mkin - Routines for fitting kinetic models with one or more state variables to chemical degradation data

Johannes Ranke
Product Safety
Harlan Laboratories Ltd.
Zelgliweg 1, CH–4452 Itingen, Switzerland

May 11, 2010

Abstract

In the regulatory evaluation of chemical substances like plant protection products (pesticides), biocides and other chemicals, degradation data play an important role. For the evaluation of pesticide degradation experiments, detailed guidance has been developed, based on nonlinear optimisation. The R add-on package **mkin** implements fitting some of the models recommended in this guidance from within R and calculates some statistical measures for data series within one or more compartments, for parent and metabolites.

Contents

Key words: Kinetics, FOCUS, nonlinear optimisation

1 Introduction

Many approaches are possible regarding the evaluation of chemical degradation data. The **kinfit** package (Ranke, 2010a) in R (R Development Core Team, 2010) implements the approach recommended in the kinetics report provided by the FOrum for Co-ordination of pesticide fate models and their USe (FOCUS Work Group on Degradation Kinetics, 2006) for simple data series for one parent compound in one compartment.

The **mkin** package (Ranke, 2010b) extends this approach to data series with metabolites and more than one compartment and includes the possibility for back reactions.

2 Example

In the following, requirements for data formatting are explained. Then the procedure for fitting the four kinetic models recommended by the FOCUS group to an example dataset for parent only given in the FOCUS kinetics report is illustrated. The explanations are kept rather verbose in order to lower the barrier for R newcomers.

2.1 Data format

The following listing shows example dataset C from the FOCUS kinetics report as distributed with the **kinfit** package

```
R> library("mkin")
```

```
This product includes software developed by the University of Chicago, as
Operator of Argonne National Laboratory.

See the LICENSE file distributed with the minpack.lm source code or
http://www.netlib.org/minpack/disclaimer for the full license.
```

```
R> FOCUS_2006_C

    name time value
1 parent    0  85.1
2 parent    1  57.9
3 parent    3  29.9
4 parent    7  14.6
5 parent   14   9.7
6 parent   28   6.6
7 parent   63   4.0
8 parent   91   3.9
9 parent  119   0.6
```

Note that the data needs to be in the format of a data frame containing a variable `name` specifying the observed variable, indicating the compound name and, if applicable, the compartment, a variable `time` containing sampling times, and a numeric variable `value` specifying the observed value of the variable. If a further variable `error` is present, this will be used to give different weights to the data points (the higher the error, the lower the weight, see the help page of the `modCost` function of the **FME** package (Soetaert and Petzoldt, 2010)). Replicate measurements are not recorded in extra columns but simply appended, leading to multiple occurrences of the sampling times `time`.

Small to medium size dataset can be conveniently entered directly as R code as shown in the following listing

```
R> example_data <- data.frame(
+   time = c(0, 1, 3, 7, 14, 28, 63, 91, 119),
+   parent = c(85.1, 57.9, 29.9, 14.6, 9.7, 6.6, 4, 3.9, 0.6)
+ )
```

2.2 Model definition

The next task is to define the model to be fitted to the data. In order to facilitate this task, a convenience function `mkinmod` is available.

```
R> SFO <- mkinmod(spec = list(parent = list(type = "SFO", to = NA, sink = TRUE)))
R> SFORB <- mkinmod(spec = list(parent = list(type = "SFORB", to = NA, sink = TRUE)))
R> SFO_SFO <- mkinmod(spec = list(
+   parent = list(type = "SFO", to = "m1", sink = TRUE),
+   m1 = list(type = "SFO", to = NA, sink = TRUE)))
R> SFORB_SFO <- mkinmod(spec = list(
+   parent = list(type = "SFORB", to = "m1", sink = TRUE),
+   m1 = list(type = "SFO", to = NA, sink = TRUE)))
```

2.3 Fitting the model

Then the model parameters should be fitted to the data. The function `mkinfit` internally creates a cost function using `modCost` from the **FME** package and the produces a fit using `modFit` from the same package.

```
R> # Do not show significance stars as they interfere with vignette generation
R> options(show.signif.stars = FALSE)
R> SFO.fit <- mkinfit(SFO, FOCUS_2006_C)
R> summary(SFO.fit)
```

```
Parameters:
              Estimate Std. Error t value Pr(>|t|)
parent_0       82.4920     4.7402  17.403 5.09e-07
k_parent_sink   0.3061     0.0459   6.668 0.000286

Residual standard error: 5.299 on 7 degrees of freedom

Parameter correlation:
              parent_0 k_parent_sink
parent_0        1.0000        0.5212
k_parent_sink   0.5212        1.0000
```

```
R> SFORB.fit <- mkinfit(SFORB, FOCUS_2006_C)
R> summary(SFORB.fit)
```

```
Parameters:
                     Estimate Std. Error t value Pr(>|t|)
parent_free_0       84.999203   0.890977  95.400  2.4e-09
k_parent_free_sink   0.394914   0.014386  27.452  1.2e-06
k_parent_free_bound  0.061482   0.007503   8.195  0.00044
k_parent_bound_free  0.020687   0.004158   4.975  0.00419

Residual standard error: 0.9341 on 5 degrees of freedom

Parameter correlation:
                    parent_free_0 k_parent_free_sink k_parent_free_bound
parent_free_0             1.00000             0.5217              0.1834
k_parent_free_sink        0.52169             1.0000              0.6718
k_parent_free_bound       0.18342             0.6718              1.0000
k_parent_bound_free       0.08489             0.3249              0.6995
                    k_parent_bound_free
parent_free_0                   0.08489
k_parent_free_sink              0.32485
k_parent_free_bound             0.69952
k_parent_bound_free             1.00000
```

```
R> SFO_SFO.fit <- mkinfit(SFO_SFO, FOCUS_2006_D)
R> summary(SFO_SFO.fit)
```

```
Parameters:
               Estimate Std. Error t value Pr(>|t|)
parent_0      9.960e+01  1.614e+00  61.720  < 2e-16
k_parent_sink 4.792e-02  3.750e-03  12.777 6.10e-15
k_m1_sink     5.261e-03  7.159e-04   7.349 1.15e-08
k_parent_m1   5.078e-02  2.094e-03  24.248  < 2e-16

Residual standard error: 3.211 on 36 degrees of freedom

Parameter correlation:
              parent_0 k_parent_sink k_m1_sink k_parent_m1
parent_0       1.00000       0.60752   -0.1701    -0.06623
k_parent_sink  0.60752       1.00000   -0.6252    -0.08738
k_m1_sink     -0.17005      -0.62525    1.0000     0.47166
k_parent_m1   -0.06623      -0.08738    0.4717     1.00000
```

```
R> SFO_SFO.fit.2 <- mkinfit(SFO_SFO, FOCUS_2006_D,
+   fixed_initials = c(FALSE, FALSE), fixed_parms = c(FALSE, TRUE, FALSE))
R> summary(SFO_SFO.fit.2)
```

```
Parameters:
              Estimate Std. Error t value Pr(>|t|)
parent_0      87.10345    7.46450  11.669 8.61e-14
m1_0          -8.07507    8.47214  -0.953  0.34688
k_parent_sink -0.05968    0.02031  -2.939  0.00571
k_parent_m1    0.11444    0.02213   5.170 8.93e-06

Residual standard error: 17.12 on 36 degrees of freedom

Parameter correlation:
              parent_0    m1_0 k_parent_sink k_parent_m1
parent_0        1.0000 -0.0960        0.4248     -0.1427
m1_0           -0.0960  1.0000        0.3812     -0.4521
k_parent_sink   0.4248  0.3812        1.0000     -0.8601
k_parent_m1    -0.1427 -0.4521       -0.8601      1.0000
```

```
R> SFO_SFO.fit.3 <- mkinfit(SFO_SFO, FOCUS_2006_D,
+   fixed_initials = c(FALSE, FALSE), fixed_parms = c(FALSE, TRUE, FALSE), lower = -0
R> summary(SFO_SFO.fit.3)
```

```
Parameters:
                Estimate Std. Error   t value Pr(>|t|)
parent_0      99.0370138  8.8487379     11.19 2.82e-13
m1_0          -0.0000001  9.2865451 -1.08e-08  1.00000
k_parent_sink -0.0000001  0.0252081 -3.97e-06  1.00000
k_parent_m1    0.0826239  0.0246628      3.35  0.00191

Residual standard error: 18.31 on 36 degrees of freedom
```

```
Parameter correlation:
              parent_0     m1_0 k_parent_sink k_parent_m1
parent_0       1.00000 -0.07706        0.4227    -0.03923
m1_0          -0.07706  1.00000        0.3834    -0.50708
k_parent_sink  0.42273  0.38335        1.0000    -0.71281
k_parent_m1   -0.03923 -0.50708       -0.7128     1.00000
```

```
R> SFORB_SFO.fit <- mkinfit(SFORB_SFO, FOCUS_2006_D)
R> summary(SFORB_SFO.fit)
```

```
Parameters:
                     Estimate Std. Error t value Pr(>|t|)
parent_free_0       1.011e+02  2.020e+00  50.034  < 2e-16
k_parent_free_sink  6.408e-02  2.691e-02   2.381   0.0230
k_parent_free_bound 1.680e-01  5.142e-01   0.327   0.7458
k_parent_bound_free 5.239e-01  8.544e-01   0.613   0.5438
k_m1_sink           5.213e-03  7.210e-04   7.230 2.29e-08
k_parent_free_m1    6.563e-02  2.542e-02   2.582   0.0143

Residual standard error: 3.219 on 34 degrees of freedom

Parameter correlation:
                    parent_free_0 k_parent_free_sink k_parent_free_bound
parent_free_0              1.0000             0.5432             0.34390
k_parent_free_sink         0.5432             1.0000             0.94317
k_parent_free_bound        0.3439             0.9432             1.00000
k_parent_bound_free        0.1950             0.8179             0.95432
k_m1_sink                 -0.1801            -0.2031            -0.09286
k_parent_free_m1           0.4402             0.9752             0.96052
                    k_parent_bound_free k_m1_sink k_parent_free_m1
parent_free_0                   0.19504  -0.18007          0.44016
k_parent_free_sink              0.81787  -0.20307          0.97518
k_parent_free_bound             0.95432  -0.09286          0.96052
k_parent_bound_free             1.00000  -0.08809          0.83987
k_m1_sink                      -0.08809   1.00000         -0.03944
k_parent_free_m1                0.83987  -0.03944          1.00000
```

References

FOCUS Work Group on Degradation Kinetics. *Guidance Document on Estimating Persistence and Degradation Kinetics from Environmental Fate Studies on Pesticides in EU Registration. Report of the FOCUS Work Group on Degradation Kinetics*, 2006. URL http://focus.jrc.ec.europa.eu/dk. EC Document Reference Sanco/10058/2005 version 2.0.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. URL http://www.R-project.org. ISBN 3-900051-07-0.

Johannes Ranke. *kinfit: Routines for fitting simple kinetic models to chemical degradation data*, 2010a. URL http://CRAN.R-project.org.

Johannes Ranke. *mkin: Routines for fitting kinetic models with one or more state variables to chemical degradation data*, 2010b. URL http://CRAN.R-project.org.

Karline Soetaert and Thomas Petzoldt. Inverse modelling, sensitivity and monte carlo analysis in R using package FME. *Journal of Statistical Software*, 33(3):1–28, 2010. URL http://www.jstatsoft.org/v33/i03/.