Testing covariate modelling in hierarchical parent degradation kinetics with residue data on mesotrione

Johannes Ranke

Last change on 4 August 2023, last compiled on 8 August 2023

Contents

Introduction

The purpose of this document is to test demonstrate how nonlinear hierarchical models (NLHM) based on the parent degradation models SFO, FOMC, DFOP and HS can be fitted with the mkin package, also considering the influence of covariates like soil pH on different degradation parameters. Because in some other case studies, the SFORB parameterisation of biexponential decline has shown some advantages over the DFOP parameterisation, SFORB was included in the list of tested models as well.

The mkin package is used in version 1.2.5, which is contains the functions that were used for the evaluations. The `saemix` package is used as a backend for fitting the NLHM, but is also loaded to make the convergence plot function available.

This document is processed with the `knitr` package, which also provides the `kable` function that is used to improve the display of tabular data in R markdown documents. For parallel processing, the `parallel` package is used.

```
library(mkin)
library(knitr)
library(saemix)
library(parallel)
n_cores <- detectCores()
if (Sys.info()["sysname"] == "Windows") {
  cl <- makePSOCKcluster(n_cores)
} else {
  cl <- makeForkCluster(n_cores)
}
```

Test data

```
data_file <- system.file(
  "testdata", "mesotrione_soil_efsa_2016.xlsx", package = "mkin")
meso_ds <- read_spreadsheet(data_file, parent_only = TRUE)
```

The following tables show the covariate data and the 18 datasets that were read in from the spreadsheet file.

```
pH <- attr(meso_ds, "covariates")
kable(pH, caption = "Covariate data")
```

Table 1: Covariate data

	pH
Richmond	6.2
Richmond 2	6.2
ERTC	6.4
Toulouse	7.7
Picket Piece	7.1
721	5.6
722	5.7
723	5.4
724	4.8
725	5.8
727	5.1
728	5.9
729	5.6
730	5.3
731	6.1
732	5.0
741	5.7
742	7.2

```
for (ds_name in names(meso_ds)) {
  print(
    kable(mkin_long_to_wide(meso_ds[[ds_name]]),
      caption = paste("Dataset", ds_name),
      booktabs = TRUE, row.names = FALSE))
}
```

Table 2: Dataset Richmond

time	meso
0.000000	91.00
1.179050	86.70
3.537149	73.60
7.074299	61.50
10.611448	55.70
15.327647	47.70
17.685747	39.50
24.760046	29.80
35.371494	19.60
68.384889	5.67
0.000000	97.90
1.179050	96.40
3.537149	89.10
7.074299	74.40
10.611448	57.40
15.327647	46.30
18.864797	35.50
27.118146	27.20
35.371494	19.10
74.280138	6.50
108.472582	3.40
142.665027	2.20

Table 3: Dataset Richmond 2

time	meso
0.000000	96.0
2.422004	82.4
5.651343	71.2
8.073348	53.1
11.302687	48.5
16.954030	33.4
22.605373	24.2
45.210746	11.9

Table 4: Dataset ERTC

time	meso
0.000000	99.9
2.755193	80.0
6.428782	42.1
9.183975	50.1

time	meso
12.857565	28.4
19.286347	39.8
25.715130	29.9
51.430259	2.5

Table 5: Dataset Toulouse

time	meso
0.000000	96.8
2.897983	63.3
6.761960	22.3
9.659942	16.6
13.523919	16.1
20.285879	17.2
27.047838	1.8

Table 6: Dataset Picket Piece

time	meso
0.000000	102.0
2.841195	73.7
6.629454	35.5
9.470649	31.8
13.258909	18.0
19.888364	3.7

Table 7: Dataset 721

time	meso
0.00000	86.4
11.24366	61.4
22.48733	49.8
33.73099	41.0
44.97466	35.1

Table 8: Dataset 722

time	meso
0.00000	90.3
11.24366	52.1
22.48733	37.4
33.73099	21.2
44.97466	14.3

Table 9: Dataset 723

time	meso
0.00000	89.3
11.24366	70.8
22.48733	51.1
33.73099	42.7
44.97466	26.7

Table 10: Dataset 724

time	meso
0.000000	89.4
9.008208	65.2
18.016415	55.8
27.024623	46.0
36.032831	41.7

Table 11: Dataset 725

time	meso
0.00000	89.0
10.99058	35.4
21.98116	18.6
32.97174	11.6
43.96232	7.6

Table 12: Dataset 727

time	meso
0.00000	91.3
10.96104	63.2
21.92209	51.1
32.88313	42.0
43.84417	40.8

Table 13: Dataset 728

time	meso
0.00000	91.8
11.24366	43.6
22.48733	22.0
33.73099	15.9
44.97466	8.8

Table 14: Dataset 729

time	meso
0.00000	91.6
11.24366	60.5
22.48733	43.5
33.73099	28.4
44.97466	20.5

Table 15: Dataset 730

time	meso
0.00000	92.7
11.07446	58.9
22.14893	44.0
33.22339	46.0
44.29785	29.3

Table 16: Dataset 731

time	meso
0.00000	92.1
11.24366	64.4
22.48733	45.3
33.73099	33.6
44.97466	23.5

Table 17: Dataset 732

time	meso
0.00000	90.3
11.24366	58.2
22.48733	40.1
33.73099	33.1
44.97466	25.8

Table 18: Dataset 741

time	meso
0.00000	90.3
10.84712	68.7
21.69424	58.0
32.54136	52.2
43.38848	48.0

Table 19: Dataset 742

time	meso
0.00000	92.0
11.24366	60.9
22.48733	36.2
33.73099	18.3
44.97466	8.7

Separate evaluations

In order to obtain suitable starting parameters for the NLHM fits, separate fits of the five models to the data for each soil are generated using the `mmkin` function from the mkin package. In a first step, constant variance is assumed. Convergence is checked with the `status` function.

```
deg_mods <- c("SFO", "FOMC", "DFOP", "SFORB", "HS")
f_sep_const <- mmkin(
  deg_mods,
  meso_ds,
  error_model = "const",
  cluster = cl,
  quiet = TRUE)
```

```
status(f_sep_const[, 1:5]) |> kable()
```

	Richmond	Richmond 2	ERTC	Toulouse	Picket Piece
SFO	OK	OK	OK	OK	OK
FOMC	OK	OK	OK	OK	C
DFOP	OK	OK	OK	OK	OK
SFORB	OK	OK	OK	OK	OK
HS	OK	OK	C	OK	OK

```
status(f_sep_const[, 6:18]) |> kable()
```

	721	722	723	724	725	727	728	729	730	731	732	741	742
SFO	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
FOMC	OK	OK	C	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
DFOP	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
SFORB	OK	OK	OK	OK	OK	OK	OK	C	OK	OK	OK	OK	OK
HS	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK

In the tables above, OK indicates convergence and C indicates failure to converge. Most separate fits with constant variance converged, with the exception of two FOMC fits, one SFORB fit and one HS fit.

```
f_sep_tc <- update(f_sep_const, error_model = "tc")
```

```
status(f_sep_tc[, 1:5]) |> kable()
```

	Richmond	Richmond 2	ERTC	Toulouse	Picket Piece
SFO	OK	OK	OK	OK	OK
FOMC	OK	OK	OK	OK	OK
DFOP	C	OK	OK	OK	OK
SFORB	OK	OK	OK	OK	OK
HS	OK	OK	C	OK	OK

```
status(f_sep_tc[, 6:18]) |> kable()
```

	721	722	723	724	725	727	728	729	730	731	732	741	742
SFO	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
FOMC	OK	OK	C	OK	C	C	OK	C	OK	C	OK	C	OK
DFOP	C	OK	OK	OK	C	OK	OK	OK	OK	C	OK	C	OK

	721	722	723	724	725	727	728	729	730	731	732	741	742
SFORB	C	OK	OK	OK	C	OK	OK	C	OK	OK	OK	C	OK
HS	OK	OK	OK	OK	OK	OK	OK	OK	OK	C	OK	OK	OK

With the two-component error model, the set of fits that did not converge is larger, with convergence problems appearing for a number of non-SFO fits.

Hierarchical model fits without covariate effect

The following code fits hierarchical kinetic models for the ten combinations of the five different degradation models with the two different error models in parallel.

```
f_saem_1 <- mhmkin(list(f_sep_const, f_sep_tc), cluster = cl)
status(f_saem_1) |> kable()
```

	const	tc
SFO	OK	OK
FOMC	OK	OK
DFOP	OK	OK
SFORB	OK	OK
HS	OK	OK

All fits terminate without errors (status OK).

```
anova(f_saem_1) |> kable(digits = 1)
```

	npar	AIC	BIC	Lik
SFO const	5	800.0	804.5	-395.0
SFO tc	6	802.1	807.4	-395.0
FOMC const	7	787.4	793.6	-386.7
FOMC tc	8	789.2	796.4	-386.6
DFOP const	9	787.6	795.6	-384.8
SFORB const	9	787.4	795.4	-384.7
HS const	9	781.9	789.9	-382.0
DFOP tc	10	787.8	796.7	-383.9
SFORB tc	10	798.1	807.0	-389.1
HS tc	10	785.4	794.3	-382.7

The model comparisons show that the fits with constant variance are consistently preferable to the corresponding fits with two-component error for these data. This is confirmed by the fact that the parameter `b.1` (the relative standard deviation in the fits obtained with the saemix package), is ill-defined in all fits.

```
illparms(f_saem_1) |> kable()
```

	const	tc
SFO	sd(meso_0)	sd(meso_0), b.1
FOMC	sd(meso_0), sd(log_beta)	sd(meso_0), sd(log_beta), b.1
DFOP	sd(meso_0), sd(log_k1)	sd(meso_0), sd(g_qlogis), b.1
SFORB	sd(meso_free_0), sd(log_k_meso_free_bound)	sd(meso_free_0), sd(log_k_meso_free_bound), sd(log_k_meso_bound_free), b.1
HS	sd(meso_0)	sd(meso_0), b.1

For obtaining fits with only well-defined random effects, we update the set of fits, excluding random effects that were ill-defined according to the `illparms` function.

```
f_saem_2 <- update(f_saem_1, no_random_effect = illparms(f_saem_1))
status(f_saem_2) |> kable()
```

	const	tc
SFO	OK	OK

	const	tc
FOMC	OK	OK
DFOP	OK	OK
SFORB	OK	OK
HS	OK	OK

The updated fits terminate without errors.

```
illparms(f_saem_2) |> kable()
```

	const	tc
SFO		b.1
FOMC		b.1
DFOP		b.1
SFORB		b.1
HS		b.1

No ill-defined errors remain in the fits with constant variance.

Hierarchical model fits with covariate effect

In the following sections, hierarchical fits including a model for the influence of pH on selected degradation parameters are shown for all parent models. Constant variance is selected as the error model based on the fits without covariate effects. Random effects that were ill-defined in the fits without pH influence are excluded. A potential influence of the soil pH is only included for parameters with a well-defined random effect, because experience has shown that only for such parameters a significant pH effect could be found.

SFO

```
sfo_pH <- saem(f_sep_const["SFO", ], no_random_effect = "meso_0", covariates = pH,
  covariate_models = list(log_k_meso ~ pH))
```

```
summary(sfo_pH)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	91.35	89.27	93.43
log_k_meso	-6.66	-7.97	-5.35
beta_pH(log_k_meso)	0.59	0.37	0.81
a.1	5.48	4.71	6.24
SD.log_k_meso	0.35	0.23	0.47

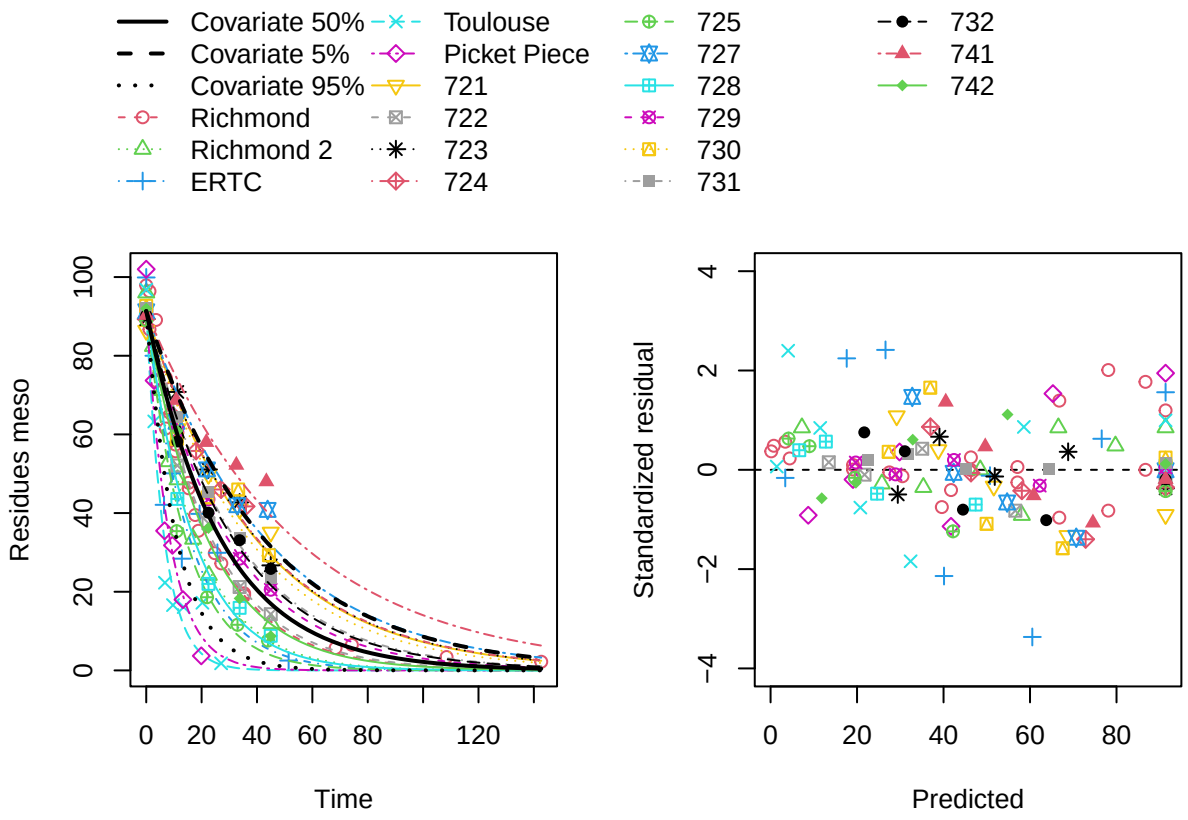
The parameter showing the pH influence in the above table is `beta_pH(log_k_meso)`. Its confidence interval does not include zero, indicating that the influence of soil pH on the log of the degradation rate constant is significantly greater than zero.

```
anova(f_saem_2[["SFO", "const"]], sfo_pH, test = TRUE)
```

```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

                           npar    AIC    BIC     Lik  Chisq Df Pr(>Chisq)
f_saem_2[["SFO", "const"]]    4 797.56 801.12 -394.78
sfo_pH                        5 783.09 787.54 -386.54 16.473  1  4.934e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The comparison with the SFO fit without covariate effect confirms that considering the soil pH improves the model, both by comparison of AIC and BIC and by the likelihood ratio test.

```
plot(sfo_pH)
```



Endpoints for a model with covariates are by default calculated for the median of the covariate values. This quantile can be adapted, or a specific covariate value can be given as shown below.

```
endpoints(sfo_pH)
```

```
$covariates
      pH
50% 5.75

$distimes
          DT50     DT90
meso 18.52069 61.52441
```

```
endpoints(sfo_pH, covariate_quantile = 0.9)
```

```
$covariates
      pH
90% 7.13

$distimes
          DT50     DT90
meso 8.237019 27.36278
```

```
endpoints(sfo_pH, covariates = c(pH = 7.0))
```

```
$covariates
     pH
User  7

$distimes
        DT50     DT90
```

```
meso 8.89035 29.5331
```

FOMC

```
fomc_pH <- saem(f_sep_const["FOMC", ], no_random_effect = "meso_0", covariates = pH,
  covariate_models = list(log_alpha ~ pH))
```

```
summary(fomc_pH)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	92.84	90.75	94.93
log_alpha	-2.21	-3.49	-0.92
beta_pH(log_alpha)	0.58	0.37	0.79
log_beta	4.21	3.44	4.99
a.1	5.03	4.32	5.73
SD.log_alpha	0.00	-23.77	23.78
SD.log_beta	0.37	0.01	0.74

As in the case of SFO, the confidence interval of the slope parameter (here `beta_pH(log_alpha)`) quantifying the influence of soil pH does not include zero, and the model comparison clearly indicates that the model with covariate influence is preferable. However, the random effect for `alpha` is not well-defined any more after inclusion of the covariate effect (the confidence interval of `SD.log_alpha` includes zero).

```
illparms(fomc_pH)
```

```
[1] "sd(log_alpha)"
```

Therefore, the model is updated without this random effect, and no ill-defined parameters remain.

```
fomc_pH_2 <- update(fomc_pH, no_random_effect = c("meso_0", "log_alpha"))
illparms(fomc_pH_2)
```

```
anova(f_saem_2[["FOMC", "const"]], fomc_pH, fomc_pH_2, test = TRUE)
```

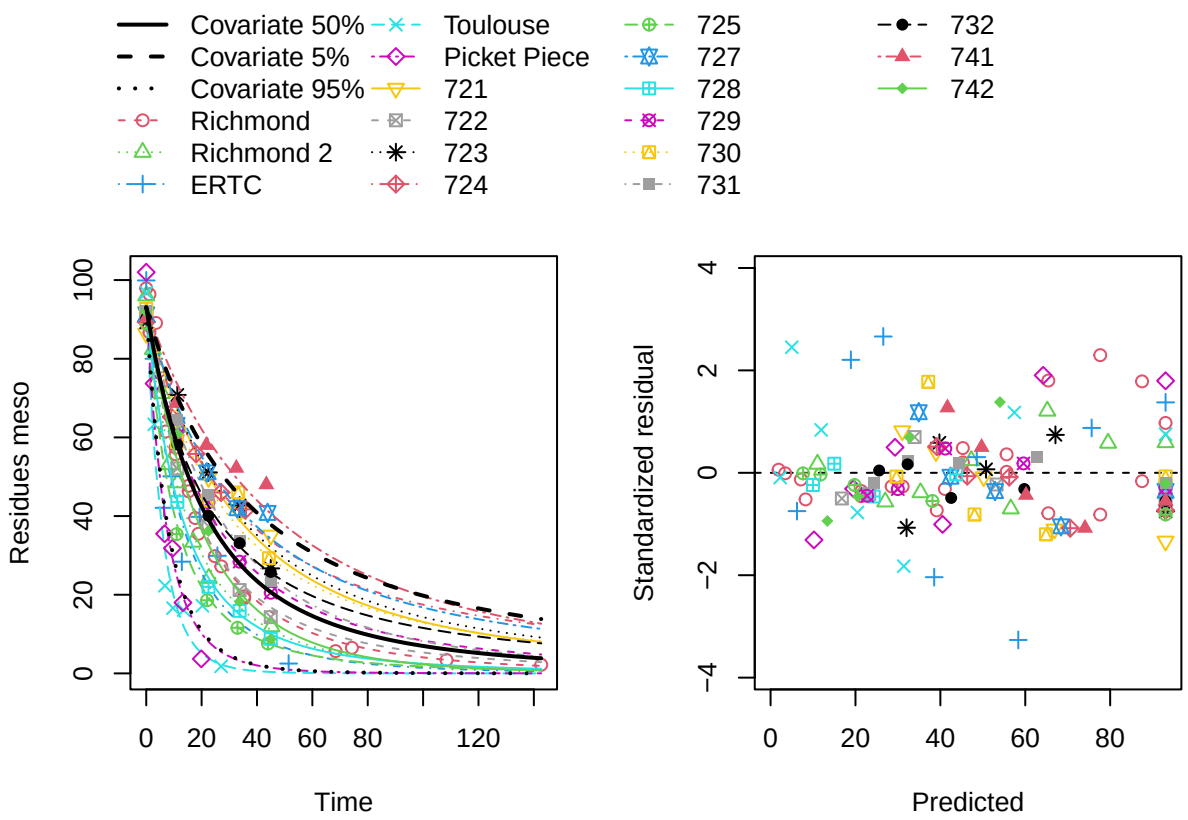
```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

                           npar    AIC    BIC     Lik  Chisq Df Pr(>Chisq)
f_saem_2[["FOMC", "const"]]    5 783.25 787.71 -386.63
fomc_pH_2                      6 767.49 772.83 -377.75 17.762  1  2.503e-05 ***
fomc_pH                        7 770.07 776.30 -378.04  0.000  1          1
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model comparison indicates that including pH dependence significantly improves the fit, and that the reduced model with covariate influence results in the most preferable FOMC fit.

```
summary(fomc_pH_2)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	93.05	90.98	95.13
log_alpha	-2.91	-4.18	-1.63
beta_pH(log_alpha)	0.66	0.44	0.87
log_beta	3.95	3.29	4.62
a.1	4.98	4.28	5.68
SD.log_beta	0.40	0.26	0.54

```
plot(fomc_pH_2)
```




```
endpoints(fomc_pH_2)
```

```
$covariates
      pH
50% 5.75

$distimes
         DT50     DT90 DT50back
meso 17.30248 82.91343 24.95943
```

```
endpoints(fomc_pH_2, covariates = c(pH = 7))
```

```
$covariates
     pH
User  7

$distimes
         DT50     DT90 DT50back
meso 6.986239 27.02927 8.136621
```

DFOP

In the DFOP fits without covariate effects, random effects for two degradation parameters (k2 and g) were identifiable.

```
summary(f_saem_2[["DFOP", "const"]])$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	93.61	91.58	95.63
log_k1	-1.53	-2.27	-0.79
log_k2	-3.42	-3.73	-3.11
g_qlogis	-1.67	-2.57	-0.77
a.1	4.74	4.02	5.45
SD.log_k2	0.60	0.38	0.81
SD.g_qlogis	0.94	0.33	1.54

A fit with pH dependent degradation parameters was obtained by excluding the same random effects as in the refined DFOP fit without covariate influence, and including covariate models for the two identifiable parameters k2 and g.

```
dfop_pH <- saem(f_sep_const["DFOP", ], no_random_effect = c("meso_0", "log_k1"),
  covariates = pH,
  covariate_models = list(log_k2 ~ pH, g_qlogis ~ pH))
```

The corresponding parameters for the influence of soil pH are beta_pH(log_k2) for the influence of soil pH on k2, and beta_pH(g_qlogis) for its influence on g.

```
summary(dfop_pH)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	92.84	90.85	94.84
log_k1	-2.82	-3.09	-2.54
log_k2	-11.48	-15.32	-7.64
beta_pH(log_k2)	1.31	0.69	1.92
g_qlogis	3.13	0.47	5.80
beta_pH(g_qlogis)	-0.57	-1.04	-0.09
a.1	4.96	4.26	5.65
SD.log_k2	0.76	0.47	1.05
SD.g_qlogis	0.01	-9.96	9.97

```
illparms(dfop_pH)
```

```
[1] "sd(g_qlogis)"
```

Confidence intervals for neither of them include zero, indicating a significant difference from zero. However, the random effect for g is now ill-defined. The fit is updated without this ill-defined random effect.

```
dfop_pH_2 <- update(dfop_pH,
  no_random_effect = c("meso_0", "log_k1", "g_qlogis"))
illparms(dfop_pH_2)
```

```
[1] "beta_pH(g_qlogis)"
```

Now, the slope parameter for the pH effect on g is ill-defined. Therefore, another attempt is made without the corresponding covariate model.

```
dfop_pH_3 <- saem(f_sep_const["DFOP", ], no_random_effect = c("meso_0", "log_k1"),
  covariates = pH,
  covariate_models = list(log_k2 ~ pH))
illparms(dfop_pH_3)
```

```
[1] "sd(g_qlogis)"
```

As the random effect for g is again ill-defined, the fit is repeated without it.

```
dfop_pH_4 <- update(dfop_pH_3, no_random_effect = c("meso_0", "log_k1", "g_qlogis"))
illparms(dfop_pH_4)
```

While no ill-defined parameters remain, model comparison suggests that the previous model dfop_pH_2 with two pH dependent parameters is preferable, based on information criteria as well as based on the likelihood ratio test.

```
anova(f_saem_2[["DFOP", "const"]], dfop_pH, dfop_pH_2, dfop_pH_3, dfop_pH_4)
```

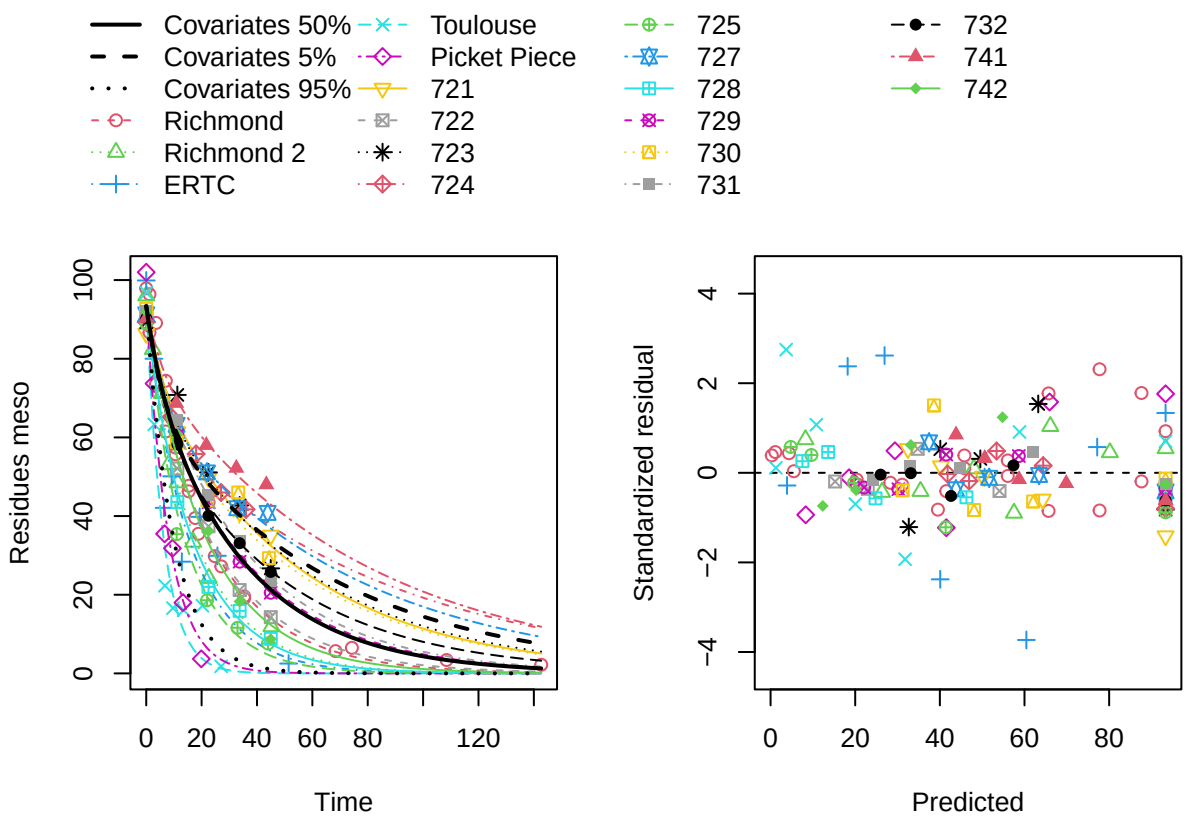
```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

                           npar    AIC    BIC     Lik
f_saem_2[["DFOP", "const"]]    7 782.94 789.18 -384.47
dfop_pH_4                      7 767.35 773.58 -376.68
dfop_pH_2                      8 765.14 772.26 -374.57
dfop_pH_3                      8 769.00 776.12 -376.50
dfop_pH                        9 769.10 777.11 -375.55
```

```
anova(dfop_pH_2, dfop_pH_4, test = TRUE)
```

```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

          npar    AIC    BIC     Lik  Chisq Df Pr(>Chisq)
dfop_pH_4    7 767.35 773.58 -376.68
dfop_pH_2    8 765.14 772.26 -374.57 4.2153  1    0.04006 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

When focussing on parameter identifiability using the test if the confidence interval includes zero, dfop_pH_4 would still be the preferred model. However, it should be kept in mind that parameter confidence intervals are constructed using a simple linearisation of the likelihood. As the confidence interval of the random effect for g only marginally includes zero, it is suggested that this is acceptable, and that dfop_pH_2 can be considered the most preferable model.

```
plot(dfop_pH_2)
```




```
endpoints(dfop_pH_2)
```

```
$covariates
      pH
50% 5.75

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 18.36876 73.51841 22.13125 4.191901 23.98672
```

```
endpoints(dfop_pH_2, covariates = c(pH = 7))
```

```
$covariates
     pH
User  7

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 8.346428 28.34437 8.532507 4.191901 8.753618
```

SFORB

```
sforb_pH <- saem(f_sep_const["SFORB", ], no_random_effect = c("meso_free_0", "log_k_meso_free_bound"),
  covariates = pH,
  covariate_models = list(log_k_meso_free ~ pH, log_k_meso_bound_free ~ pH))
```

```
summary(sforb_pH)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_free_0	93.42	91.32	95.52
log_k_meso_free	-5.37	-6.94	-3.81
beta_pH(log_k_meso_free)	0.42	0.18	0.67
log_k_meso_free_bound	-3.49	-4.92	-2.05
log_k_meso_bound_free	-9.98	-19.22	-0.74
beta_pH(log_k_meso_bound_free)	1.23	-0.21	2.67
a.1	4.90	4.18	5.63
SD.log_k_meso_free	0.35	0.23	0.47
SD.log_k_meso_bound_free	0.13	-1.95	2.20

The confidence interval of `beta_pH(log_k_meso_bound_free)` includes zero, indicating that the influence of soil pH on `k_meso_bound_free` cannot reliably be quantified. Also, the confidence interval for the random effect on this parameter (`SD.log_k_meso_bound_free`) includes zero.

Using the `illparms` function, these ill-defined parameters can be found more conveniently.

```
illparms(sforb_pH)
```

```
[1] "sd(log_k_meso_bound_free)"      "beta_pH(log_k_meso_bound_free)"
```

To remove the ill-defined parameters, a second variant of the SFORB model with pH influence is fitted. No ill-defined parameters remain.

```
sforb_pH_2 <- update(sforb_pH,
  no_random_effect = c("meso_free_0", "log_k_meso_free_bound", "log_k_meso_bound_free"),
  covariate_models = list(log_k_meso_free ~ pH))
illparms(sforb_pH_2)
```

The model comparison of the SFORB fits includes the refined model without covariate effect, and both versions of the SFORB fit with covariate effect.

```
anova(f_saem_2[["SFORB", "const"]], sforb_pH, sforb_pH_2, test = TRUE)
```

```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

                            npar    AIC    BIC     Lik   Chisq Df Pr(>Chisq)
f_saem_2[["SFORB", "const"]]    7 783.40 789.63 -384.70
sforb_pH_2                      7 770.94 777.17 -378.47 12.4616  0
sforb_pH                        9 768.81 776.83 -375.41  6.1258  2    0.04675 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

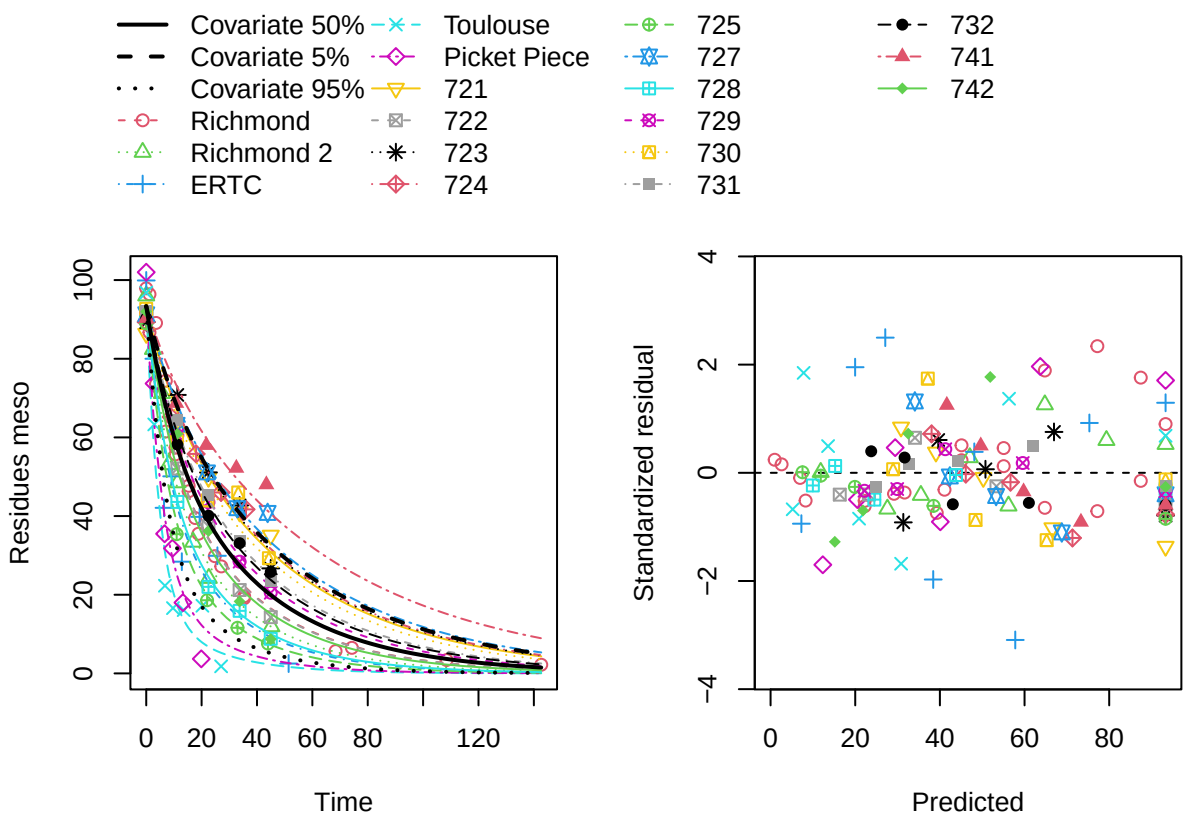
The first model including pH influence is preferable based on information criteria and the likelihood ratio test. However, as it is not fully identifiable, the second model is selected.

```
summary(sforb_pH_2)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_free_0	93.32	91.16	95.48

	est.	lower	upper
log_k_meso_free	-6.15	-7.43	-4.86
beta_pH(log_k_meso_free)	0.54	0.33	0.75
log_k_meso_free_bound	-3.80	-5.20	-2.40
log_k_meso_bound_free	-2.95	-4.26	-1.64
a.1	5.08	4.38	5.79
SD.log_k_meso_free	0.33	0.22	0.45

```
plot(sforb_pH_2)
```




```
endpoints(sforb_pH_2)
```

```
$covariates
      pH
50% 5.75

$ff
meso_free
        1

$SFORB
   meso_b1    meso_b2     meso_g
0.09735824 0.02631699 0.31602120

$distimes
         DT50     DT90 DT50back DT50_meso_b1 DT50_meso_b2
meso 16.86549 73.15824 22.02282     7.119554     26.33839
```

```
endpoints(sforb_pH_2, covariates = c(pH = 7))
```

```
$covariates
     pH
User  7

$ff
meso_free
        1

$SFORB
   meso_b1    meso_b2     meso_g
```

```
0.13315233 0.03795988 0.61186191

$distimes
         DT50     DT90 DT50back DT50_meso_b1 DT50_meso_b2
meso 7.932495 36.93311 11.11797     5.205671        18.26
```

HS

```
hs_pH <- saem(f_sep_const["HS", ], no_random_effect = c("meso_0"),
  covariates = pH,
  covariate_models = list(log_k1 ~ pH, log_k2 ~ pH, log_tb ~ pH))
```

```
summary(hs_pH)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	93.33	91.47	95.19
log_k1	-5.81	-7.27	-4.36
beta_pH(log_k1)	0.47	0.23	0.72
log_k2	-6.80	-8.76	-4.83
beta_pH(log_k2)	0.54	0.21	0.87
log_tb	3.25	1.25	5.25
beta_pH(log_tb)	-0.10	-0.43	0.23
a.1	4.49	3.78	5.21
SD.log_k1	0.37	0.24	0.51
SD.log_k2	0.29	0.10	0.48
SD.log_tb	0.25	-0.07	0.57

```
illparms(hs_pH)
```

```
[1] "sd(log_tb)"      "beta_pH(log_tb)"
```

According to the output of the `illparms` function, the random effect on the break time `tb` cannot reliably be quantified, neither can the influence of soil pH on `tb`. The fit is repeated without the corresponding covariate model, and no ill-defined parameters remain.

```
hs_pH_2 <- update(hs_pH, covariate_models = list(log_k1 ~ pH, log_k2 ~ pH))
illparms(hs_pH_2)
```

Model comparison confirms that this model is preferable to the fit without covariate influence, and also to the first version with covariate influence.

```
anova(f_saem_2[["HS", "const"]], hs_pH, hs_pH_2, test = TRUE)
```

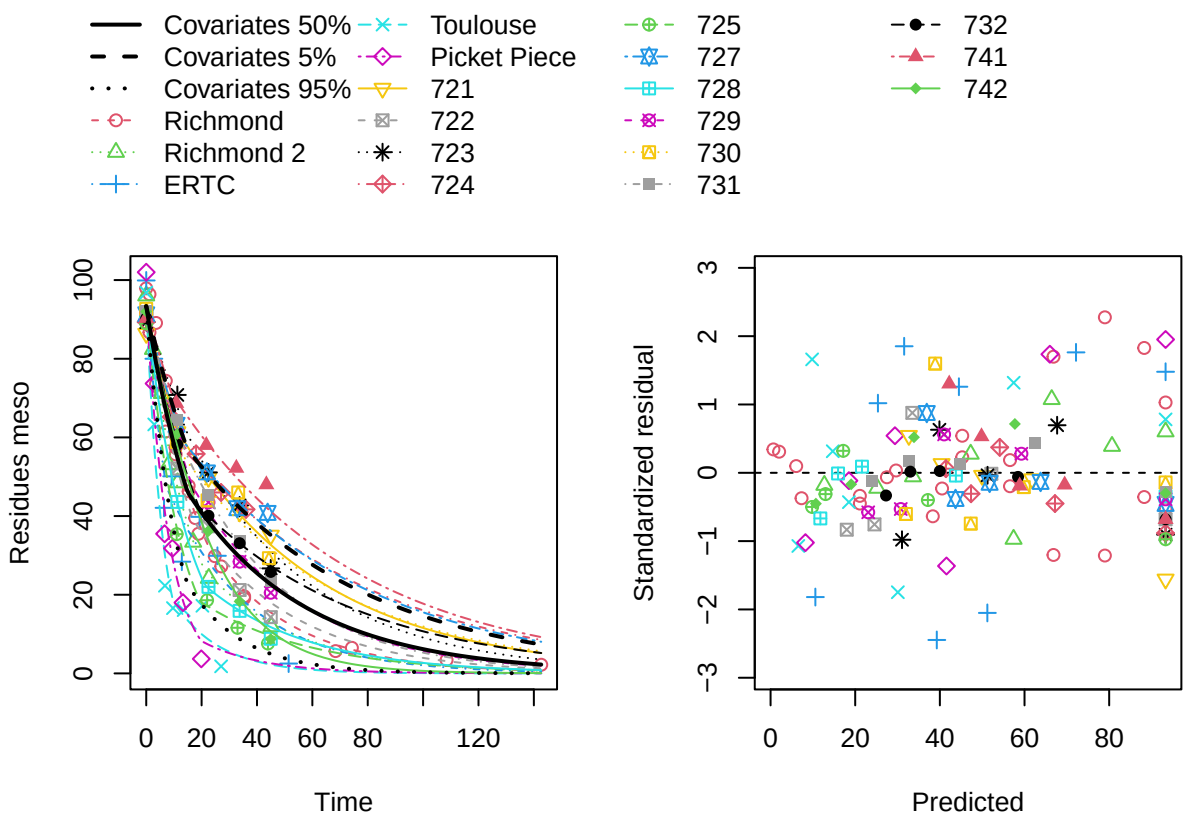
```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

                          npar    AIC    BIC     Lik  Chisq Df Pr(>Chisq)
f_saem_2[["HS", "const"]]    8 780.08 787.20 -382.04
hs_pH_2                     10 766.47 775.37 -373.23 17.606  2  0.0001503 ***
hs_pH                       11 769.80 779.59 -373.90  0.000  1  1.0000000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(hs_pH_2)$confint_trans |> kable(digits = 2)
```

	est.	lower	upper
meso_0	93.33	91.50	95.15
log_k1	-5.68	-7.09	-4.27
beta_pH(log_k1)	0.46	0.22	0.69
log_k2	-6.61	-8.34	-4.88
beta_pH(log_k2)	0.50	0.21	0.79
log_tb	2.70	2.33	3.08
a.1	4.45	3.74	5.16
SD.log_k1	0.36	0.22	0.49

	est.	lower	upper
SD.log__k2	0.23	0.02	0.43
SD.log__tb	0.55	0.25	0.85

```
plot(hs_pH_2)
```




```
endpoints(hs_pH_2)
```

```
$covariates
      pH
50% 5.75

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 14.68725 82.45287 24.82079 14.68725 29.29299
```

```
endpoints(hs_pH_2, covariates = c(pH = 7))
```

```
$covariates
     pH
User  7

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 8.298536 38.85371 11.69613 8.298536 15.71561
```

Comparison across parent models

After model reduction for all models with pH influence, they are compared with each other.

```
anova(sfo_pH, fomc_pH_2, dfop_pH_2, dfop_pH_4, sforb_pH_2, hs_pH_2)
```

```
Data: 116 observations of 1 variable(s) grouped in 18 datasets

           npar    AIC    BIC     Lik
sfo_pH        5 783.09 787.54 -386.54
fomc_pH_2     6 767.49 772.83 -377.75
dfop_pH_4     7 767.35 773.58 -376.68
sforb_pH_2    7 770.94 777.17 -378.47
dfop_pH_2     8 765.14 772.26 -374.57
hs_pH_2      10 766.47 775.37 -373.23
```

The DFOP model with pH influence on `k2` and `g` and a random effect only on `k2` is finally selected as the best fit.

The endpoints resulting from this model are listed below. Please refer to the Appendix for a detailed listing.

```
endpoints(dfop_pH_2)
```

```
$covariates
      pH
50% 5.75

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 18.36876 73.51841 22.13125 4.191901 23.98672
```

```
endpoints(dfop_pH_2, covariates = c(pH = 7))
```

```
$covariates
     pH
User  7

$distimes
         DT50     DT90 DT50back  DT50_k1  DT50_k2
meso 8.346428 28.34437 8.532507 4.191901 8.753618
```

Conclusions

These evaluations demonstrate that covariate effects can be included for all types of parent degradation models. These models can then be further refined to make them fully identifiable.

Appendix

Hierarchical fit listings

Fits without covariate effects

Listing 1: Hierarchical SFO fit with constant variance

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:46:32 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - k_meso * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 1.427 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
    meso_0 log_k_meso
    90.832     -3.192

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
           meso_0 log_k_meso
meso_0      6.752     0.0000
log_k_meso  0.000     0.9155

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
  AIC   BIC logLik
  800 804.5   -395

Optimised parameters:
                est.    lower   upper
meso_0        92.0705  89.9917 94.1493
log_k_meso    -3.1641  -3.4286 -2.8996
a.1            5.4628   4.6421  6.2835
SD.meso_0      0.0611 -98.3545 98.4767
SD.log_k_meso  0.5616   0.3734  0.7499

Correlation:
           meso_0
log_k_meso 0.1132

Random effects:
                est.    lower   upper
SD.meso_0     0.0611 -98.3545 98.4767
SD.log_k_meso 0.5616   0.3734  0.7499

Variance model:
     est. lower upper
a.1 5.463 4.642 6.284

Backtransformed parameters:
           est.    lower    upper
meso_0 92.07053 89.99172 94.14933
k_meso  0.04225  0.03243  0.05505

Estimated disappearance times:
      DT50 DT90
meso 16.41 54.5
```

Listing 2: Hierarchical FOMC fit with constant variance

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:46:33 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - (alpha/beta) * 1/((time/beta) + 1) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 2.248 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
   meso_0 log_alpha  log_beta
  93.0520    0.6008    3.4176

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
          meso_0 log_alpha log_beta
meso_0     6.287      0.00    0.000
log_alpha  0.000      1.53    0.000
log_beta   0.000      0.00    1.724

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  787.4 793.6 -386.7

Optimised parameters:
               est.     lower   upper
meso_0       93.5648  91.42864 95.7009
log_alpha     0.7645   0.28068  1.2484
log_beta      3.6597   3.05999  4.2594
a.1           5.0708   4.29823  5.8435
SD.meso_0     0.1691 -34.01517 34.3535
SD.log_alpha  0.3764   0.05834  0.6945
SD.log_beta   0.3903  -0.06074  0.8414

Correlation:
          meso_0  log_lph
log_alpha -0.2839
log_beta  -0.3443  0.8855

Random effects:
               est.     lower   upper
SD.meso_0    0.1691 -34.01517 34.3535
SD.log_alpha 0.3764   0.05834  0.6945
SD.log_beta  0.3903  -0.06074  0.8414

Variance model:
     est. lower upper
a.1 5.071 4.298 5.843

Backtransformed parameters:
         est.  lower  upper
meso_0 93.565 91.429 95.701
alpha   2.148  1.324  3.485
beta   38.850 21.327 70.770

Estimated disappearance times:
     DT50  DT90 DT50back
meso 14.8 74.64    22.47
```

Listing 3: Hierarchical DFOP fit with constant variance

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:46:35 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ((k1 * g * exp(-k1 * time) + k2 * (1 - g) * exp(-k2 *
           time)) / (g * exp(-k1 * time) + (1 - g) * exp(-k2 * time)))
           * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 4.095 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
  meso_0   log_k1   log_k2 g_qlogis
93.14689 -2.05241 -3.53079 -0.09522

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
         meso_0 log_k1 log_k2 g_qlogis
meso_0    6.418  0.000  0.000     0.00
log_k1    0.000  1.018  0.000     0.00
log_k2    0.000  0.000  1.694     0.00
g_qlogis  0.000  0.000  0.000     2.37

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  787.6 795.6 -384.8

Optimised parameters:
               est.     lower   upper
meso_0      93.6684  91.63599 95.7008
log_k1      -1.7354  -2.61433 -0.8565
log_k2      -3.4015  -3.73323 -3.0697
g_qlogis    -1.6341  -2.66133 -0.6069
a.1          4.7803   4.01269  5.5479
SD.meso_0    0.1661 -30.97086 31.3031
SD.log_k1    0.1127  -2.59680  2.8223
SD.log_k2    0.6394   0.41499  0.8638
SD.g_qlogis  0.8166   0.09785  1.5353

Correlation:
         meso_0  log_k1  log_k2
log_k1    0.1757
log_k2    0.0199  0.2990
g_qlogis  0.0813 -0.7431 -0.3826

Random effects:
              est.     lower   upper
SD.meso_0   0.1661 -30.97086 31.3031
SD.log_k1   0.1127  -2.59680  2.8223
SD.log_k2   0.6394   0.41499  0.8638
SD.g_qlogis 0.8166   0.09785  1.5353

Variance model:
    est. lower upper
a.1 4.78 4.013 5.548

Backtransformed parameters:
           est.    lower    upper
meso_0 93.66841 91.63599 95.70082
k1      0.17633  0.07322  0.42466
k2      0.03332  0.02392  0.04643
g       0.16327  0.06529  0.35277

Estimated disappearance times:
```

```
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 16.04 63.75    19.19   3.931    20.8
```

Listing 4: Hierarchical SFORB fit with constant variance

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:46:35 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso_free/dt = - k_meso_free * meso_free - k_meso_free_bound *
          meso_free + k_meso_bound_free * meso_bound
d_meso_bound/dt = + k_meso_free_bound * meso_free - k_meso_bound_free *
          meso_bound

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 3.545 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
          meso_free_0       log_k_meso_free log_k_meso_free_bound
                93.147                -2.305                -4.230
log_k_meso_bound_free
                -3.761

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
                      meso_free_0 log_k_meso_free log_k_meso_free_bound
meso_free_0                 6.418          0.0000                 0.000
log_k_meso_free             0.000          0.9276                 0.000
log_k_meso_free_bound       0.000          0.0000                 2.272
log_k_meso_bound_free       0.000          0.0000                 0.000
                      log_k_meso_bound_free
meso_free_0                           0.000
log_k_meso_free                       0.000
log_k_meso_free_bound                 0.000
log_k_meso_bound_free                 1.447

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  787.4 795.4 -384.7

Optimised parameters:
                            est.    lower  upper
meso_free_0              93.6285  91.6262 95.631
log_k_meso_free          -2.8314  -3.1375 -2.525
log_k_meso_free_bound    -3.2213  -4.4695 -1.973
log_k_meso_bound_free    -2.4246  -3.5668 -1.282
a.1                       4.7372   3.9542  5.520
SD.meso_free_0            0.1634 -32.7769 33.104
SD.log_k_meso_free        0.4885   0.3080  0.669
SD.log_k_meso_free_bound  0.2876  -1.7955  2.371
SD.log_k_meso_bound_free  0.9942   0.2181  1.770

Correlation:
                      ms_fr_0 lg_k_m_ lg_k_ms_f_
log_k_meso_free        0.2332
log_k_meso_free_bound  0.1100  0.5964
log_k_meso_bound_free -0.0413  0.3697  0.8025

Random effects:
                           est.    lower  upper
SD.meso_free_0           0.1634 -32.7769 33.104
SD.log_k_meso_free       0.4885   0.3080  0.669
SD.log_k_meso_free_bound 0.2876  -1.7955  2.371
SD.log_k_meso_bound_free 0.9942   0.2181  1.770

Variance model:
     est. lower upper
a.1 4.737 3.954  5.52
```

```
Backtransformed parameters:
                     est.    lower    upper
meso_free_0       93.62849 91.62622 95.63075
k_meso_free        0.05893  0.04339  0.08004
k_meso_free_bound  0.03990  0.01145  0.13903
k_meso_bound_free  0.08851  0.02825  0.27736

Estimated Eigenvalues of SFORB model(s):
meso_b1 meso_b2  meso_g
0.15333 0.03402 0.20881

Resulting formation fractions:
          ff
meso_free  1

Estimated disappearance times:
      DT50  DT90 DT50back DT50_meso_b1 DT50_meso_b2
meso 14.79 60.81     18.3        4.521        20.37
```

Listing 5: Hierarchical HS fit with constant variance

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:46:35 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ifelse(time <= tb, k1, k2) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 3.694 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
meso_0 log_k1 log_k2 log_tb
92.920 -2.409 -3.295  2.471

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
       meso_0 log_k1 log_k2 log_tb
meso_0  6.477 0.0000 0.0000   0.00
log_k1  0.000 0.8675 0.0000   0.00
log_k2  0.000 0.0000 0.4035   0.00
log_tb  0.000 0.0000 0.0000   1.16

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  781.9 789.9   -382

Optimised parameters:
              est.    lower   upper
meso_0    93.34242  91.4730 95.2118
log_k1    -2.77312  -3.0826 -2.4637
log_k2    -3.61854  -3.8430 -3.3941
log_tb     2.00266   1.3357  2.6696
a.1        4.47693   3.7059  5.2479
SD.meso_0  0.07963 -63.1661 63.3253
SD.log_k1  0.47817   0.2467  0.7097
SD.log_k2  0.39216   0.2137  0.5706
SD.log_tb  0.94683   0.4208  1.4728

Correlation:
       meso_0  log_k1  log_k2
log_k1  0.1627
log_k2  0.0063 -0.0301
log_tb  0.0083 -0.3931 -0.1225

Random effects:
             est.    lower   upper
SD.meso_0 0.07963 -63.1661 63.3253
SD.log_k1 0.47817   0.2467  0.7097
SD.log_k2 0.39216   0.2137  0.5706
SD.log_tb 0.94683   0.4208  1.4728

Variance model:
     est. lower upper
a.1 4.477 3.706 5.248

Backtransformed parameters:
           est.    lower    upper
meso_0 93.34242 91.47303 95.21181
k1      0.06247  0.04584  0.08512
k2      0.02682  0.02143  0.03357
tb      7.40872  3.80282 14.43376

Estimated disappearance times:
     DT50 DT90 DT50back DT50_k1 DT50_k2
meso   16   76    22.88    11.1   25.84
```

Fits with covariate effects

Listing 6: Hierarchichal SFO fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:08 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - k_meso * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 2.609 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
    meso_0 log_k_meso
    90.832     -3.192

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
           meso_0 log_k_meso
meso_0      6.752     0.0000
log_k_meso  0.000     0.9155

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  783.1 787.5 -386.5

Optimised parameters:
                       est.   lower   upper
meso_0              91.3481 89.2688 93.4275
log_k_meso          -6.6614 -7.9715 -5.3514
beta_pH(log_k_meso)  0.5871  0.3684  0.8059
a.1                  5.4750  4.7085  6.2415
SD.log_k_meso        0.3471  0.2258  0.4684

Correlation:
                     meso_0  lg_k_ms
log_k_meso           0.0414
beta_pH(log_k_meso) -0.0183 -0.9917

Random effects:
                est.  lower  upper
SD.log_k_meso 0.3471 0.2258 0.4684

Variance model:
     est. lower upper
a.1 5.475 4.709 6.242

Backtransformed parameters:
            est.     lower     upper
meso_0 91.348139 8.927e+01 93.427476
k_meso  0.001279 3.452e-04  0.004741

Covariates used for endpoints below:
      pH
50% 5.75

Estimated disappearance times:
      DT50  DT90
meso 18.52 61.52
```

Listing 7: Hierarchichal FOMC fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:13 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - (alpha/beta) * 1/((time/beta) + 1) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 3.123 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
   meso_0 log_alpha  log_beta
  93.0520    0.6008    3.4176

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
          meso_0 log_alpha log_beta
meso_0     6.287      0.00    0.000
log_alpha  0.000      1.53    0.000
log_beta   0.000      0.00    1.724

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  770.1 776.3   -378

Optimised parameters:
                        est.      lower   upper
meso_0             92.840646  90.750461 94.9308
log_alpha          -2.206602  -3.494546 -0.9187
beta_pH(log_alpha)  0.577505   0.369805  0.7852
log_beta            4.214099   3.438851  4.9893
a.1                 5.027768   4.322028  5.7335
SD.log_alpha        0.004034 -23.766993 23.7751
SD.log_beta         0.374640   0.009252  0.7400

Correlation:
                   meso_0  log_lph bt_H(_)
log_alpha          -0.0865
beta_pH(log_alpha) -0.0789 -0.8704
log_beta           -0.3544  0.3302  0.1628

Random effects:
                 est.      lower upper
SD.log_alpha 0.004034 -23.766993 23.78
SD.log_beta  0.374640   0.009252  0.74

Variance model:
     est. lower upper
a.1 5.028 4.322 5.734

Backtransformed parameters:
          est.    lower    upper
meso_0 92.8406 90.75046  94.9308
alpha   0.1101  0.03036   0.3991
beta   67.6332 31.15113 146.8404

Covariates used for endpoints below:
      pH
50% 5.75

Estimated disappearance times:
      DT50  DT90 DT50back
meso 17.28 76.37    22.99
```

Listing 8: Refined hierarchichal FOMC fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:21 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - (alpha/beta) * 1/((time/beta) + 1) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 5.989 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
   meso_0 log_alpha  log_beta
  93.0520    0.6008    3.4176

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
          meso_0 log_alpha log_beta
meso_0     6.287      0.00    0.000
log_alpha  0.000      1.53    0.000
log_beta   0.000      0.00    1.724

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  767.5 772.8 -377.7

Optimised parameters:
                       est.   lower   upper
meso_0              93.0536 90.9771 95.1300
log_alpha          -2.9054 -4.1803 -1.6304
beta_pH(log_alpha)  0.6590  0.4437  0.8744
log_beta            3.9549  3.2860  4.6239
a.1                 4.9784  4.2815  5.6754
SD.log_beta         0.4019  0.2632  0.5406

Correlation:
                   meso_0  log_lph bt_H(_)
log_alpha          -0.0397
beta_pH(log_alpha) -0.0899 -0.9146
log_beta           -0.3473  0.2038  0.1919

Random effects:
              est.  lower  upper
SD.log_beta 0.4019 0.2632 0.5406

Variance model:
     est. lower upper
a.1 4.978 4.281 5.675

Backtransformed parameters:
           est.    lower    upper
meso_0 93.05359 90.97713  95.1300
alpha   0.05473  0.01529   0.1958
beta   52.19251 26.73597 101.8874

Covariates used for endpoints below:
      pH
50% 5.75

Estimated disappearance times:
     DT50  DT90 DT50back
meso 17.3 82.91    24.96
```

Listing 9: Hierarchichal DFOP fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:31 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ((k1 * g * exp(-k1 * time) + k2 * (1 - g) * exp(-k2 *
           time)) / (g * exp(-k1 * time) + (1 - g) * exp(-k2 * time)))
           * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 8.366 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
  meso_0   log_k1   log_k2 g_qlogis
93.14689 -2.05241 -3.53079 -0.09522

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
         meso_0 log_k1 log_k2 g_qlogis
meso_0    6.418  0.000  0.000     0.00
log_k1    0.000  1.018  0.000     0.00
log_k2    0.000  0.000  1.694     0.00
g_qlogis  0.000  0.000  0.000     2.37

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  769.1 777.1 -375.5

Optimised parameters:
                        est.    lower    upper
meso_0             92.843344  90.8464 94.84028
log_k1             -2.815685  -3.0888 -2.54261
log_k2            -11.479779 -15.3203 -7.63923
beta_pH(log_k2)     1.308417   0.6948  1.92203
g_qlogis            3.133036   0.4657  5.80035
beta_pH(g_qlogis)  -0.565988  -1.0394 -0.09262
a.1                 4.955518   4.2597  5.65135
SD.log_k2           0.758963   0.4685  1.04943
SD.g_qlogis         0.005215  -9.9561  9.96656

Correlation:
                  meso_0  log_k1  log_k2  b_H(_2) g_qlogs
log_k1             0.2706
log_k2            -0.0571  0.1096
beta_pH(log_k2)    0.0554 -0.1291 -0.9937
g_qlogis          -0.1125 -0.5062 -0.1305  0.1294
beta_pH(g_qlogis)  0.1267  0.4226  0.0419 -0.0438 -0.9864

Random effects:
                est.   lower upper
SD.log_k2   0.758963  0.4685 1.049
SD.g_qlogis 0.005215 -9.9561 9.967

Variance model:
     est. lower upper
a.1 4.956  4.26 5.651

Backtransformed parameters:
            est.     lower     upper
meso_0 9.284e+01 9.085e+01 9.484e+01
k1     5.986e-02 4.556e-02 7.866e-02
k2     1.034e-05 2.221e-07 4.812e-04
g      9.582e-01 6.144e-01 9.970e-01

Covariates used for endpoints below:
```

```
      pH
50% 5.75

Estimated disappearance times:
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 20.23 88.45    26.62   11.58   36.23
```

Listing 10: Refined hierarchical DFOP fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:41 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ((k1 * g * exp(-k1 * time) + k2 * (1 - g) * exp(-k2 *
           time)) / (g * exp(-k1 * time) + (1 - g) * exp(-k2 * time)))
           * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 8.863 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
  meso_0   log_k1   log_k2 g_qlogis
93.14689 -2.05241 -3.53079 -0.09522

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
         meso_0 log_k1 log_k2 g_qlogis
meso_0    6.418  0.000  0.000     0.00
log_k1    0.000  1.018  0.000     0.00
log_k2    0.000  0.000  1.694     0.00
g_qlogis  0.000  0.000  0.000     2.37

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  765.1 772.3 -374.6

Optimised parameters:
                      est.    lower    upper
meso_0            93.3333  91.2427 95.42394
log_k1            -1.7997  -2.9124 -0.68698
log_k2            -8.1810 -10.1819 -6.18008
beta_pH(log_k2)    0.8064   0.4903  1.12257
g_qlogis           3.3513  -1.1792  7.88182
beta_pH(g_qlogis) -0.8672  -1.7661  0.03177
a.1                4.9158   4.2277  5.60390
SD.log_k2          0.3946   0.2565  0.53281

Correlation:
                  meso_0  log_k1  log_k2  b_H(_2) g_qlogs
log_k1             0.1730
log_k2             0.0442  0.5370
beta_pH(log_k2)   -0.0392 -0.4880 -0.9923
g_qlogis          -0.1536  0.1431 -0.1129  0.1432
beta_pH(g_qlogis)  0.1504 -0.3151 -0.0196 -0.0212 -0.9798

Random effects:
            est.  lower  upper
SD.log_k2 0.3946 0.2565 0.5328

Variance model:
     est. lower upper
a.1 4.916 4.228 5.604

Backtransformed parameters:
            est.     lower    upper
meso_0 9.333e+01 9.124e+01 95.42394
k1     1.654e-01 5.435e-02  0.50309
k2     2.799e-04 3.785e-05  0.00207
g      9.661e-01 2.352e-01  0.99962

Covariates used for endpoints below:
      pH
50% 5.75
```

```
Estimated disappearance times:
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 18.37 73.52    22.13   4.192   23.99
```

Listing 11: Further refined hierarchical DFOP fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:47:58 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ((k1 * g * exp(-k1 * time) + k2 * (1 - g) * exp(-k2 *
           time)) / (g * exp(-k1 * time) + (1 - g) * exp(-k2 * time)))
           * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 8.501 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
  meso_0   log_k1   log_k2 g_qlogis
93.14689 -2.05241 -3.53079 -0.09522

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
         meso_0 log_k1 log_k2 g_qlogis
meso_0    6.418  0.000  0.000     0.00
log_k1    0.000  1.018  0.000     0.00
log_k2    0.000  0.000  1.694     0.00
g_qlogis  0.000  0.000  0.000     2.37

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  767.4 773.6 -376.7

Optimised parameters:
                    est.    lower   upper
meso_0          93.3011  91.1905 95.4118
log_k1          -2.1487  -2.7607 -1.5367
log_k2          -8.1039 -10.4225 -5.7853
beta_pH(log_k2)  0.7821   0.4126  1.1517
g_qlogis        -1.0373  -1.9337 -0.1409
a.1              5.0095   4.3082  5.7108
SD.log_k2        0.4622   0.3009  0.6235

Correlation:
                meso_0  log_k1  log_k2  b_H(_2)
log_k1           0.2179
log_k2           0.0337  0.5791
beta_pH(log_k2) -0.0326 -0.5546 -0.9932
g_qlogis         0.0237 -0.8479 -0.6571  0.6123

Random effects:
            est.  lower  upper
SD.log_k2 0.4622 0.3009 0.6235

Variance model:
     est. lower upper
a.1 5.009 4.308 5.711

Backtransformed parameters:
            est.     lower     upper
meso_0 9.330e+01 9.119e+01 95.411751
k1     1.166e-01 6.325e-02  0.215084
k2     3.024e-04 2.975e-05  0.003072
g      2.617e-01 1.263e-01  0.464832

Covariates used for endpoints below:
      pH
50% 5.75

Estimated disappearance times:
```

```
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 17.09 73.67    22.18   5.943   25.54
```

Listing 12: Hierarchichal SFORB fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:48:03 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso_free/dt = - k_meso_free * meso_free - k_meso_free_bound *
          meso_free + k_meso_bound_free * meso_bound
d_meso_bound/dt = + k_meso_free_bound * meso_free - k_meso_bound_free *
          meso_bound

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 4.259 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
          meso_free_0       log_k_meso_free log_k_meso_free_bound
                93.147                -2.305                -4.230
log_k_meso_bound_free
                -3.761

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
                      meso_free_0 log_k_meso_free log_k_meso_free_bound
meso_free_0                 6.418          0.0000                 0.000
log_k_meso_free             0.000          0.9276                 0.000
log_k_meso_free_bound       0.000          0.0000                 2.272
log_k_meso_bound_free       0.000          0.0000                 0.000
                      log_k_meso_bound_free
meso_free_0                           0.000
log_k_meso_free                       0.000
log_k_meso_free_bound                 0.000
log_k_meso_bound_free                 1.447

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  768.8 776.8 -375.4

Optimised parameters:
                                 est.    lower   upper
meso_free_0                   93.4204  91.3213 95.5195
log_k_meso_free               -5.3742  -6.9366 -3.8117
beta_pH(log_k_meso_free)       0.4232   0.1769  0.6695
log_k_meso_free_bound         -3.4889  -4.9243 -2.0535
log_k_meso_bound_free         -9.9797 -19.2232 -0.7362
beta_pH(log_k_meso_bound_free)  1.2290  -0.2107  2.6687
a.1                             4.9031   4.1795  5.6268
SD.log_k_meso_free              0.3454   0.2252  0.4656
SD.log_k_meso_bound_free        0.1277  -1.9459  2.2012

Correlation:
                              ms_fr_0 lg_k_m_ b_H(___) lg_k_ms_f_ lg_k_ms_b_
log_k_meso_free                0.1493
beta_pH(log_k_meso_free)      -0.0930 -0.9854
log_k_meso_free_bound          0.2439  0.4621 -0.3492
log_k_meso_bound_free          0.2188  0.1292 -0.0339   0.7287
beta_pH(log_k_meso_bound_free) -0.2216 -0.0797 -0.0111  -0.6566    -0.9934

Random effects:
                          est.   lower  upper
SD.log_k_meso_free       0.3454  0.2252 0.4656
SD.log_k_meso_bound_free 0.1277 -1.9459 2.2012

Variance model:
     est. lower upper
a.1 4.903  4.18 5.627
```

```
Backtransformed parameters:
                       est.     lower    upper
meso_free_0       9.342e+01 9.132e+01 95.51946
k_meso_free       4.635e-03 9.716e-04  0.02211
k_meso_free_bound 3.054e-02 7.268e-03  0.12829
k_meso_bound_free 4.633e-05 4.482e-09  0.47894

Covariates used for endpoints below:
      pH
50% 5.75

Estimated Eigenvalues of SFORB model(s):
meso_b1 meso_b2  meso_g
 0.1121  0.0256  0.3148

Resulting formation fractions:
          ff
meso_free  1

Estimated disappearance times:
      DT50 DT90 DT50back DT50_meso_b1 DT50_meso_b2
meso 16.42 75.2    22.64        6.185        27.08
```

Listing 13: Refined hierarchichal SFORB fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:48:11 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso_free/dt = - k_meso_free * meso_free - k_meso_free_bound *
          meso_free + k_meso_bound_free * meso_bound
d_meso_bound/dt = + k_meso_free_bound * meso_free - k_meso_bound_free *
          meso_bound

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 6.928 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
          meso_free_0       log_k_meso_free log_k_meso_free_bound
                93.147                -2.305                -4.230
log_k_meso_bound_free
                -3.761

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
                      meso_free_0 log_k_meso_free log_k_meso_free_bound
meso_free_0                 6.418          0.0000                 0.000
log_k_meso_free             0.000          0.9276                 0.000
log_k_meso_free_bound       0.000          0.0000                 2.272
log_k_meso_bound_free       0.000          0.0000                 0.000
                      log_k_meso_bound_free
meso_free_0                           0.000
log_k_meso_free                       0.000
log_k_meso_free_bound                 0.000
log_k_meso_bound_free                 1.447

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  770.9 777.2 -378.5

Optimised parameters:
                             est.   lower   upper
meso_free_0              93.3196 91.1633 95.4760
log_k_meso_free          -6.1460 -7.4306 -4.8614
beta_pH(log_k_meso_free)  0.5435  0.3329  0.7542
log_k_meso_free_bound    -3.8001 -5.2027 -2.3975
log_k_meso_bound_free    -2.9462 -4.2565 -1.6359
a.1                       5.0825  4.3793  5.7856
SD.log_k_meso_free        0.3338  0.2175  0.4502

Correlation:
                         ms_fr_0 lg_k_m_ b_H(___ lg_k_ms_f_
log_k_meso_free           0.1086
beta_pH(log_k_meso_free) -0.0426 -0.9821
log_k_meso_free_bound     0.2513  0.1717 -0.0409
log_k_meso_bound_free     0.1297  0.1171 -0.0139  0.9224

Random effects:
                     est.  lower  upper
SD.log_k_meso_free 0.3338 0.2175 0.4502

Variance model:
     est. lower upper
a.1 5.082 4.379 5.786

Backtransformed parameters:
                       est.     lower    upper
meso_free_0       93.319649 9.116e+01 95.47601
k_meso_free        0.002142 5.928e-04  0.00774
```

```
k_meso_free_bound  0.022369 5.502e-03  0.09095
k_meso_bound_free  0.052539 1.417e-02  0.19478

Covariates used for endpoints below:
      pH
50% 5.75

Estimated Eigenvalues of SFORB model(s):
meso_b1 meso_b2  meso_g
0.09736 0.02632 0.31602

Resulting formation fractions:
          ff
meso_free  1

Estimated disappearance times:
      DT50  DT90 DT50back DT50_meso_b1 DT50_meso_b2
meso 16.87 73.16    22.02         7.12        26.34
```

Listing 14: Hierarchichal HS fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:48:16 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ifelse(time <= tb, k1, k2) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 4.058 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
meso_0 log_k1 log_k2 log_tb
92.920 -2.409 -3.295  2.471

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
       meso_0 log_k1 log_k2 log_tb
meso_0  6.477 0.0000 0.0000   0.00
log_k1  0.000 0.8675 0.0000   0.00
log_k2  0.000 0.0000 0.4035   0.00
log_tb  0.000 0.0000 0.0000   1.16

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  769.8 779.6 -373.9

Optimised parameters:
                    est.   lower   upper
meso_0          93.32599 91.4658 95.1862
log_k1          -5.81463 -7.2710 -4.3583
beta_pH(log_k1)  0.47472  0.2334  0.7160
log_k2          -6.79633 -8.7605 -4.8322
beta_pH(log_k2)  0.54151  0.2124  0.8706
log_tb           3.24674  1.2470  5.2465
beta_pH(log_tb) -0.09889 -0.4258  0.2280
a.1              4.49487  3.7766  5.2132
SD.log_k1        0.37191  0.2370  0.5068
SD.log_k2        0.29210  0.0994  0.4848
SD.log_tb        0.25353 -0.0664  0.5735

Correlation:
                meso_0  log_k1  b_H(_1) log_k2  b_H(_2) log_tb
log_k1           0.0744
beta_pH(log_k1) -0.0452 -0.9915
log_k2           0.0066 -0.0363  0.0376
beta_pH(log_k2) -0.0071  0.0372 -0.0391 -0.9939
log_tb          -0.0238 -0.1483  0.1362 -0.3836  0.3696
beta_pH(log_tb)  0.0097  0.1359 -0.1265  0.3736 -0.3653 -0.9905

Random effects:
            est.   lower  upper
SD.log_k1 0.3719  0.2370 0.5068
SD.log_k2 0.2921  0.0994 0.4848
SD.log_tb 0.2535 -0.0664 0.5735

Variance model:
     est. lower upper
a.1 4.495 3.777 5.213

Backtransformed parameters:
            est.     lower     upper
meso_0 93.325994 9.147e+01 9.519e+01
k1      0.002984 6.954e-04 1.280e-02
k2      0.001118 1.568e-04 7.969e-03
tb     25.706437 3.480e+00 1.899e+02
```

```
Covariates used for endpoints below:
      pH
50% 5.75

Estimated disappearance times:
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 15.65 79.63    23.97   15.16   27.55
```

Listing 15: Refined hierarchichal HS fit with pH influence

```
saemix version used for fitting:      3.1
mkin version used for pre-fitting:  1.2.5
R version used for fitting:         4.3.1
Date of fit:     Tue Aug  8 15:48:21 2023
Date of summary: Tue Aug  8 15:48:22 2023

Equations:
d_meso/dt = - ifelse(time <= tb, k1, k2) * meso

Data:
116 observations of 1 variable(s) grouped in 18 datasets

Model predictions using solution type analytical

Fitted in 4.118 s
Using 300, 100 iterations and 3 chains

Variance model: Constant variance

Starting values for degradation parameters:
meso_0 log_k1 log_k2 log_tb
92.920 -2.409 -3.295  2.471

Fixed degradation parameter values:
None

Starting values for random effects (square root of initial entries in omega):
       meso_0 log_k1 log_k2 log_tb
meso_0  6.477 0.0000 0.0000   0.00
log_k1  0.000 0.8675 0.0000   0.00
log_k2  0.000 0.0000 0.4035   0.00
log_tb  0.000 0.0000 0.0000   1.16

Starting values for error model parameters:
a.1
  1

Results:

Likelihood computed by importance sampling
    AIC   BIC logLik
  766.5 775.4 -373.2

Optimised parameters:
                   est.    lower   upper
meso_0          93.3251 91.49823 95.1520
log_k1          -5.6796 -7.08789 -4.2714
beta_pH(log_k1)  0.4567  0.22400  0.6894
log_k2          -6.6083 -8.33839 -4.8781
beta_pH(log_k2)  0.4982  0.20644  0.7899
log_tb           2.7040  2.33033  3.0777
a.1              4.4452  3.73537  5.1551
SD.log_k1        0.3570  0.22104  0.4930
SD.log_k2        0.2252  0.01864  0.4318
SD.log_tb        0.5488  0.24560  0.8521

Correlation:
                meso_0  log_k1  b_H(_1) log_k2  b_H(_2)
log_k1           0.0740
beta_pH(log_k1) -0.0453 -0.9912
log_k2           0.0115 -0.0650  0.0661
beta_pH(log_k2) -0.0116  0.0649 -0.0667 -0.9936
log_tb          -0.0658 -0.1135  0.0913 -0.1500  0.1210

Random effects:
            est.   lower  upper
SD.log_k1 0.3570 0.22104 0.4930
SD.log_k2 0.2252 0.01864 0.4318
SD.log_tb 0.5488 0.24560 0.8521

Variance model:
     est. lower upper
a.1 4.445 3.735 5.155

Backtransformed parameters:
            est.     lower     upper
meso_0 93.325134 9.150e+01 95.152036
k1      0.003415 8.352e-04  0.013962
k2      0.001349 2.392e-04  0.007611
tb     14.939247 1.028e+01 21.707445

Covariates used for endpoints below:
```

```
      pH
50% 5.75

Estimated disappearance times:
      DT50  DT90 DT50back DT50_k1 DT50_k2
meso 14.69 82.45    24.82   14.69   29.29
```

Session info

```
R version 4.3.1 (2023-06-16)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 22.04.3 LTS

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.10.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.10.0

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

time zone: Europe/Zurich
tzcode source: system (glibc)

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] saemix_3.1 npde_3.3   knitr_1.43 mkin_1.2.5

loaded via a namespace (and not attached):
 [1] vctrs_0.6.3       nlme_3.1-162      cli_3.6.1         rlang_1.1.1
 [5] xfun_0.39         mclust_6.0.0      generics_0.1.3    zoo_1.8-12
 [9] glue_1.6.2        colorspace_2.1-0  htmltools_0.5.5   gridExtra_2.3
[13] lmtest_0.9-40     readxl_1.4.2      fansi_1.0.4       scales_1.2.1
[17] rmarkdown_2.23    cellranger_1.1.0  grid_4.3.1        tibble_3.2.1
[21] evaluate_0.21     munsell_0.5.0     fastmap_1.1.1     yaml_2.3.7
[25] lifecycle_1.0.3   compiler_4.3.1    codetools_0.2-19  dplyr_1.1.2
[29] pkgconfig_2.0.3   rstudioapi_0.15.0 lattice_0.21-8    digest_0.6.33
[33] R6_2.5.1          tidyselect_1.2.0  utf8_1.2.3        pillar_1.9.0
[37] magrittr_2.0.3    tools_4.3.1       gtable_0.3.3      ggplot2_3.4.2
```

Hardware info

```
CPU model: Intel(R) Xeon(R) Gold 6134 CPU @ 3.20GHz

MemTotal:       247605664 kB
```