

MOSAIC_{bioacc} REPORT

2021-04-15

This report is provided by the MOSAIC_{bioacc} application available here:
<https://mosaic.univ-lyon1.fr/bioacc>

Contact: sandrine.charles@univ-lyon1.fr

MOSAIC_{bioacc} uses the JAGS (version 4.3.0) and R (version 4.0.2) software, and in particular packages RJags (version 4.10), jagsUI (version 1.5.1) and Shiny (version 1.6.0).

The MOSAIC_{bioacc} application is a turn-key web tool providing bioaccumulation factors (BCF/BSAF/BMF) from a toxicokinetic (TK) model fitted to accumulation-depuration data. It is designed to fulfil the requirements of regulators when examining applications for market authorization of active substances.

Data summary

File used: Gammarus_PRZ_1d_Fu2018.txt

Exposure: 100 $\mu\text{g.mL}^{-1}$

Accumulation phase duration: 1 days

Number of replicates: 2

Times: 0, 0.021, 0.063, 0.104, 0.229, 0.396, 0.729, 1, 1.042, 1.083, 1.167, 1.292, 1.458, 1.75, 2.083, 2.708, 3.958, 4.958, 6

Exposure routes: water

Elimination routes: excretion biotransformation

Bayesian inference

Three MCMC chains were used to estimate model parameters.

Number of iterations: 89904

Thin: 24

TK Model

The TK model used for these calculations was:

$$\frac{dC_p(t)}{dt} = k_{uw} \times c_w - (k_{ee} + k_{m1} + k_{m2}) \times C_p(t) \quad \text{for } 0 \leq t \leq t_c$$

$$\frac{dC_p(t)}{dt} = -(k_{ee} + k_{m1} + k_{m2}) \times C_p(t) \quad \text{for } t > t_c$$

$$\frac{dC_{m1}(t)}{dt} = k_{m1} \times C_p(t) - k_{em1} \times C_{m1}(t)$$

$$\frac{dC_{m2}(t)}{dt} = k_{m2} \times C_p(t) - k_{em2} \times C_{m2}(t)$$

with:

t : time (expressed in days)

t_c : duration of the accumulation phase (expressed in days)

$C_p(t)$: internal concentration of the parent compound at time (expressed in $\mu g \cdot g^{-1}$)

k_{ee} : elimination rates of excretion (expressed per days $^{-1}$)

c_w : exposure concentration of water route (expressed in $\mu g \cdot mL^{-1}$)

k_{uw} : uptake rate of water exposure (expressed per days $^{-1}$)

$C_{m\ell}(t)$: internal concentration of metabolite ℓ (expressed in $\mu g \cdot g^{-1}$)

ℓ : index of metabolites, $\ell = 1 \dots L$ with L total number of metabolites

$k_{m\ell}$: metabolization rate of metabolite ℓ (expressed per days $^{-1}$)

$k_{em\ell}$: elimination rates of metabolite ℓ (expressed per days $^{-1}$)

Bioaccumulation factor calculation

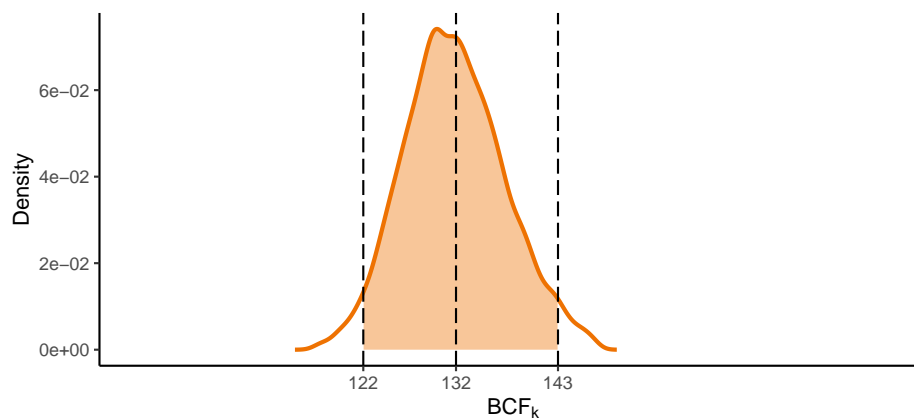
Calculations

$$BCF_k = \frac{k_{uw}}{k_{ee} + k_{m1} + k_{m2}}$$

$$BCF_{ss} = \frac{C_p(t_c)}{c_w}$$

Bioconcentration factor (BCF)

BCF_k plot

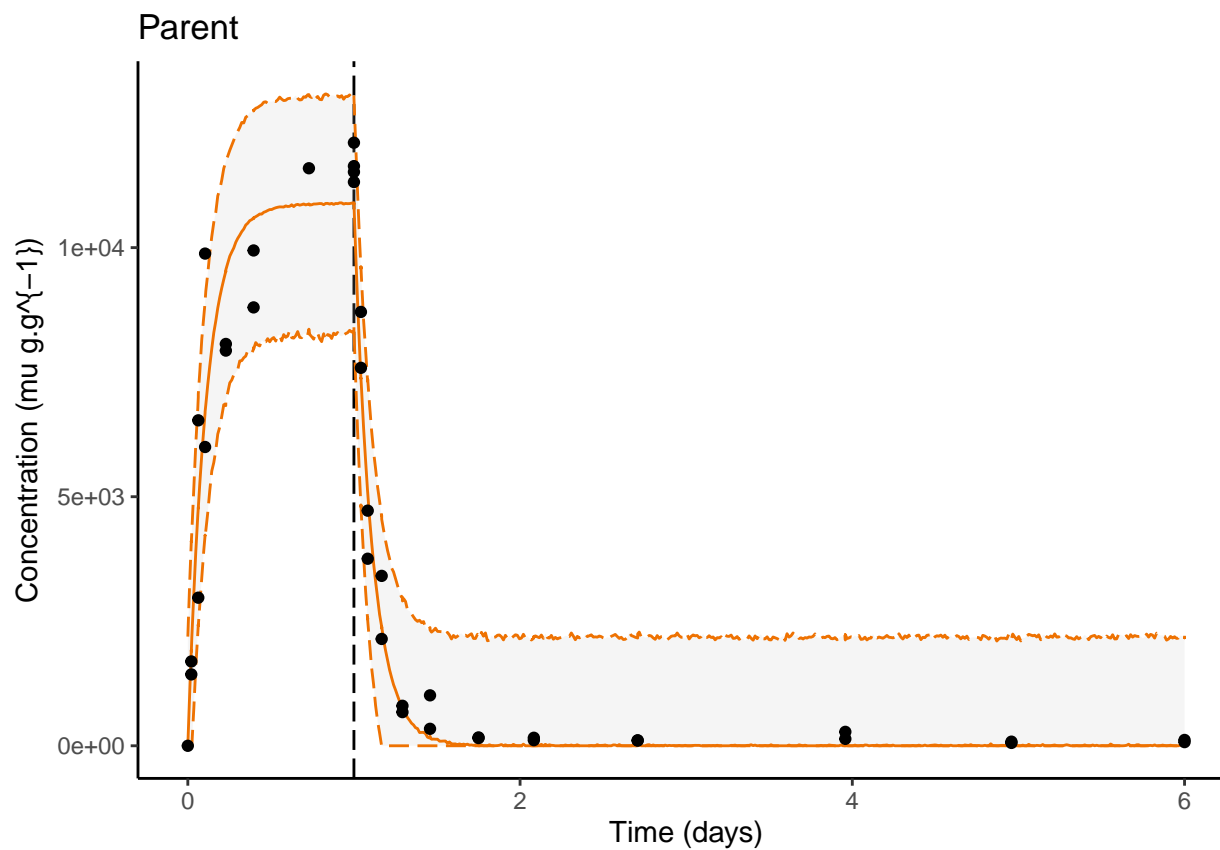


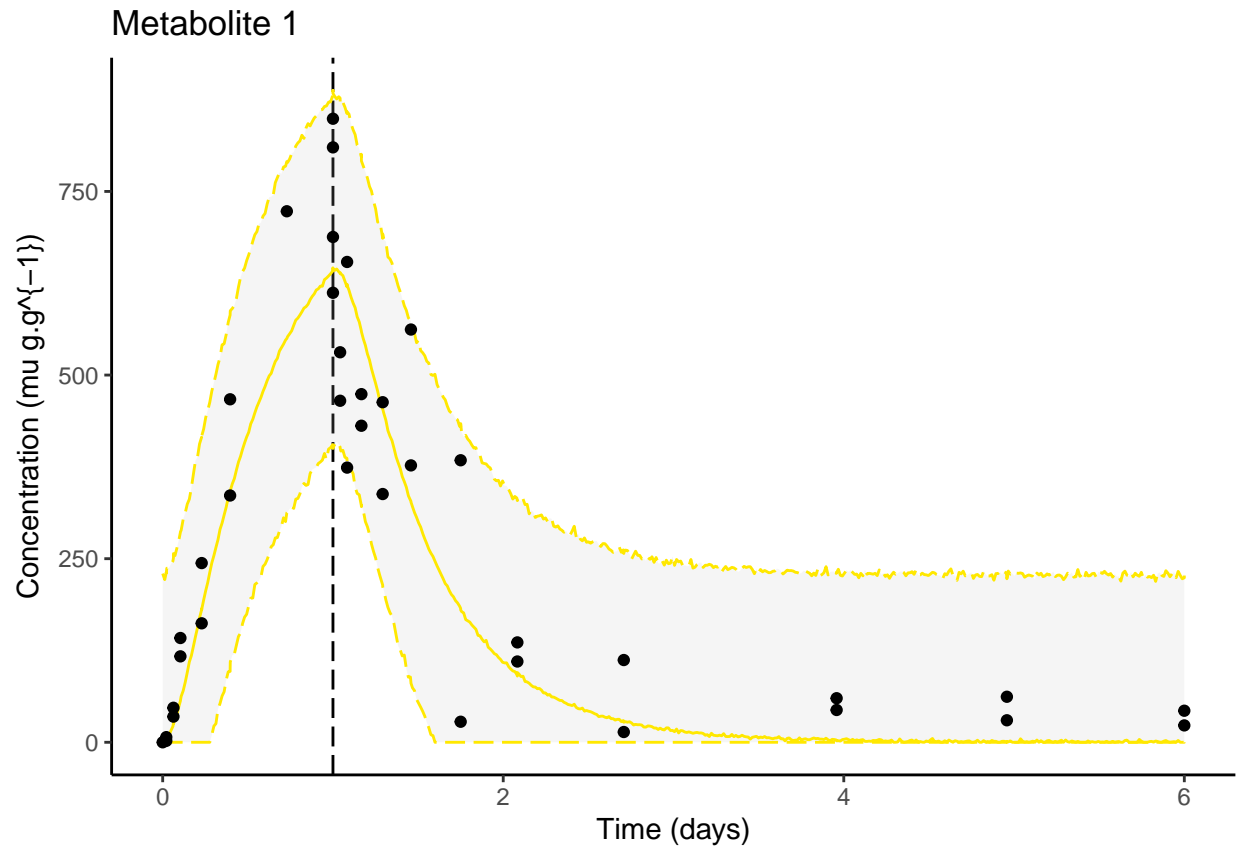
BCF summary

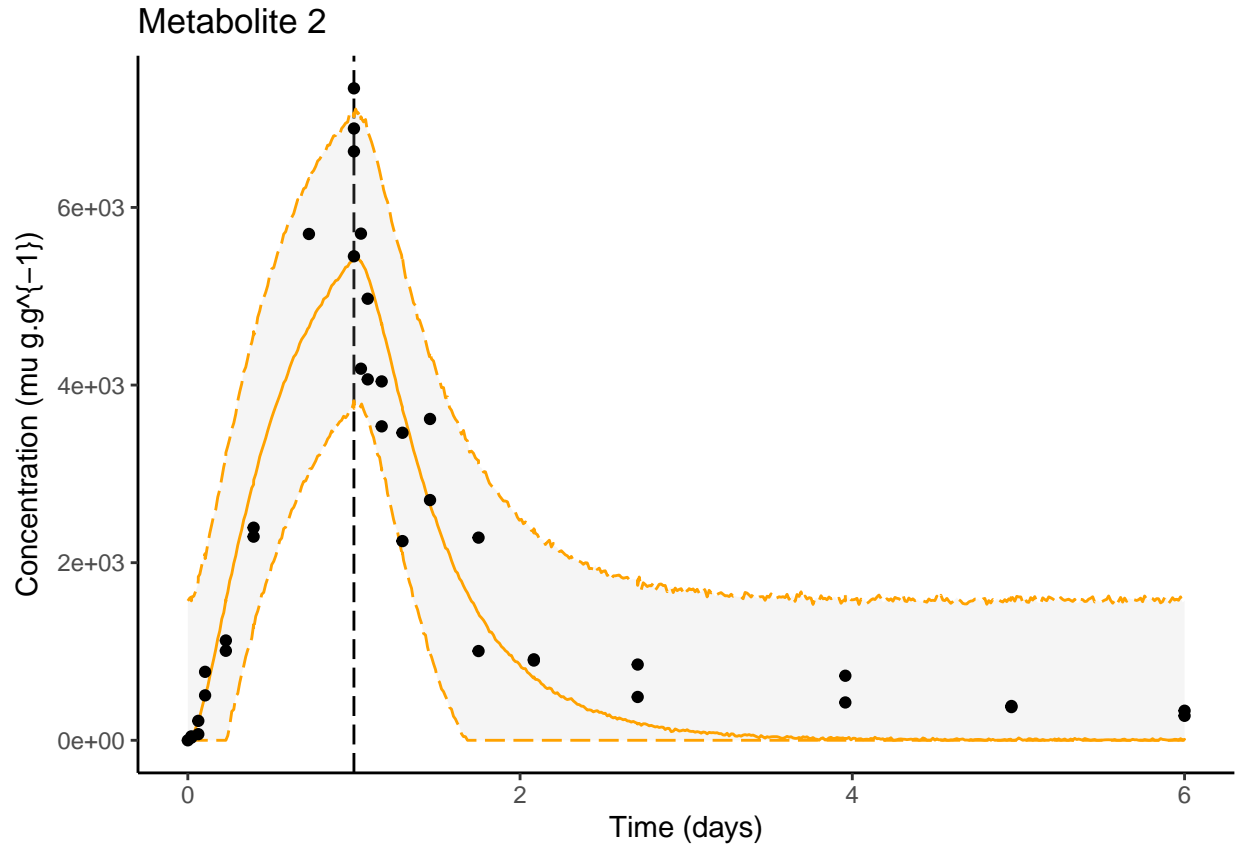
	2.5%	50%	97.5%	CV
BCFk	122	132	143	4e-02

Fitting results

Fit plot







Quantiles of estimated parameters

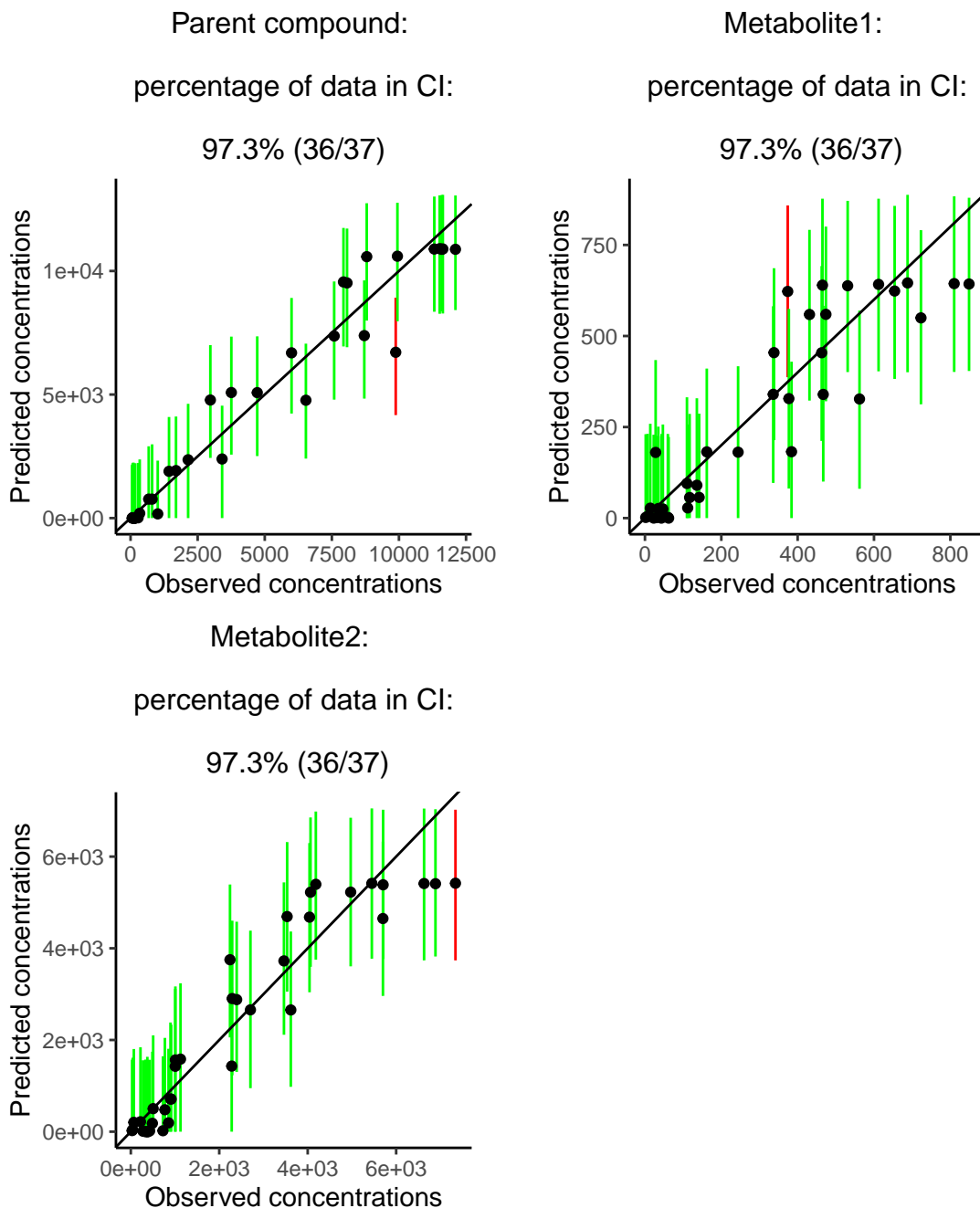
	2.5%	50%	97.5%	
k_{uw}	804.6	999.1	1198	d^{-1}
k_{ee}	5.951	7.723	9.817	d^{-1}
k_{m1}	0.1078	0.1479	0.2343	d^{-1}
k_{m2}	0.9976	1.286	1.858	d^{-1}
k_{em1}	1.357	2.093	3.173	d^{-1}
k_{em2}	1.531	2.196	3.085	d^{-1}
σ_p	820.1	1024	1441	$\mu g.g^{-1}$
σ_{met1}	90.59	113	146.9	$\mu g.g^{-1}$
σ_{met2}	629.5	783.4	1019	$\mu g.g^{-1}$

Goodness-of-fit criteria

Posterior Predictive Check

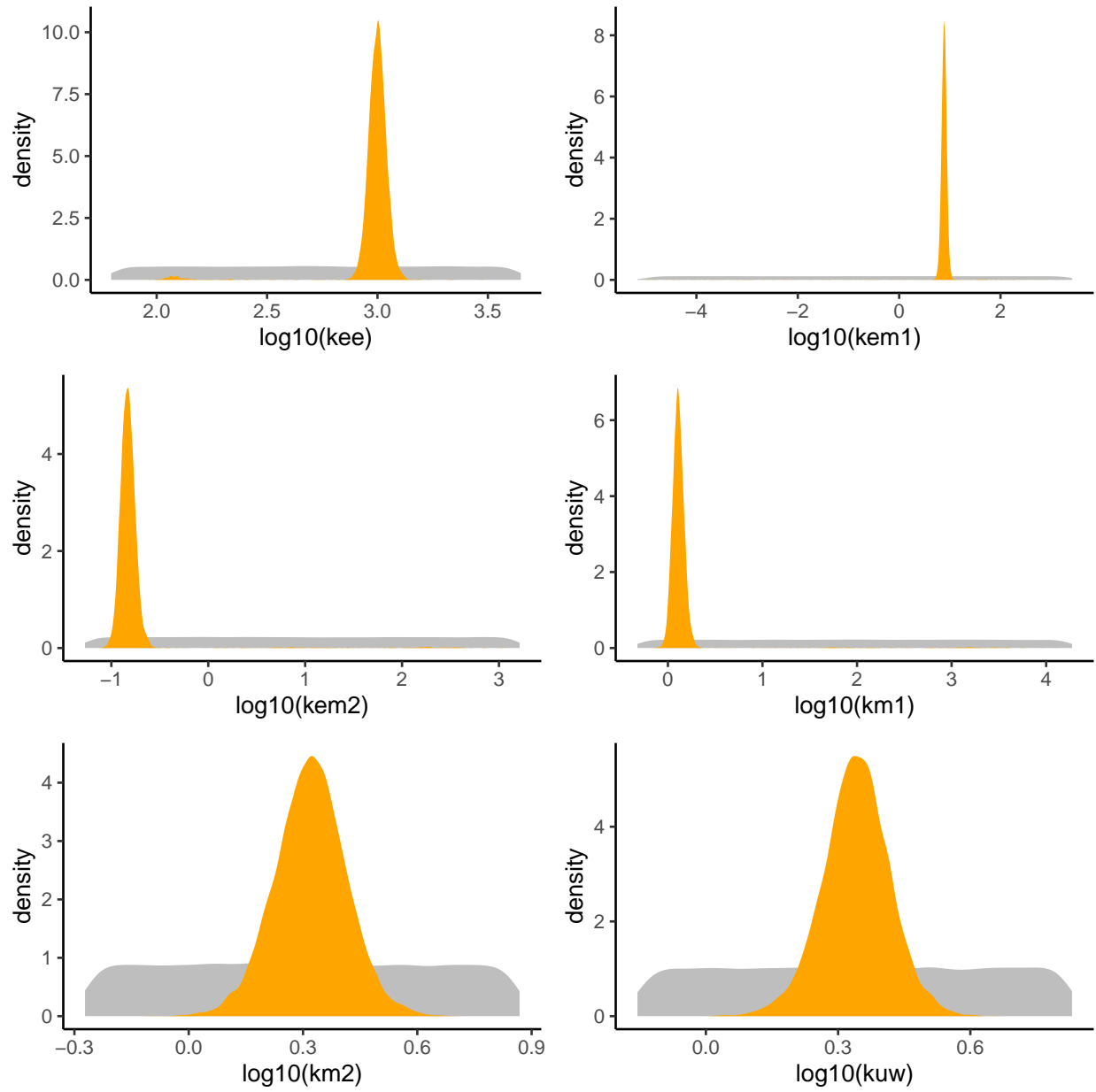
The PPC shows the observed values against their corresponding estimated predictions (black dots), along with their 95% credible interval (vertical segments). If the fit is correct, we expect to see 95% of the data within the intervals. Ideally observations and predictions should coincide, so we would expect to see black

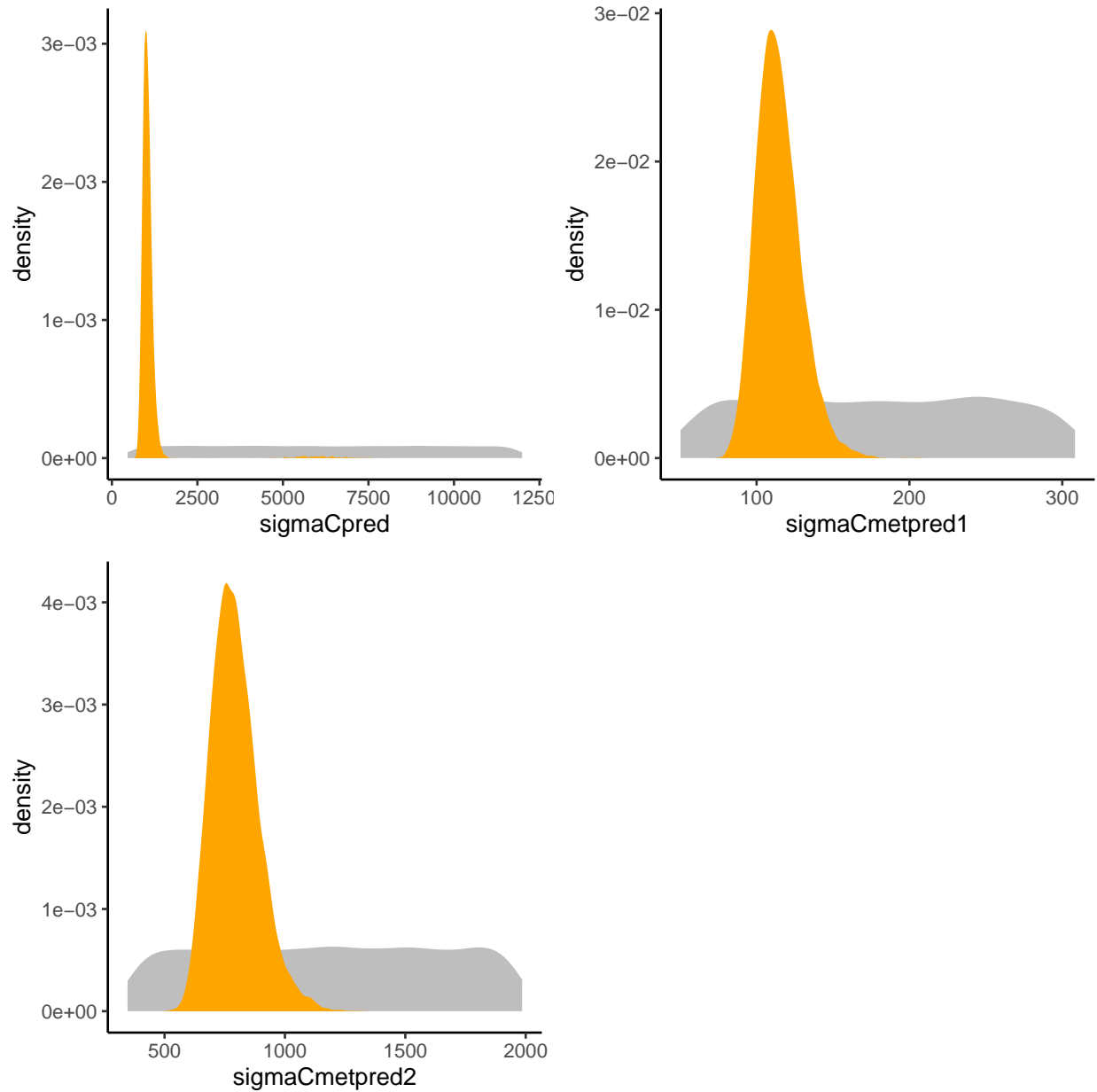
dots along the first bisector $y = x$ (plain black line). The 95% credible intervals are colored in green if they overlap this line, in red otherwise.



Priors and posteriors

The prior distribution is represented by the gray area and the posterior distribution by the orange area. The accuracy of the model parameter estimation can be visualized by comparing prior and posterior distributions: the overall expectation is to get a narrower posterior distribution compared to the prior one, what reflects that data contributed enough to precisely estimate parameters.

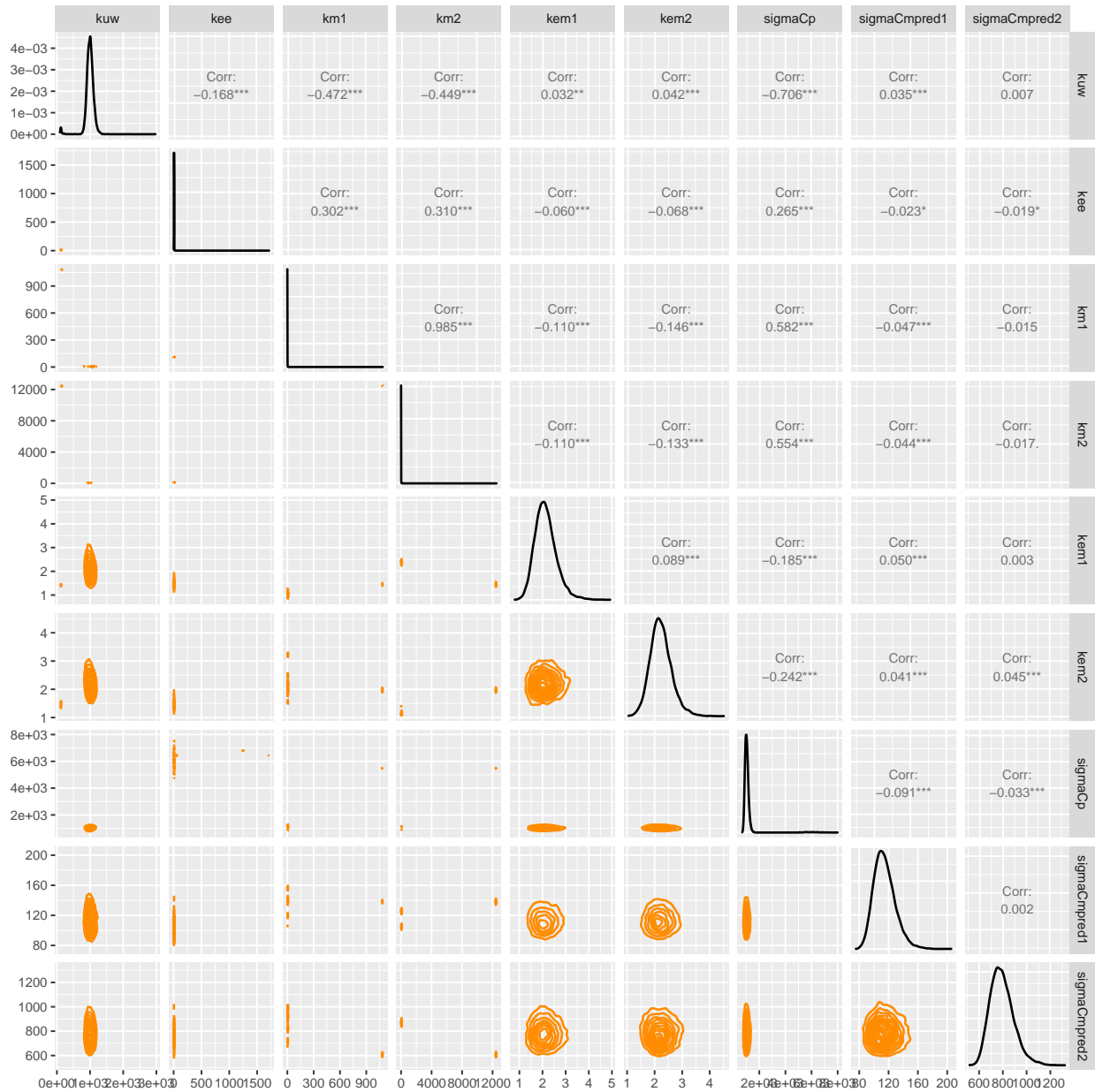




Correlation between parameters

If you want to see the coloured matrix giving a summary of parameter correlations, you need to import the corresponding figure directly from the application, page bottom, section “Downloads”, then choose Download an output and select “GOF” then “parameter correlation”. You can select the output format you prefer.

Correlations between parameters are visualized by projecting the joint posterior distribution in a plot matrix with planes of parameter pairs (lower triangular elements), marginal posterior distribution of each model parameter (diagonal), and Pearson correlation coefficients (upper triangular elements). Correlations are expected to be low (reflected by “potatoid” shapes of density lines in orange); a leaning elliptical shape translates high correlations (positive if leaning to the right, negative if leaning to the left).



Potential Scale Reduction Factors

Convergence of the MCMC chains can be checked with the Gelman-Rubin diagnostic expressed with the potential scale reduction factor (PSRF). Approximate convergence is diagnosed when the PSRF is below 1.01.

	PSRF
kuw	1.167
kee	1.293
km1	1.302
km2	1.301
kem1	1.001
kem2	1.005
sigmaCpred	1.309
sigmaCmetpred1	1

PSRF	
sigmaCmetpred2	1

Watanabe–Akaike information criterion

Information criteria offer a computationally appealing way of estimating the generalization performance of the model. A fully Bayesian criterion is the widely applicable information criterion (WAIC) by Watanabe a penalized deviance statistics accounting for the uncertainty in the parameters and can be used also for singular models. WAIC is widely used in model comparison for a same dataset (e.g., with or without k_{ee}). Sub-models with lower WAIC values will be preferred.

$$\text{WAIC} = 1683$$

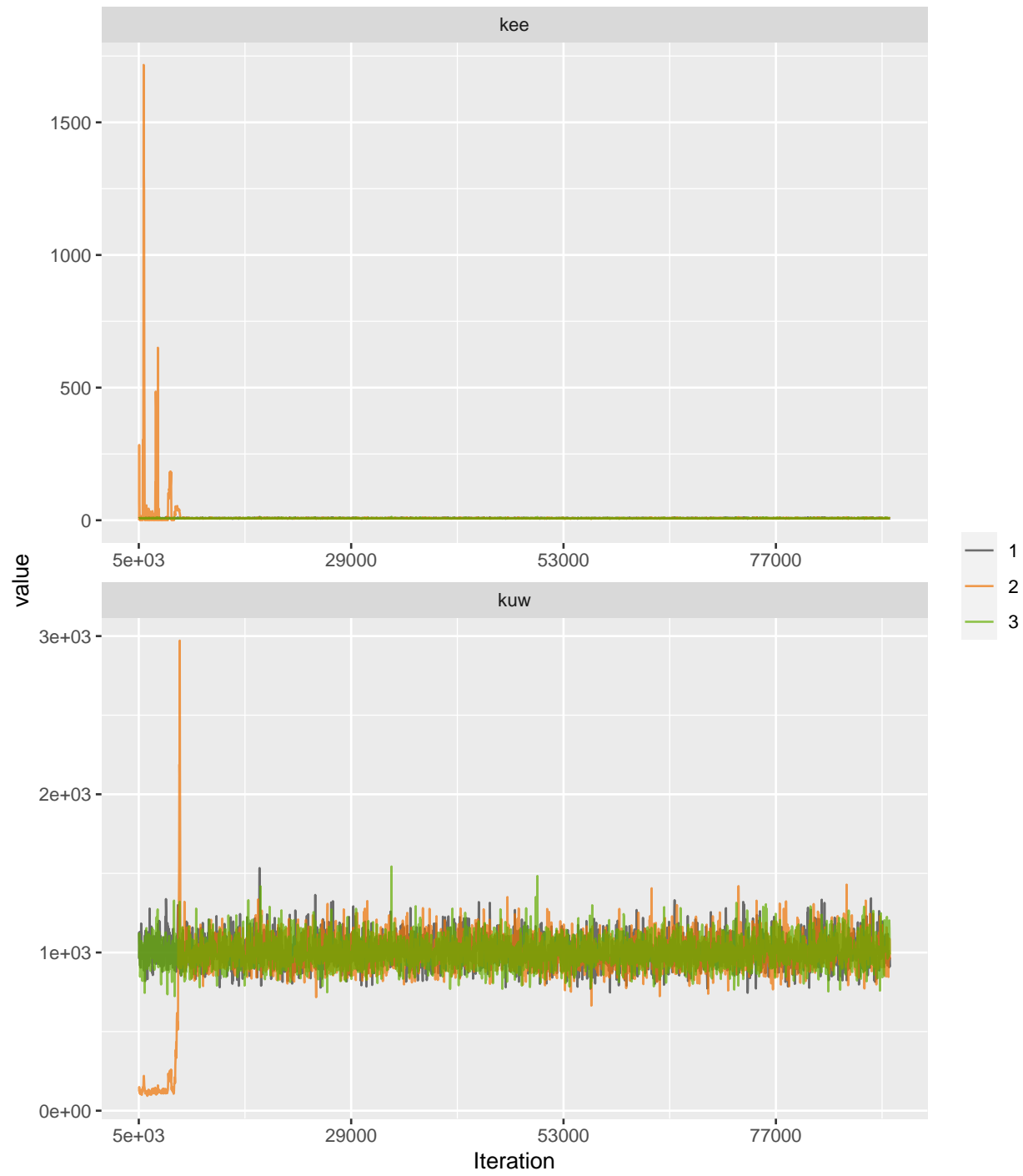
Deviance Information Criterion

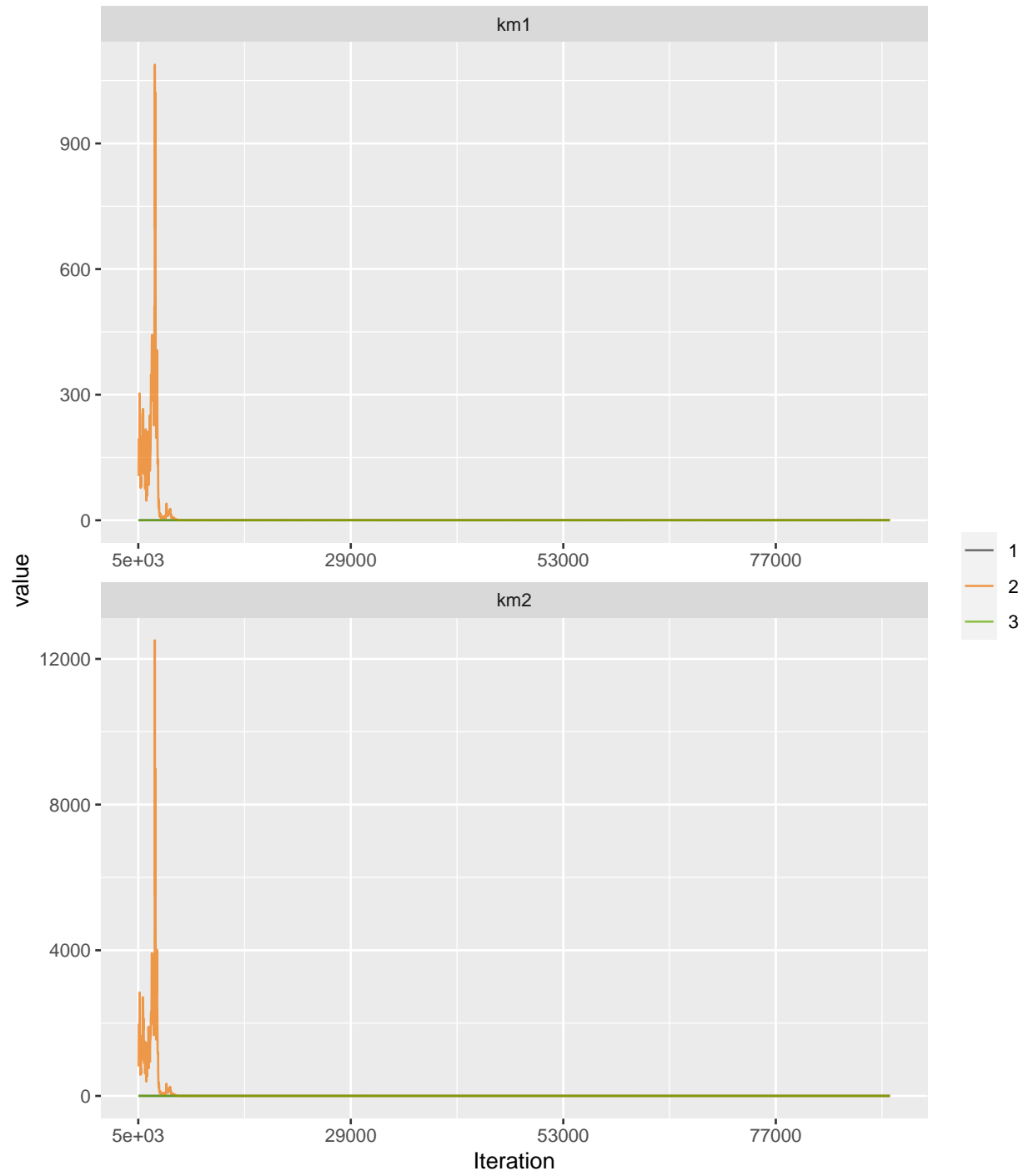
This criteria, denoted DIC, is a penalized deviance statistics accounting for the number of parameters for use in model comparison for a same dataset (e.g., with or without k_{ee}). Sub-models with lower DIC values will be preferred.

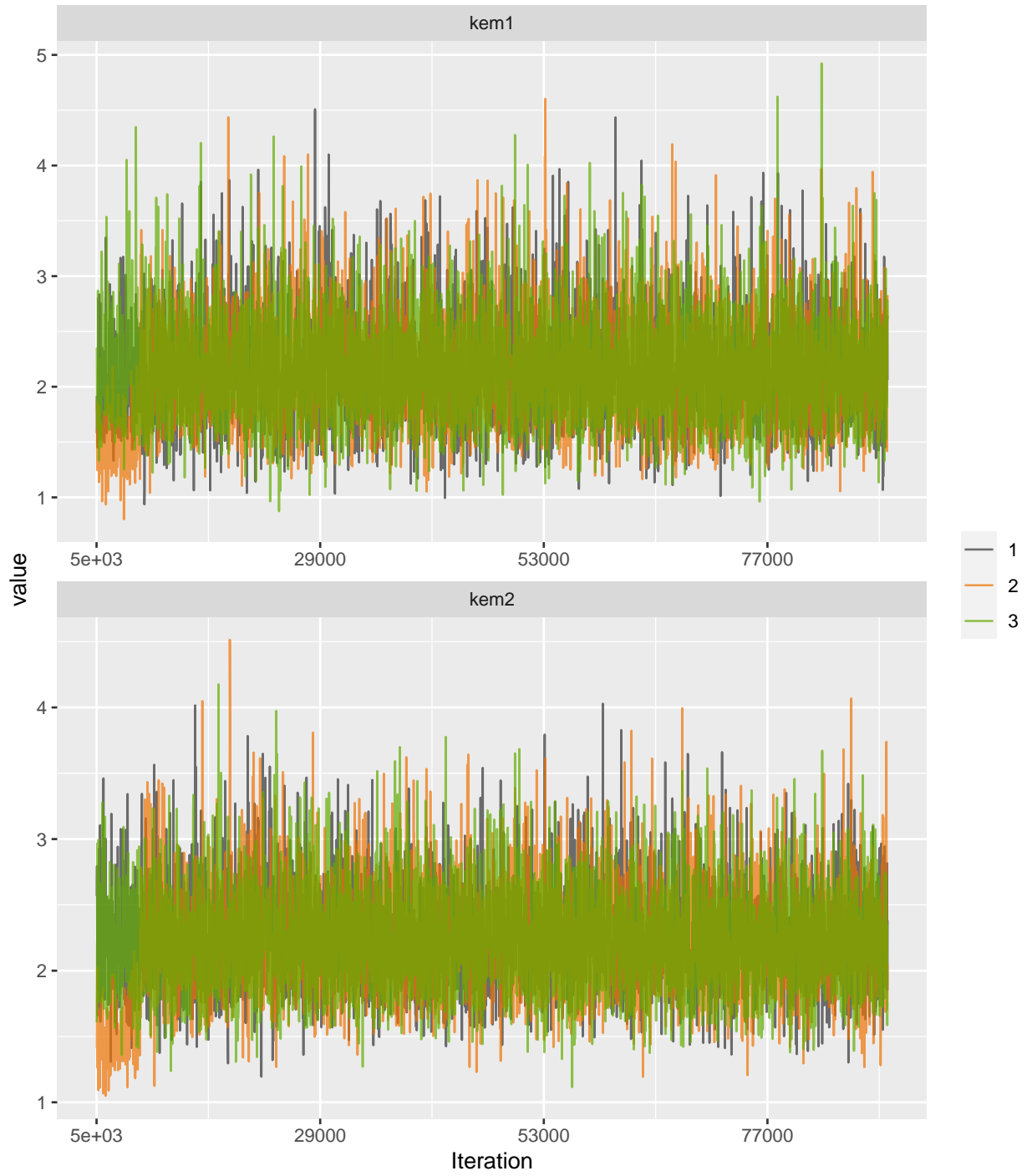
$$\text{DIC} = 1807$$

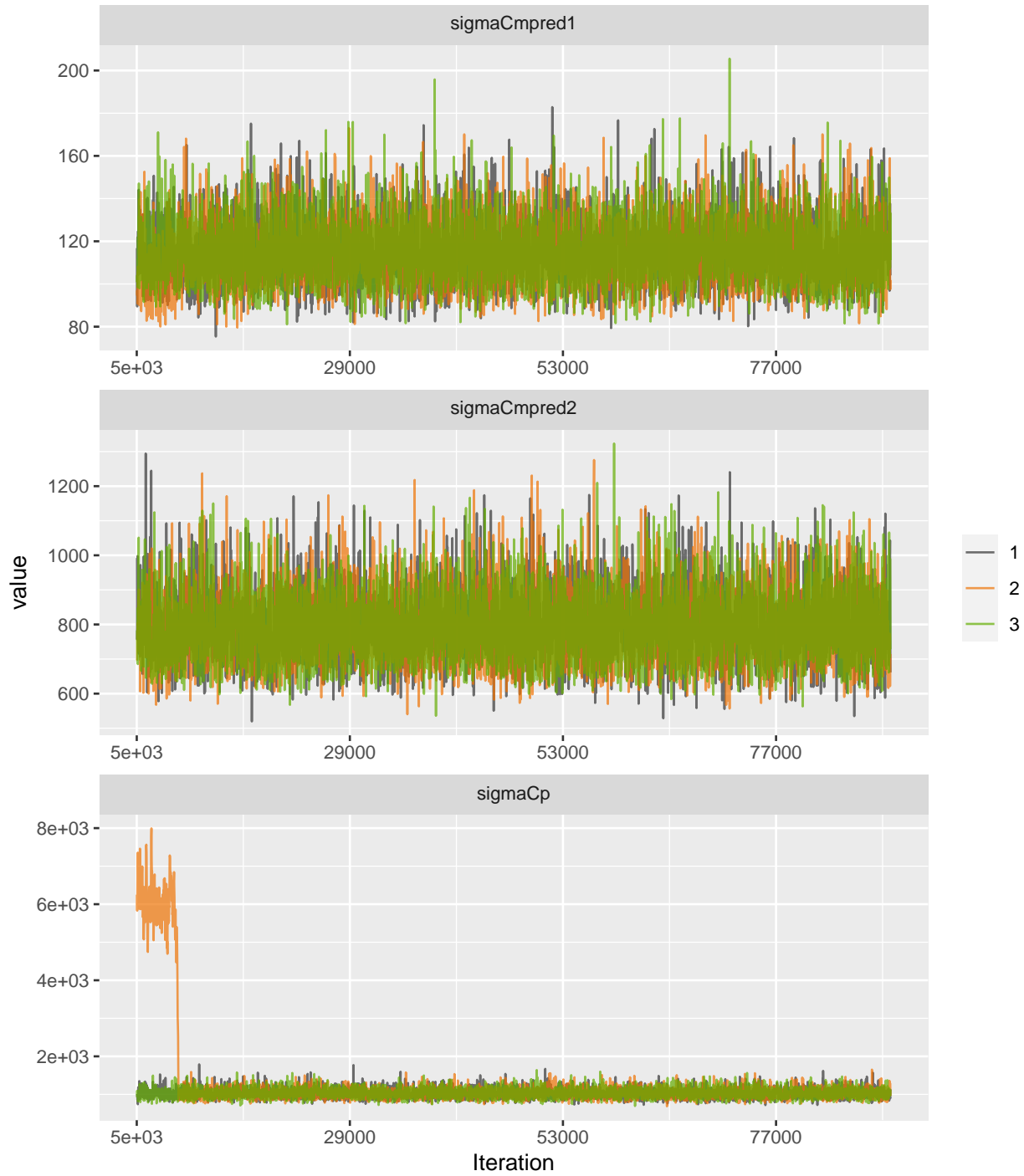
Traces of MCMC iterations

A traceplot is an essential plot for assessing convergence and diagnosing of MCMC chains. It shows the time series of the sampling process leading to the posterior distribution. Different colors are used for each of the chains (here 3) to assess within-chain convergence.









Data Table

time	expw	conc	concm1	concm2	replicate
0.000	100	0	0	0	1
0.021	100	1693	7	44	1
0.021	100	1433	2	30	2
0.063	100	2973	35	69	1
0.063	100	6533	47	219	2
0.104	100	5998	117	506	1
0.104	100	9880	142	771	2
0.229	100	8066	162	1008	1
0.229	100	7933	244	1124	2
0.396	100	9943	336	2394	1
0.396	100	8800	467	2293	2
0.729	100	11593	723	5701	2
1.000	100	11519	612	5450	1
1.000	100	11636	849	6889	2
1.000	100	12106	688	7341	1
1.000	100	11317	810	6631	2
1.042	100	7587	465	4184	1
1.042	100	8709	531	5705	2
1.083	100	3755	654	4972	1
1.083	100	4720	374	4064	2
1.167	100	2145	474	4041	1
1.167	100	3411	431	3535	2
1.292	100	803	338	2243	1
1.292	100	676	463	3463	2
1.458	100	341	377	2705	1
1.458	100	1013	562	3618	2
1.750	100	168	384	2282	1
1.750	100	155	28	1005	2
2.083	100	113	136	912	1
2.083	100	161	110	898	2
2.708	100	110	112	853	1
2.708	100	106	14	488	2
3.958	100	139	44	426	1
3.958	100	280	60	727	2
4.958	100	83	30	374	1
4.958	100	58	62	386	2
6.000	100	117	43	276	1
6.000	100	74	23	333	2