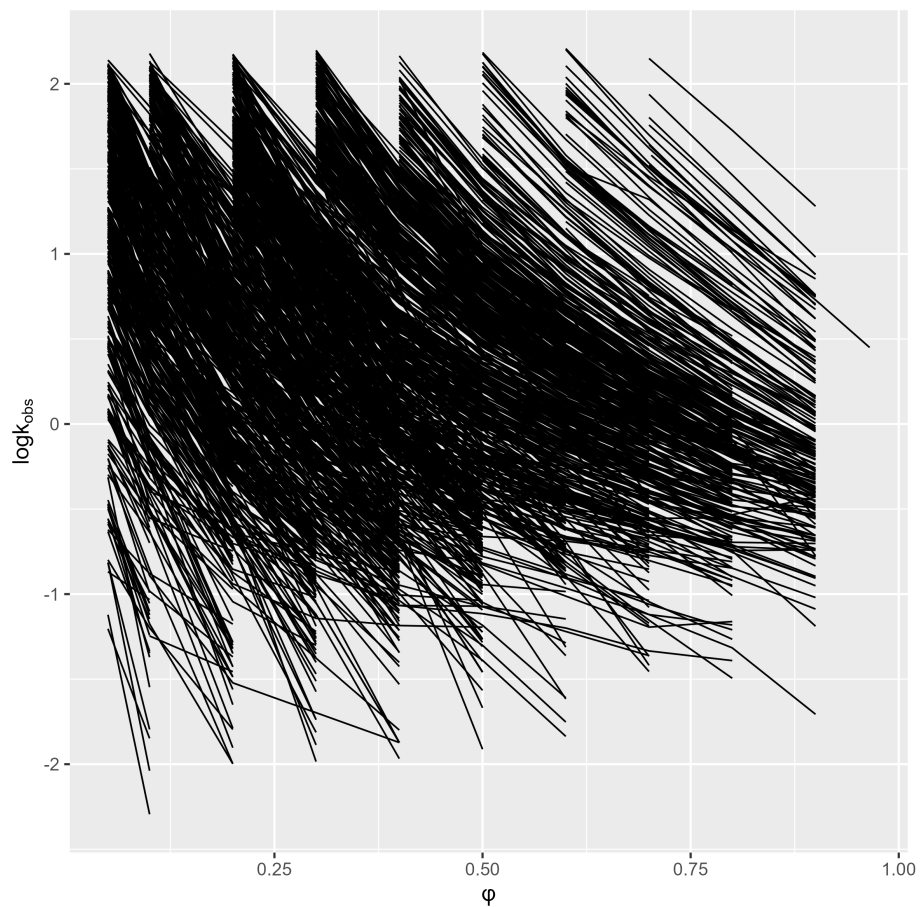
1 Table S1. Summary of the MCMC simulations of the marginal posterior distributions of population-level parameters

Mean denotes sample mean, MCSE denotes Monte Carlo Standard Error, Std-Dev denotes sample standard deviation, 5%, 50%, 95% denote corresponding quantiles, N_Eff denotes effective sample size, R_Hat denotes a measure of chain equilibrium must be within 0.05 of 1.0.

variable	mean	median	sd	mad	q5	q95	rhat	ess_bulk	ess_tail
logkwHat	3.05	3.05	0.19	0.19	2.74	3.36	1.00	2926	3497
S1Hat	3.67	3.67	0.16	0.16	3.40	3.94	1.00	2598	3029
dlogkwHat	-0.35	-0.35	0.12	0.12	-0.55	-0.15	1.00	4163	3537
dS1Hat	0.30	0.30	0.09	0.09	0.16	0.45	1.00	3403	3506
beta[1]	0.86	0.86	0.05	0.06	0.77	0.95	1.00	3235	3620
beta[2]	0.62	0.62	0.05	0.05	0.54	0.70	1.00	3131	3035
S2Hat	2.90	2.90	0.03	0.03	2.85	2.95	1.02	430	935
omega[1]	1.41	1.41	0.04	0.04	1.35	1.47	1.00	2650	2471
omega[2]	1.13	1.13	0.03	0.03	1.08	1.18	1.00	2317	3017
rho[1,2]	0.88	0.88	0.01	0.01	0.87	0.90	1.00	1933	2928
sigma	0.05	0.05	0.00	0.00	0.05	0.05	1.00	1880	2927
alpha	0.71	0.71	0.04	0.04	0.64	0.77	1.00	5245	3475
sdpi[1]	0.15	0.14	0.02	0.02	0.11	0.18	1.01	655	1162
sdpi[2]	0.13	0.13	0.02	0.02	0.10	0.17	1.01	641	1372
sdpi[3]	0.34	0.33	0.06	0.06	0.23	0.44	1.00	2816	3270
sdpi[4]	0.19	0.18	0.07	0.07	0.08	0.30	1.01	939	779

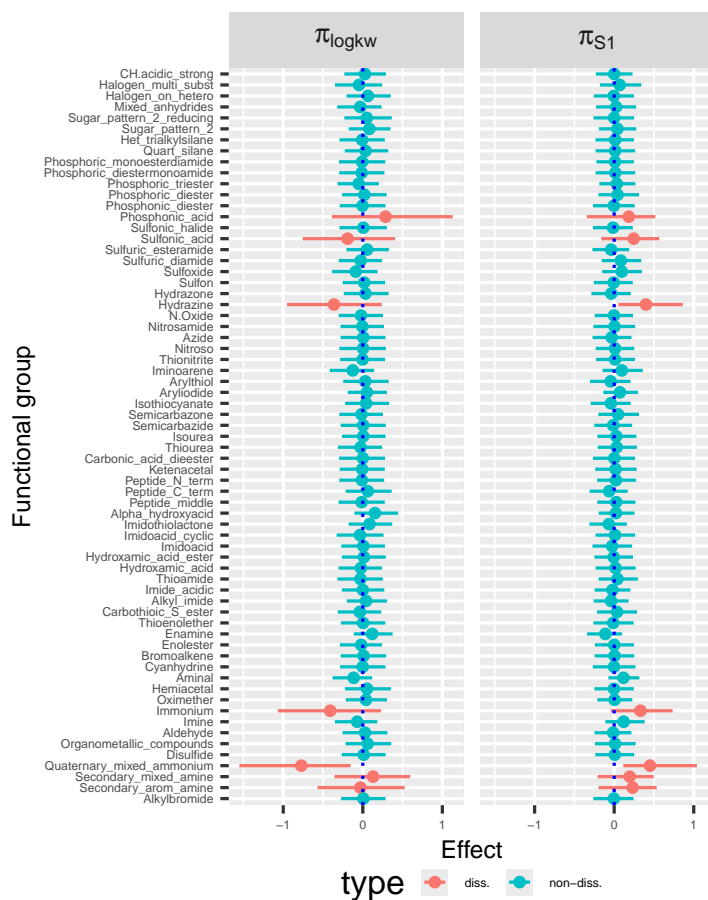
2 Figure S1. Raw data

The relationship between observed logk and organic modifier content for all the analytes.



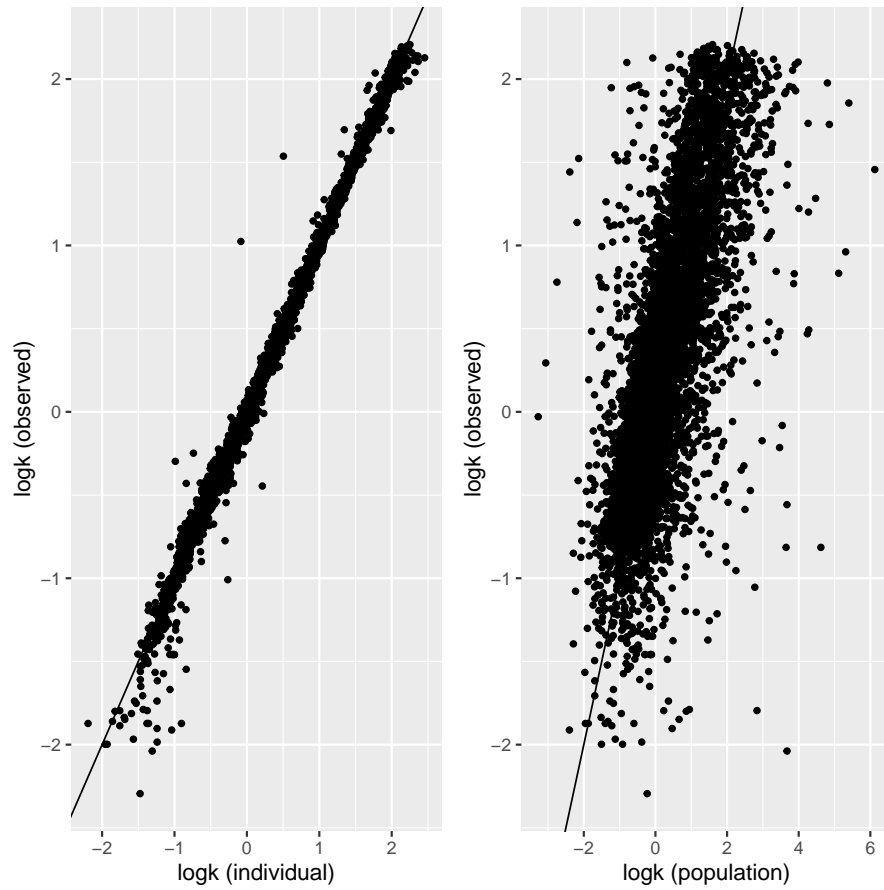
3 Figure S2. Functional group effects

A summary of the marginal posterior distributions of the functional group effects. Red represents the effect of dissociated functional groups, and green represents the effect of non-dissociated functional groups. The least common functional groups (≤ 10 across all analytes) are presented. The plot for the remaining functional groups is presented in Figure 2.



4 Figure S3. Goodness of fit plots (GOF)

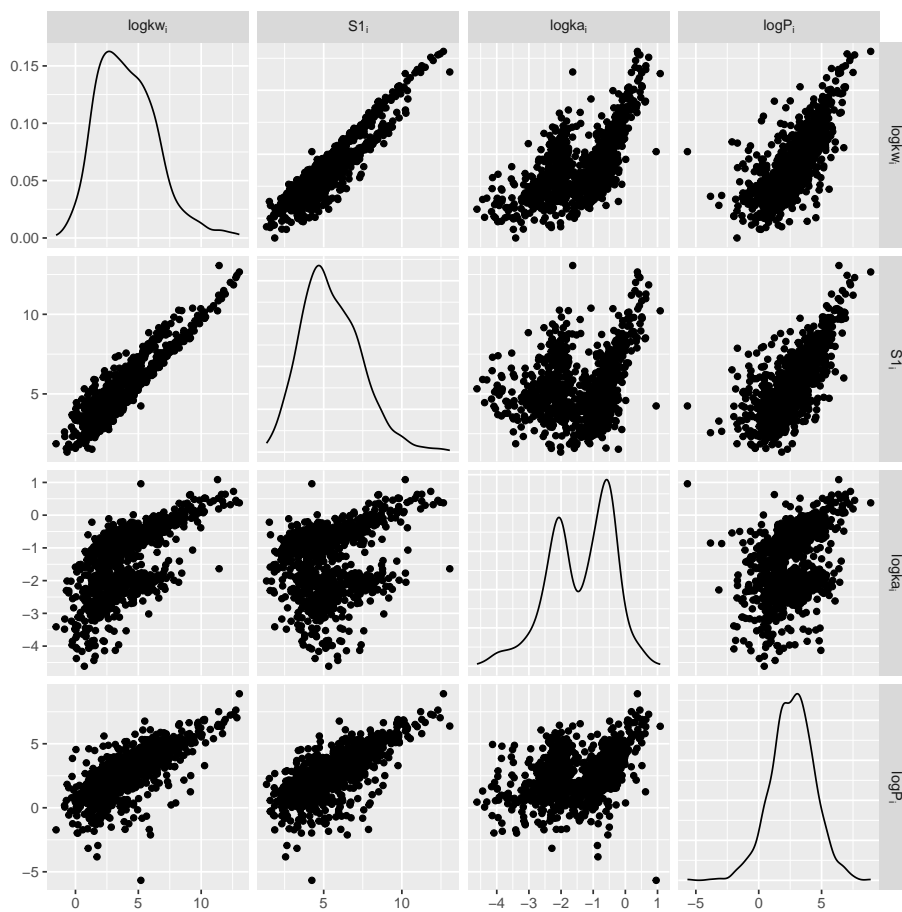
Observed vs. population and individual predictions:



5 Figure S4. Individual parameters, R_i

$\log k_a$ is the (extrapolated) log retention in 100% ACN:

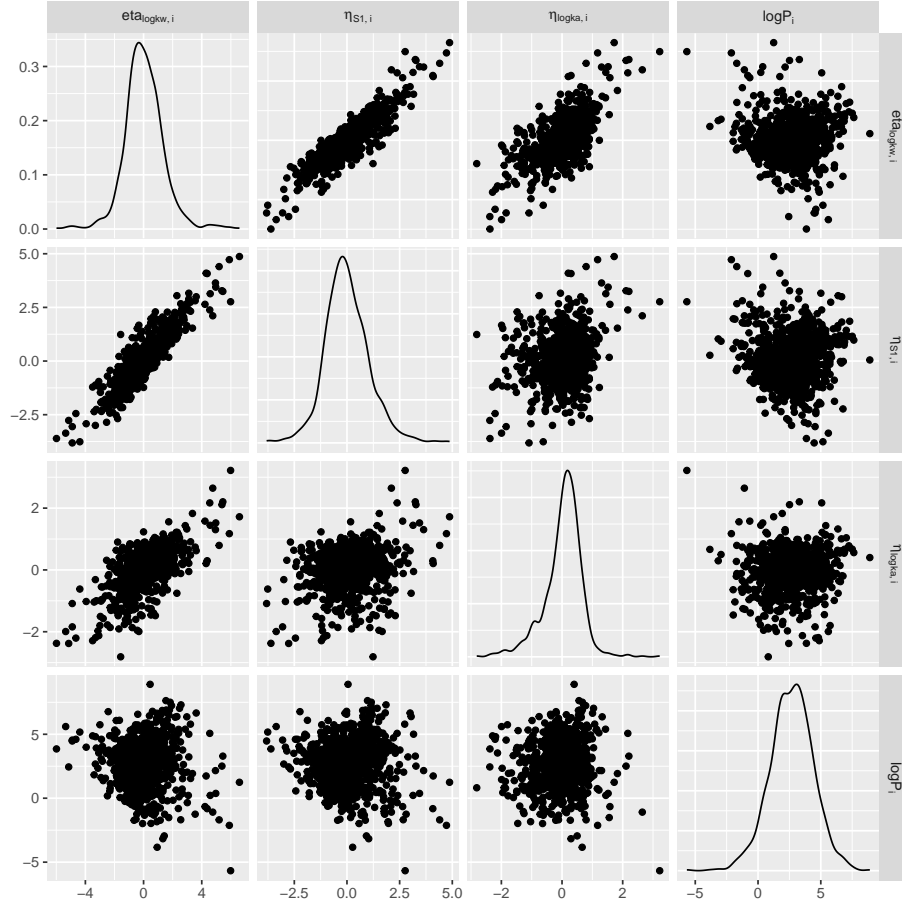
$$\log k_{a_i} = \log k_{w_i} - S_{1,i}$$



6 Figure S5. Individual parameters, $\text{eta}_{R,i}$

$\text{eta}_{R,i}$ represents the deviation of an individual analyte's parameter from its expected (population-level) value.

$$\eta_{R,i} = R_i - (\theta_R + \beta_R \cdot \log P_i + \pi_{R1} \cdot X_{1i} + \pi_{R2} \cdot X_{2i})$$



Licenses

- Code & copy; 2023, Paweł Wiczling, licensed under BSD-3.
- Text & copy; 2023, Paweł Wiczling, licensed under CC-BY-NC 4.0.

Original Computing Environment

R version 4.3.1 (2023-06-16 ucrt)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19045)

Matrix products: default

locale:

```
[1] LC_COLLATE=Polish_Poland.utf8 LC_CTYPE=Polish_Poland.utf8
[3] LC_MONETARY=Polish_Poland.utf8 LC_NUMERIC=C
[5] LC_TIME=Polish_Poland.utf8
```

time zone: Europe/Warsaw

tzcode source: internal

attached base packages:

```
[1] stats      graphics  grDevices  utils      datasets  methods    base
```

other attached packages:

```
[1] cmdstanr_0.8.0.9000 kableExtra_1.4.0    dplyr_1.1.4
```

loaded via a namespace (and not attached):

```
[1] jsonlite_1.8.8      compiler_4.3.1      tidyselect_1.2.1
[4] xml2_1.3.6          stringr_1.5.1       systemfonts_1.1.0
[7] scales_1.3.0        textshaping_0.4.0   yaml_2.3.8
[10] fastmap_1.2.0       R6_2.5.1            generics_0.1.3
[13] distributional_0.4.0 knitr_1.46          backports_1.5.0
[16] checkmate_2.3.1     tibble_3.2.1        munsell_0.5.1
[19] svglite_2.1.3       pillar_1.9.0        posterior_1.6.0
[22] rlang_1.1.3         utf8_1.2.4          stringi_1.8.4
[25] xfun_0.44           viridisLite_0.4.2   cli_3.6.2
[28] withr_3.0.2         magrittr_2.0.3      ps_1.7.6
[31] processx_3.8.4      digest_0.6.35       rstudioapi_0.16.0
[34] lifecycle_1.0.4     vctrs_0.6.5         data.table_1.15.4
[37] tensorA_0.36.2.1    evaluate_1.0.3      glue_1.7.0
[40] codetools_0.2-19    ragg_1.3.2          abind_1.4-8
[43] fansi_1.0.6         colorspace_2.1-0    rmarkdown_2.27
[46] matrixStats_1.3.0   tools_4.3.1         pkgconfig_2.0.3
```


[49] `htmltools_0.5.8.1`