Examples for kinetic evaluations using mkin

Johannes Ranke

Eurofins Regulatory AG
Weidenweg 15, CH–4310 Rheinfelden, Switzerland

and

University of Bremen

February 18, 2013

Contents

Key words: Kinetics, FOCUS, nonlinear optimisation

1 Kinetic evaluations for parent compounds

These examples are also evaluated in a parallel vignette of the **kinfit** package (**?**). The datasets are from Appendix 3, of the FOCUS kinetics report (**??**).

1.1 Laboratory Data L1

The following code defines example dataset L1 from the FOCUS kinetics report, p. 284

```
R> library("mkin")
R> FOCUS_2006_L1 = data.frame(
+   t = rep(c(0, 1, 2, 3, 5, 7, 14, 21, 30), each = 2),
+   parent = c(88.3, 91.4, 85.6, 84.5, 78.9, 77.6,
+              72.0, 71.9, 50.3, 59.4, 47.0, 45.1,
+              27.7, 27.3, 10.0, 10.4, 2.9, 4.0))
R> FOCUS_2006_L1_mkin <- mkin_wide_to_long(FOCUS_2006_L1)
```

The next step is to set up the models used for the kinetic analysis. Note that the model definitions contain the names of the observed variables in the data. In this case, there is only one variable called `parent`.

```
R> SFO <- mkinmod(parent = list(type = "SFO"))
R> FOMC <- mkinmod(parent = list(type = "FOMC"))
R> DFOP <- mkinmod(parent = list(type = "DFOP"))
```

The three models cover the first assumption of simple first order (SFO), the case of declining rate constant over time (FOMC) and the case of two different phases of the kinetics (DFOP). For a more detailed discussion of the models, please see the FOCUS kinetics report.

The following two lines fit the model and produce the summary report of the model fit. This covers the numerical analysis given in the FOCUS report.

```
R> m.L1.SFO <- mkinfit(SFO, FOCUS_2006_L1_mkin, quiet=TRUE)
R> summary(m.L1.SFO)
```

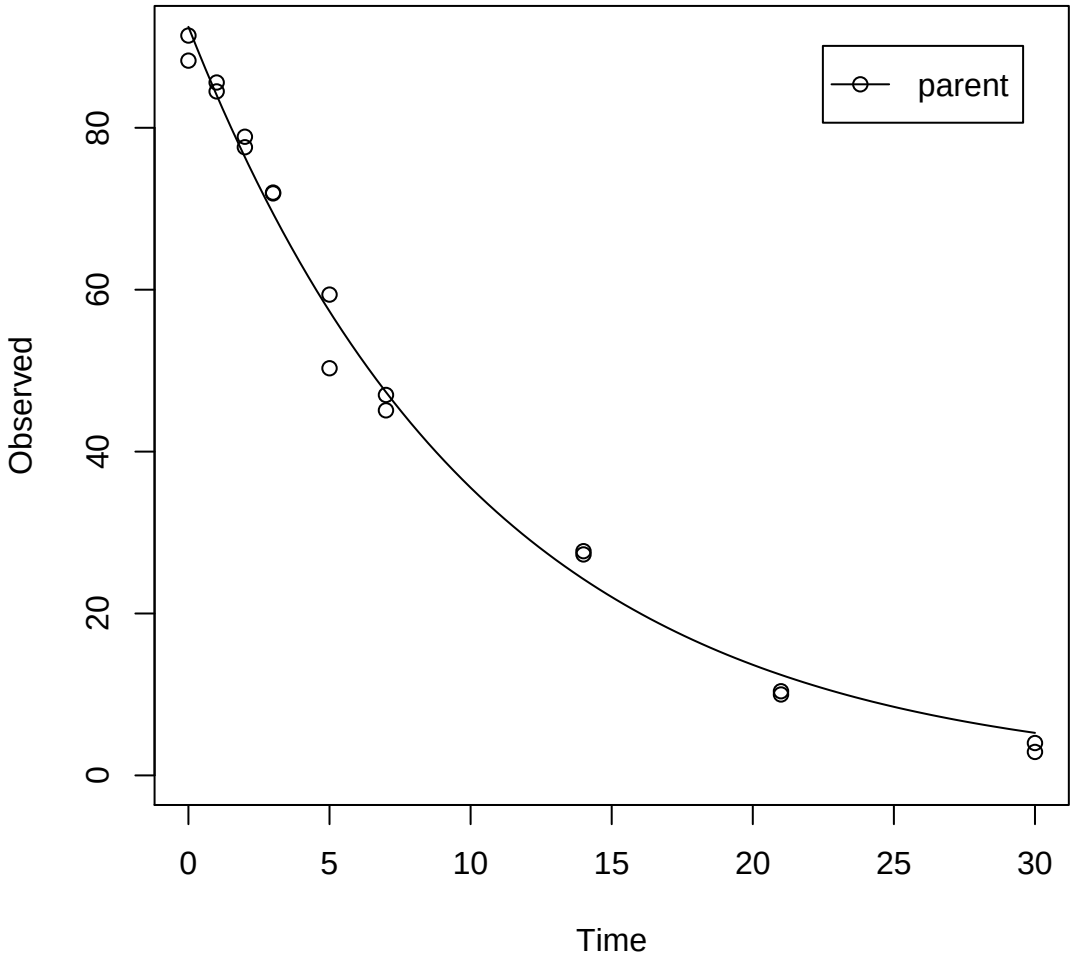
```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:41 2013
```

```
Date of summary: Mon Feb 18 09:52:41 2013

Equations:
[1] d_parent = - k_parent_sink * parent

Starting values for optimised parameters:
              initial   type transformed
parent_0        100.0  state  100.000000
k_parent_sink     0.1 deparm   -2.302585

Fixed parameter values:
None

Optimised, transformed parameters:
              Estimate Std. Error
parent_0        92.471      1.368
k_parent_sink   -2.347      0.041

Backtransformed parameters:
              Estimate
parent_0        92.471
k_parent_sink    0.096

Residual standard error: 2.948 on 16 degrees of freedom

Chi2 error levels in percent:
         err.min n.optim df
All data   3.424       2  7
parent     3.424       2  7

Estimated disappearance times:
        DT50  DT90
parent 7.249 24.08

Estimated formation fractions:
            ff
parent_sink  1

Parameter correlation:
              parent_0 k_parent_sink
parent_0        1.0000        0.6248
k_parent_sink   0.6248        1.0000

Data:
 time variable observed predicted residual
    0   parent     88.3    92.471  -4.1710
    0   parent     91.4    92.471  -1.0710
    1   parent     85.6    84.039   1.5610
    1   parent     84.5    84.039   0.4610
```

```
 2   parent     78.9    76.376   2.5241
 2   parent     77.6    76.376   1.2241
 3   parent     72.0    69.412   2.5884
 3   parent     71.9    69.412   2.4884
 5   parent     50.3    57.330  -7.0301
 5   parent     59.4    57.330   2.0699
 7   parent     47.0    47.352  -0.3515
 7   parent     45.1    47.352  -2.2515
14   parent     27.7    24.247   3.4527
14   parent     27.3    24.247   3.0527
21   parent     10.0    12.416  -2.4163
21   parent     10.4    12.416  -2.0163
30   parent      2.9     5.251  -2.3513
30   parent      4.0     5.251  -1.2513
```

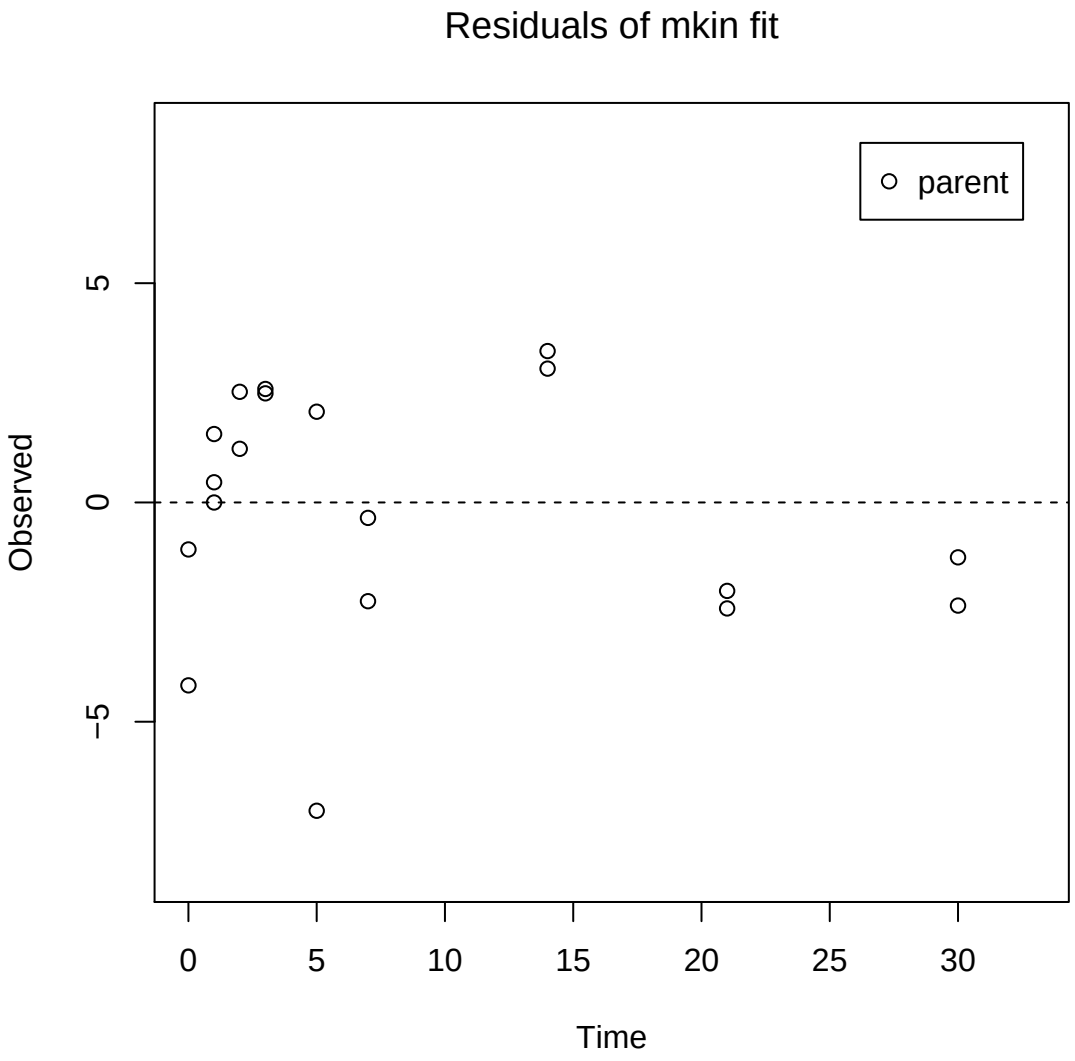
A plot of the fit is obtained with the plot function for mkinfit objects.

```
R> plot(m.L1.SFO)
```



The residual plot can be easily obtained by

```
R> mkinresplot(m.L1.SFO, ylab = "Observed", xlab = "Time")
```



For comparison, the FOMC model is fitted as well, and the χ^2 error level is checked.

```
R> m.L1.FOMC <- mkinfit(FOMC, FOCUS_2006_L1_mkin, quiet=TRUE)
R> s.m.L1.FOMC <- summary(m.L1.FOMC)
R> s.m.L1.FOMC$errmin
```

```
            err.min n.optim df
All data 0.03618911       3  6
parent   0.03618911       3  6
```

Due to the higher number of parameters, and the lower number of degrees of freedom of the fit, the χ^2 error level is actually higher for the FOMC model (3.6%) than for the SFO model (3.4%).

1.2 Laboratory Data L2

The following code defines example dataset L2 from the FOCUS kinetics report, p. 287

```
R> FOCUS_2006_L2 = data.frame(
+   t = rep(c(0, 1, 3, 7, 14, 28), each = 2),
+   parent = c(96.1, 91.8, 41.4, 38.7,
+              19.3, 22.3, 4.6, 4.6,
+              2.6, 1.2, 0.3, 0.6))
R> FOCUS_2006_L2_mkin <- mkin_wide_to_long(FOCUS_2006_L2)
```

Again, the SFO model is fitted and a summary is obtained.

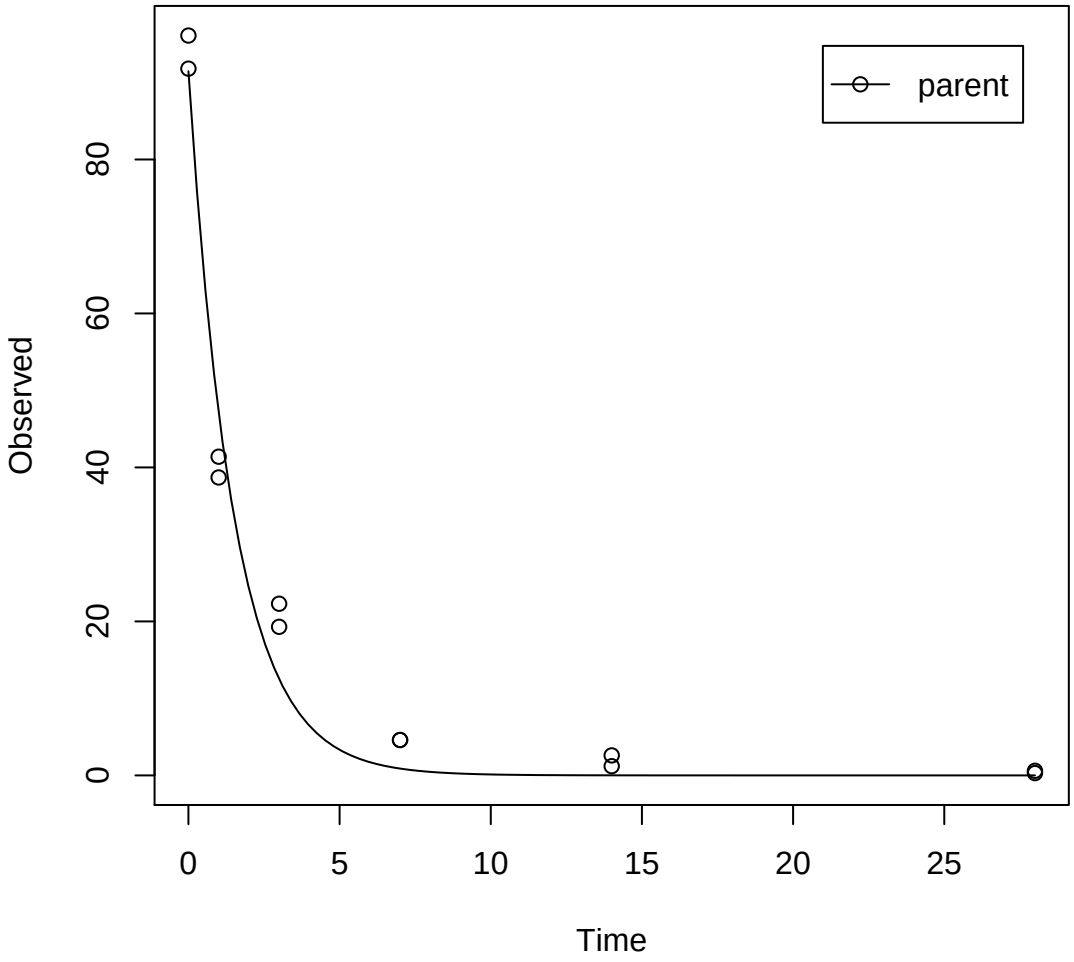
```
R> m.L2.SFO <- mkinfit(SFO, FOCUS_2006_L2_mkin, quiet=TRUE)
R> summary(m.L2.SFO)
```

```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:42 2013
Date of summary: Mon Feb 18 09:52:42 2013

Equations:
[1] d_parent = - k_parent_sink * parent

Starting values for optimised parameters:
              initial   type transformed
parent_0        100.0  state  100.000000
k_parent_sink     0.1 deparm   -2.302585

Fixed parameter values:
None

Optimised, transformed parameters:
              Estimate Std. Error
parent_0       91.4656      3.807
k_parent_sink  -0.4112      0.107

Backtransformed parameters:
              Estimate
parent_0        91.466
k_parent_sink    0.663

Residual standard error: 5.51 on 10 degrees of freedom

Chi2 error levels in percent:
         err.min n.optim df
All data   14.38       2  4
parent     14.38       2  4
```

```
Estimated disappearance times:
        DT50  DT90
parent 1.046 3.474

Estimated formation fractions:
            ff
parent_sink  1

Parameter correlation:
              parent_0 k_parent_sink
parent_0        1.0000        0.4295
k_parent_sink   0.4295        1.0000

Data:
 time variable observed     predicted residual
    0   parent     96.1 91.4656079103   4.6344
    0   parent     91.8 91.4656079103   0.3344
    1   parent     41.4 47.1395280371  -5.7395
    1   parent     38.7 47.1395280371  -8.4395
    3   parent     19.3 12.5210295280   6.7790
    3   parent     22.3 12.5210295280   9.7790
    7   parent      4.6  0.8833842647   3.7166
    7   parent      4.6  0.8833842647   3.7166
   14   parent      2.6  0.0085318162   2.5915
   14   parent      1.2  0.0085318162   1.1915
   28   parent      0.3  0.0000007958   0.3000
   28   parent      0.6  0.0000007958   0.6000
```

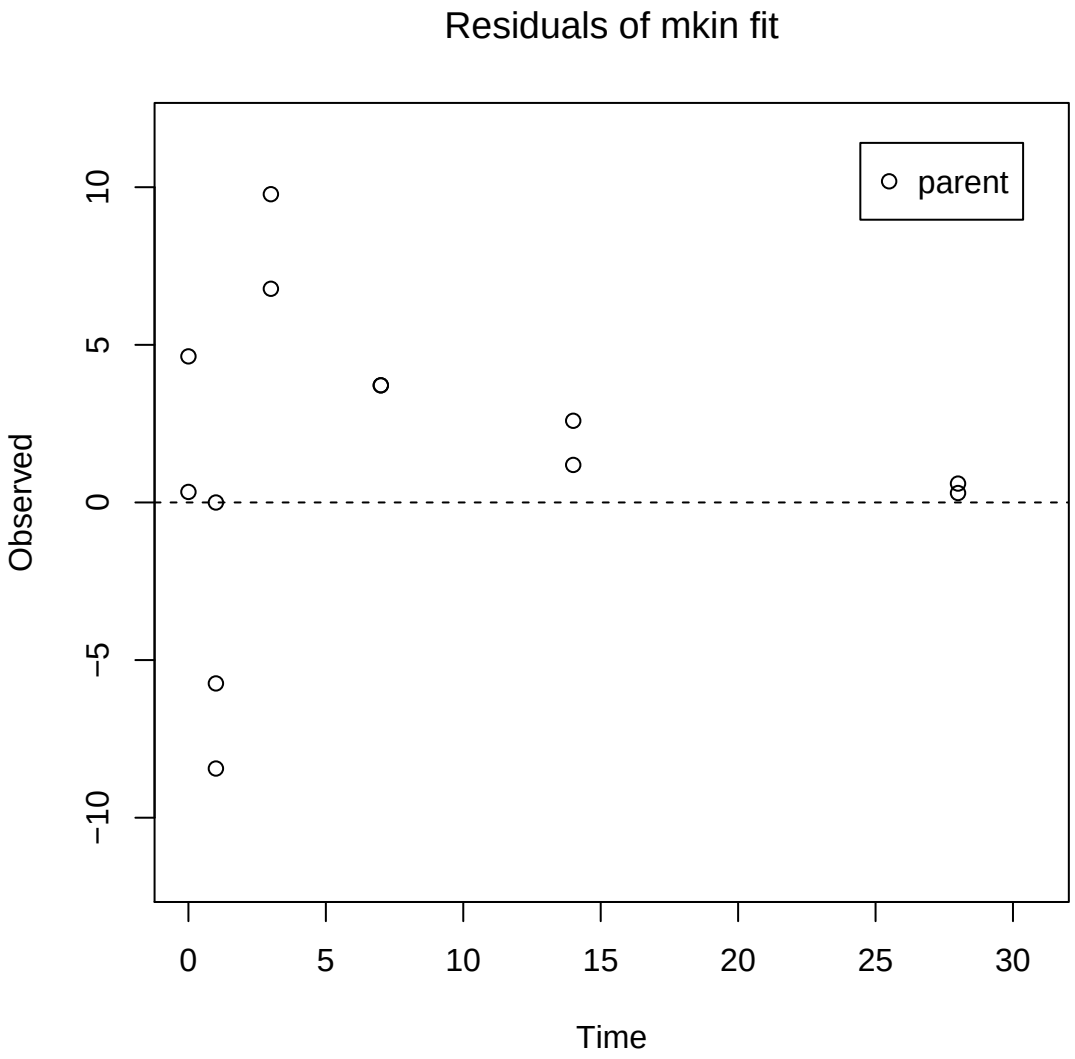
The χ^2 error level of 14% suggests that the model does not fit very well. This is also obvious from the plots of the fit and the residuals.

```
R> plot(m.L2.SFO)
```



In the FOCUS kinetics report, it is stated that there is no apparent systematic error observed from the residual plot up to the measured DT90 (approximately at day 5), and there is an underestimation beyond that point.

```
R> mkinresplot(m.L2.SFO, ylab = "Observed", xlab = "Time")
```

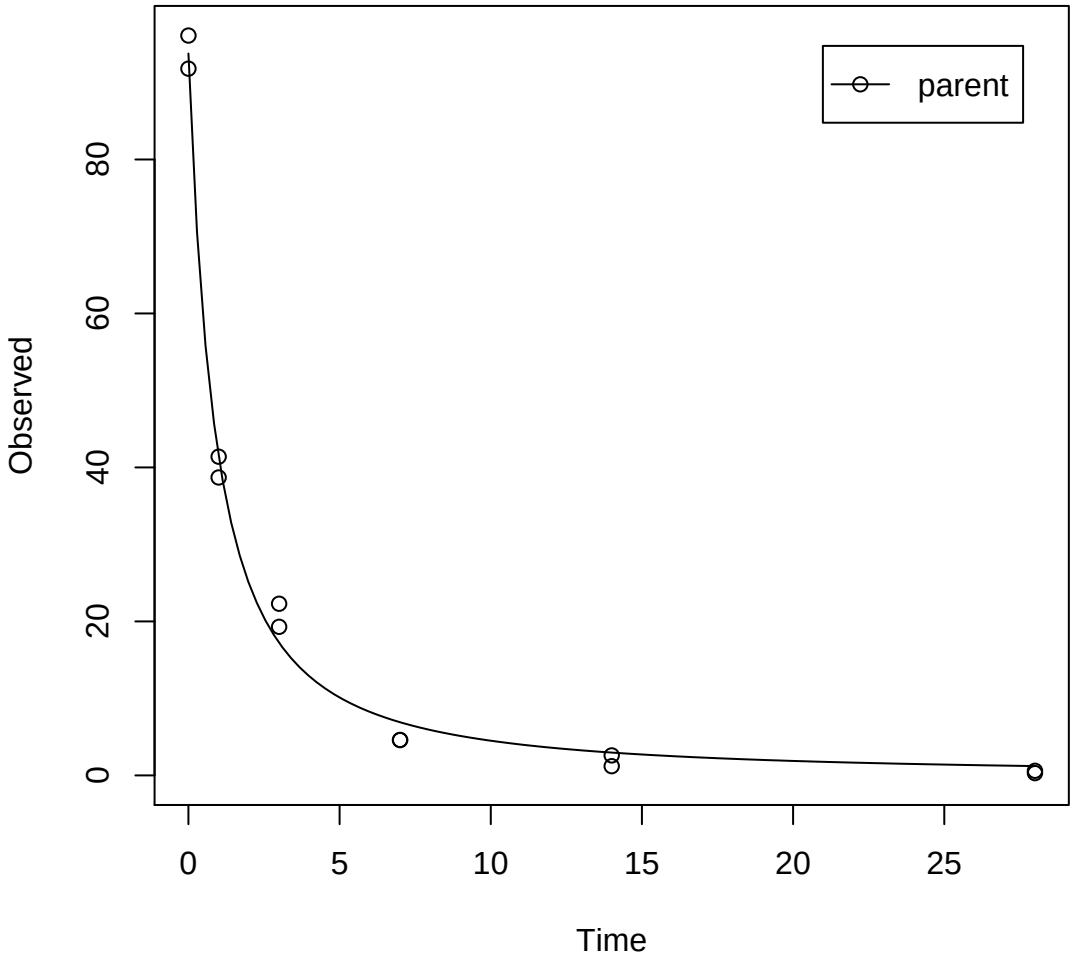


We may add that it is difficult to judge the random nature of the residuals just from the three samplings at days 0, 1 and 3. Also, it is not clear why a consistent underestimation after the approximate DT90 should be irrelevant.

For comparison, the FOMC model is fitted as well, and the χ^2 error level is checked.

```
R> m.L2.FOMC <- mkinfit(FOMC, FOCUS_2006_L2_mkin, quiet=TRUE)
R> plot(m.L2.FOMC)
R> s.m.L2.FOMC <- summary(m.L2.FOMC)
R> s.m.L2.FOMC$errmin
```

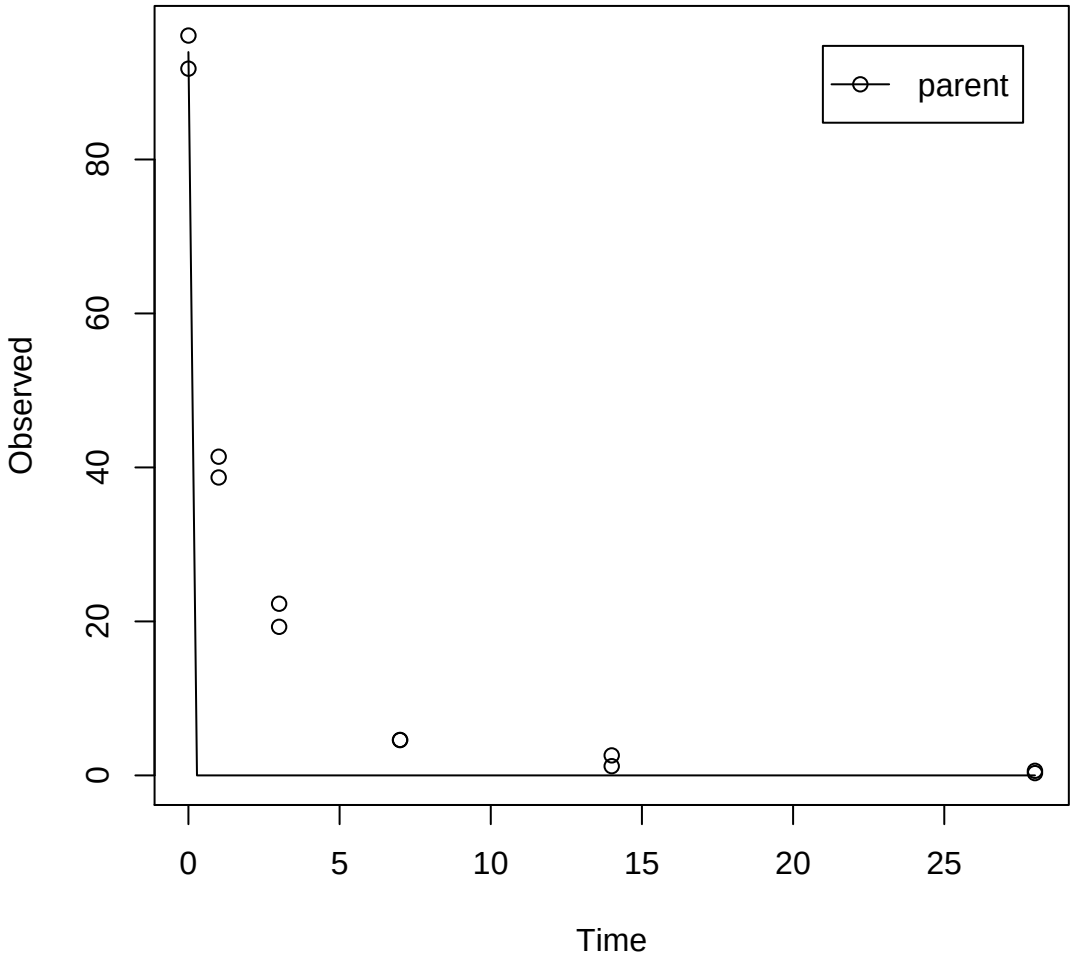
```
            err.min n.optim df
All data 0.06204245       3  3
parent   0.06204245       3  3
```



The error level at which the χ^2 test passes is much lower in this case. Therefore, the FOMC model provides a better description of the data, as less experimental error has to be assumed in order to explain the data.

Fitting the four parameter DFOP model does not further reduce the χ^2 error level.

```
R> m.L2.DFOP <- mkinfit(DFOP, FOCUS_2006_L2_mkin, quiet=TRUE)
R> plot(m.L2.DFOP)
```



Here, the default starting parameters for the DFOP model obviously do not lead to a reasonable solution. Therefore the fit is repeated with different starting parameters.

```
R> m.L2.DFOP <- mkinfit(DFOP, FOCUS_2006_L2_mkin,
+   parms.ini = c(k1 = 1, k2 = 0.01, g = 0.8),
+   quiet=TRUE)
R> plot(m.L2.DFOP)
R> summary(m.L2.DFOP)
```

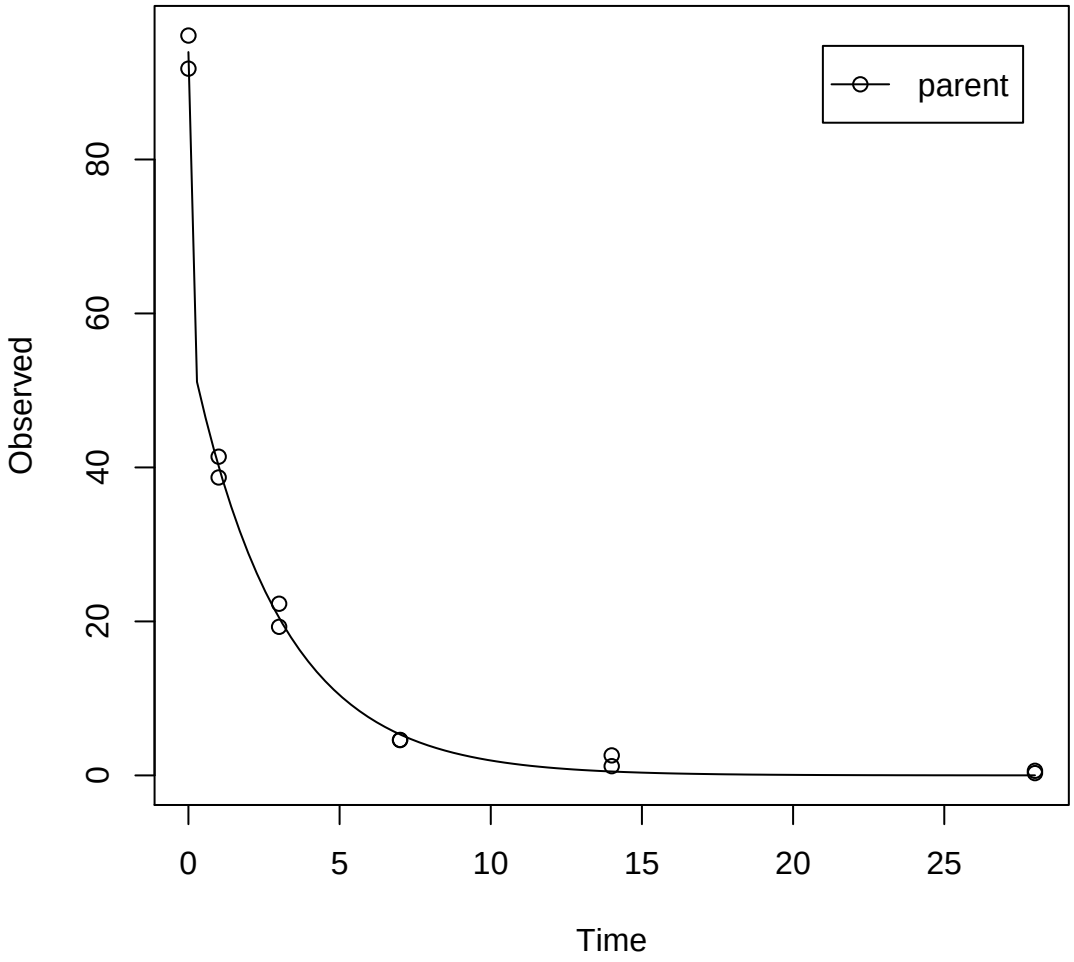
```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:43 2013
Date of summary: Mon Feb 18 09:52:43 2013
```

```
Equations:
[1] d_parent = - ((k1 * g * exp(-k1 * time) + k2 * (1 - g) * exp(-k2 * time)) / (g * exp

Starting values for optimised parameters:
         initial   type transformed
parent_0   1e+02  state 100.0000000
k1         1e+00 deparm   0.0000000
k2         1e-02 deparm  -4.6051702
g          8e-01 deparm   0.9802581

Fixed parameter values:
None

Optimised, transformed parameters:
         Estimate Std. Error
parent_0  93.9500         NA
k1         4.9589         NA
k2        -1.0880         NA
g         -0.2821         NA

Backtransformed parameters:
         Estimate
parent_0   93.950
k1        142.434
k2          0.337
g           0.402

Residual standard error: 1.732 on 8 degrees of freedom

Chi2 error levels in percent:
         err.min n.optim df
All data   2.529       4  2
parent     2.529       4  2

Estimated disappearance times:
       DT50 DT90
parent   NA   NA

Estimated formation fractions:
[1] ff
<0 rows> (or 0-length row.names)

Data:
 time variable observed predicted residual
    0   parent     96.1 93.950000   2.1500
    0   parent     91.8 93.950000  -2.1500
    1   parent     41.4 40.143423   1.2566
    1   parent     38.7 40.143423  -1.4434
```

```
    3   parent     19.3 20.464500  -1.1645
    3   parent     22.3 20.464500   1.8355
    7   parent      4.6  5.318322  -0.7183
    7   parent      4.6  5.318322  -0.7183
   14   parent      2.6  0.503070   2.0969
   14   parent      1.2  0.503070   0.6969
   28   parent      0.3  0.004501   0.2955
   28   parent      0.6  0.004501   0.5955
```

```
R> s.m.L2.DFOP <- summary(m.L2.DFOP)
R> s.m.L2.DFOP$errmin
```

```
            err.min n.optim df
All data 0.02528763       4  2
parent   0.02528763       4  2
```



Therefore, the FOMC model is clearly the best-fit model based on the χ^2 error level criterion.

1.3 Laboratory Data L3

The following code defines example dataset L3 from the FOCUS kinetics report, p. 290

```
R> FOCUS_2006_L3 = data.frame(
+   t = c(0, 3, 7, 14, 30, 60, 91, 120),
+   parent = c(97.8, 60, 51, 43, 35, 22, 15, 12))
R> FOCUS_2006_L3_mkin <- mkin_wide_to_long(FOCUS_2006_L3)
```

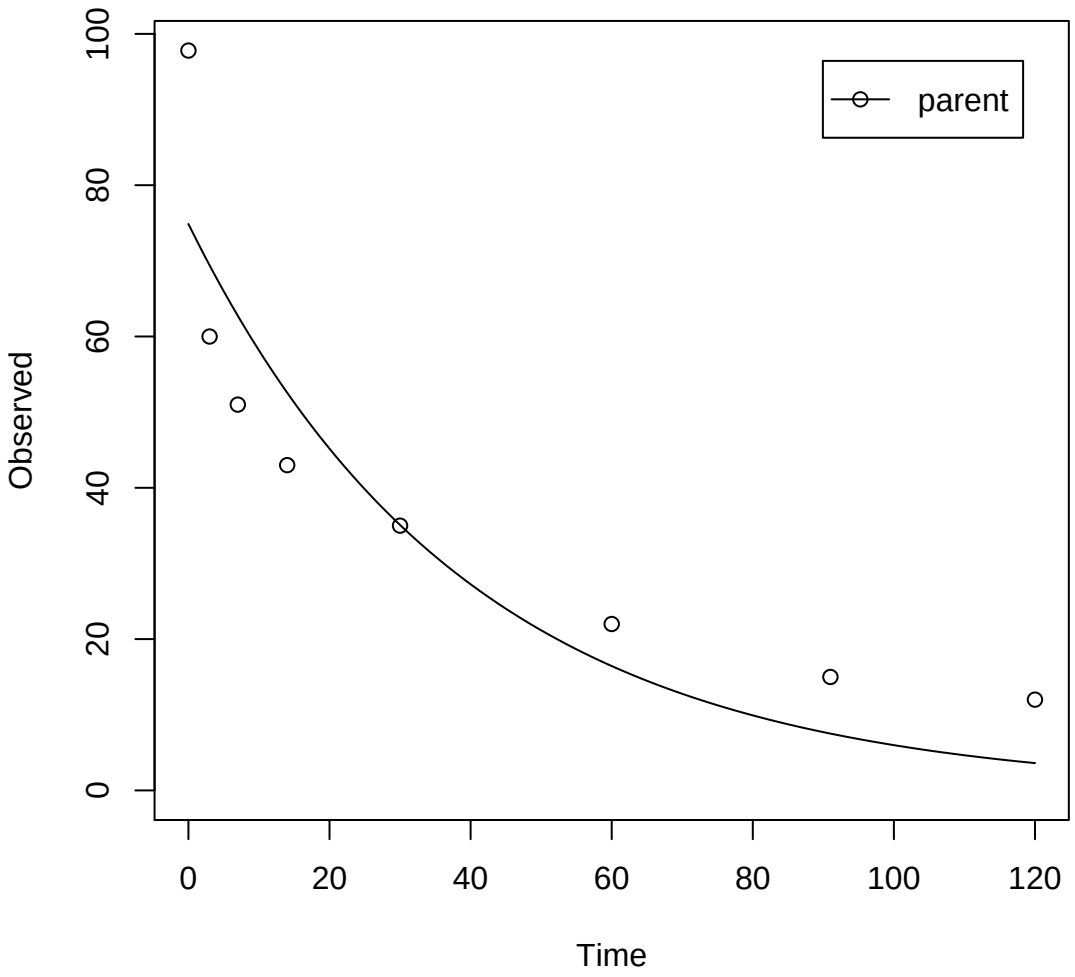
SFO model, summary and plot:

```
R> m.L3.SFO <- mkinfit(SFO, FOCUS_2006_L3_mkin, quiet=TRUE)
R> summary(m.L3.SFO)
```

```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:43 2013
Date of summary: Mon Feb 18 09:52:43 2013

Equations:
[1] d_parent = - k_parent_sink * parent

Starting values for optimised parameters:
              initial   type transformed
parent_0        100.0  state  100.000000
k_parent_sink     0.1 deparm   -2.302585

Fixed parameter values:
None

Optimised, transformed parameters:
              Estimate Std. Error
parent_0        74.873      8.458
k_parent_sink   -3.678      0.326

Backtransformed parameters:
              Estimate
parent_0        74.873
k_parent_sink    0.025

Residual standard error: 12.91 on 6 degrees of freedom

Chi2 error levels in percent:
         err.min n.optim df
All data   21.24       2  6
parent     21.24       2  6

Estimated disappearance times:
        DT50  DT90
parent 27.43 91.12

Estimated formation fractions:
            ff
parent_sink  1

Parameter correlation:
              parent_0 k_parent_sink
parent_0        1.0000        0.5484
```

```
k_parent_sink   0.5484        1.0000

Data:
 time variable observed predicted  residual
    0   parent     97.8    74.873  22.92734
    3   parent     60.0    69.407  -9.40654
    7   parent     51.0    62.734 -11.73403
   14   parent     43.0    52.563  -9.56336
   30   parent     35.0    35.083  -0.08281
   60   parent     22.0    16.439   5.56137
   91   parent     15.0     7.510   7.48961
  120   parent     12.0     3.609   8.39083
```

```
R> plot(m.L3.SFO)
```



The χ^2 error level of 22% as well as the plot suggest that the model does not fit very well.

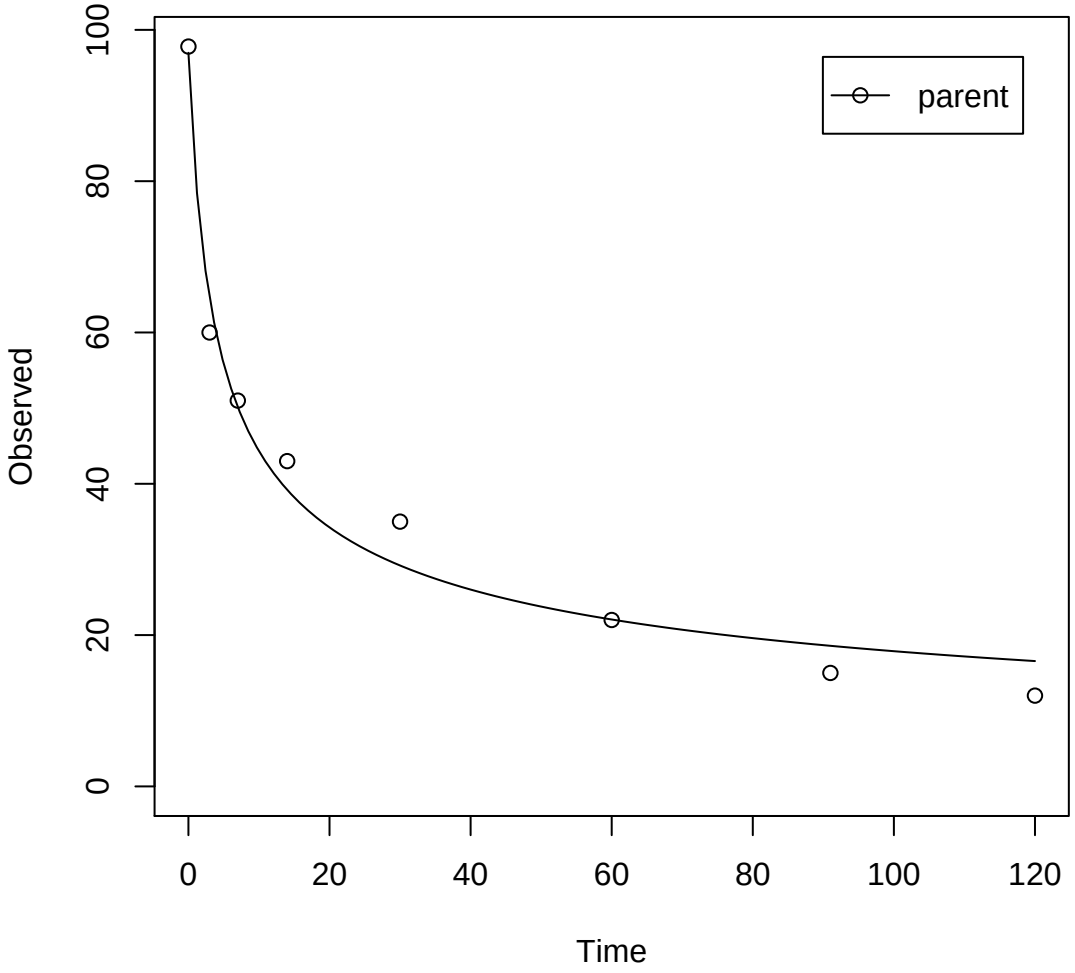
The FOMC model performs better:

```
R> m.L3.FOMC <- mkinfit(FOMC, FOCUS_2006_L3_mkin, quiet=TRUE)
R> plot(m.L3.FOMC)
R> s.m.L3.FOMC <- summary(m.L3.FOMC)
R> s.m.L3.FOMC$errmin
```

```
            err.min n.optim df
All data 0.07321867       3  5
parent   0.07321867       3  5
```

```
R> endpoints(m.L3.FOMC)
```

```
$distimes
           DT50     DT90
parent 7.729478 431.2428

$ff
logical(0)

$SFORB
logical(0)
```

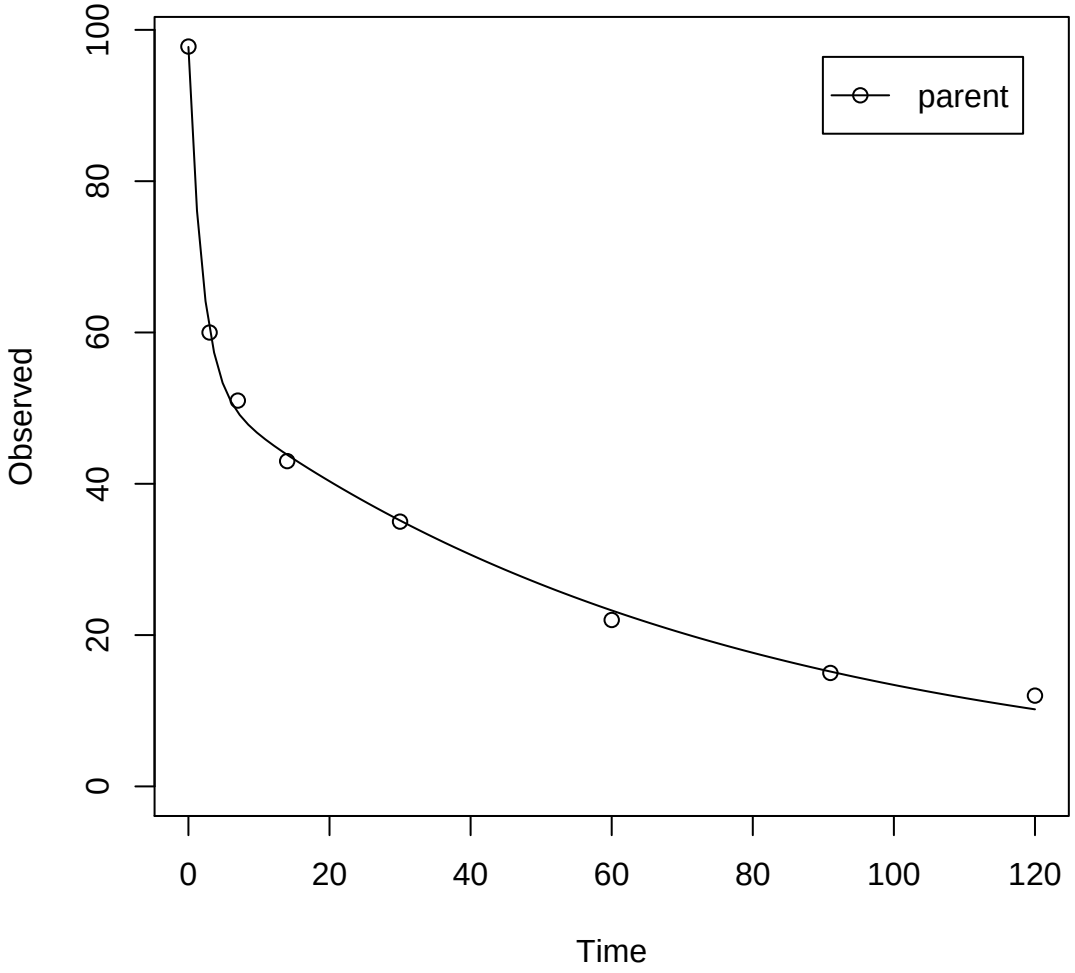


The error level at which the χ^2 test passes is 7% in this case.

Fitting the four parameter DFOP model further reduces the χ^2 error level considerably:

```
R> m.L3.DFOP <- mkinfit(DFOP, FOCUS_2006_L3_mkin, quiet=TRUE)
R> plot(m.L3.DFOP)
R> s.m.L3.DFOP <- summary(m.L3.DFOP)
R> s.m.L3.DFOP$errmin
```

```
            err.min n.optim df
All data 0.02223992       4  4
parent   0.02223992       4  4
```



Therefore, the DFOP model is the best-fit model based on the χ^2 error level criterion for laboratory data L3.

1.4 Laboratory Data L4

The following code defines example dataset L4 from the FOCUS kinetics report, p. 293

```
R> FOCUS_2006_L4 = data.frame(
+   t = c(0, 3, 7, 14, 30, 60, 91, 120),
+   parent = c(96.6, 96.3, 94.3, 88.8, 74.9, 59.9, 53.5, 49.0))
R> FOCUS_2006_L4_mkin <- mkin_wide_to_long(FOCUS_2006_L4)
```

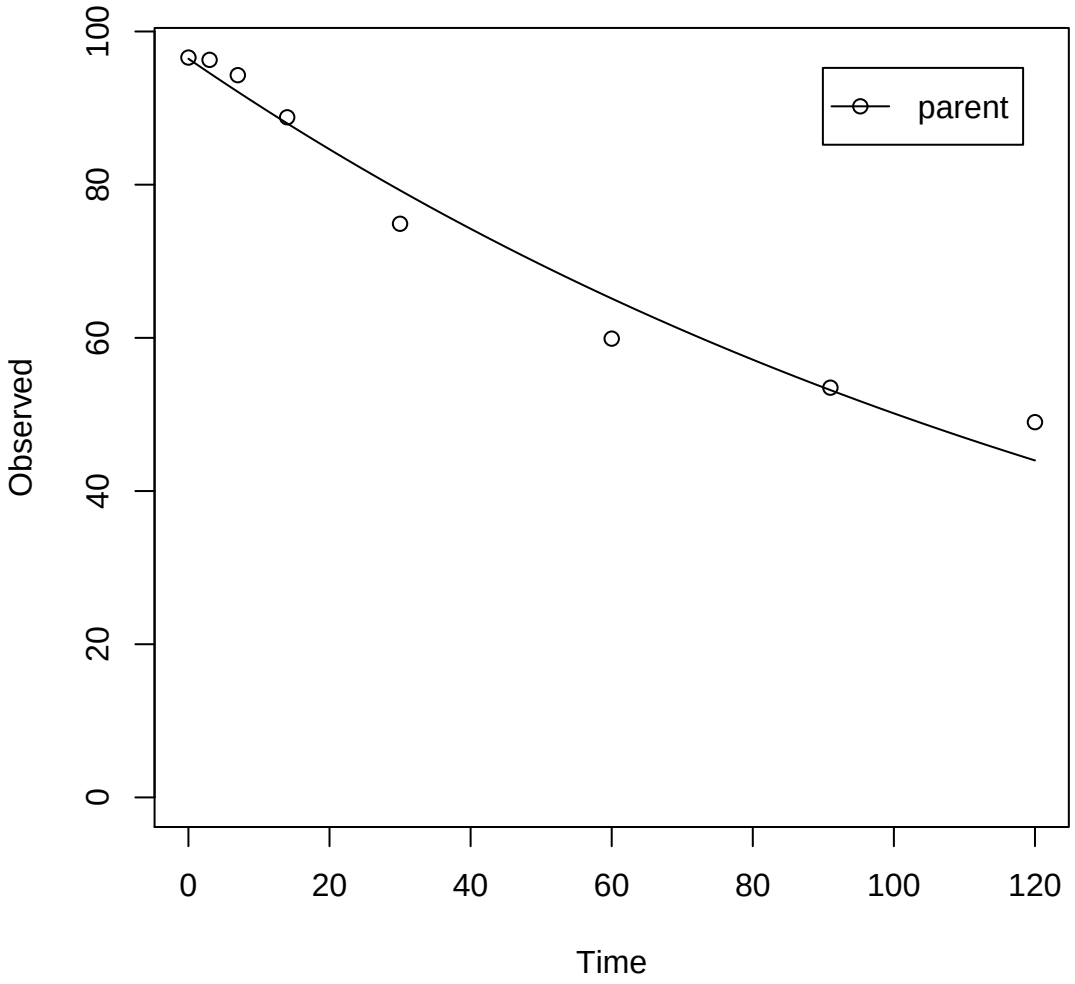
SFO model, summary and plot:

```
R> m.L4.SFO <- mkinfit(SFO, FOCUS_2006_L4_mkin, quiet=TRUE)
R> summary(m.L4.SFO)
```

```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:44 2013
Date of summary: Mon Feb 18 09:52:44 2013

Equations:
[1] d_parent = - k_parent_sink * parent

Starting values for optimised parameters:
              initial   type transformed
parent_0        100.0  state  100.000000
k_parent_sink     0.1 deparm   -2.302585

Fixed parameter values:
None

Optimised, transformed parameters:
              Estimate Std. Error
parent_0         96.44      1.949
k_parent_sink    -5.03      0.080

Backtransformed parameters:
              Estimate
parent_0        96.442
k_parent_sink    0.007

Residual standard error: 3.651 on 6 degrees of freedom

Chi2 error levels in percent:
         err.min n.optim df
All data   3.288       2  6
parent     3.288       2  6

Estimated disappearance times:
       DT50 DT90
parent  106  352

Estimated formation fractions:
            ff
parent_sink  1

Parameter correlation:
              parent_0 k_parent_sink
parent_0        1.0000        0.5865
```

```
k_parent_sink   0.5865        1.0000

Data:
 time variable observed predicted residual
    0   parent     96.6     96.44   0.1585
    3   parent     96.3     94.57   1.7324
    7   parent     94.3     92.13   2.1744
   14   parent     88.8     88.00   0.7972
   30   parent     74.9     79.26  -4.3589
   60   parent     59.9     65.14  -5.2376
   91   parent     53.5     53.18   0.3167
  120   parent     49.0     43.99   5.0054
```

```
R> plot(m.L4.SFO)
```

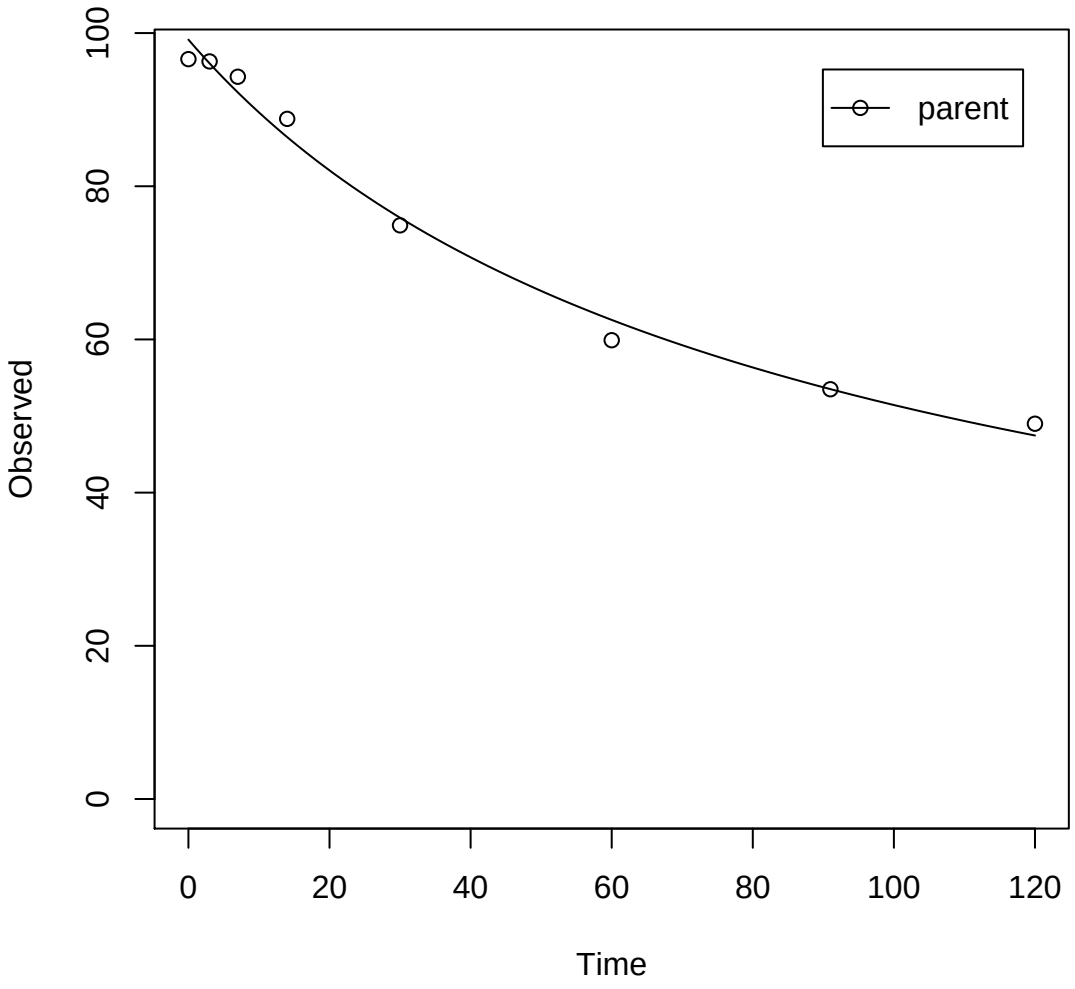


The χ^2 error level of 3.3% as well as the plot suggest that the model fits very well.

The FOMC model for comparison

```
R> m.L4.FOMC <- mkinfit(FOMC, FOCUS_2006_L4_mkin, quiet=TRUE)
R> plot(m.L4.FOMC)
R> s.m.L4.FOMC <- summary(m.L4.FOMC)
R> s.m.L4.FOMC$errmin
```

```
            err.min n.optim df
All data 0.02027643       3  5
parent   0.02027643       3  5
```



The error level at which the χ^2 test passes is slightly lower for the FOMC model. However, the difference appears negligible.

2 Kinetic evaluations for parent and metabolites

2.1 Laboratory Data for example compound Z

The following code defines the example dataset from Appendix 7 to the FOCUS kinetics report, p.350

```
R> LOD = 0.5
R> FOCUS_2006_Z = data.frame(
+   t = c(0, 0.04, 0.125, 0.29, 0.54, 1, 2, 3, 4, 7, 10, 14, 21, 42, 61, 96, 124),
+   Z = c(100, 81.7, 70.4, 51.1, 41.2, 6.6, 4.6, 3.9, 4.6, 4.3, 6.8, 2.9, 3.5,
+             5.3, 4.4, 1.2, 0.7),
+   Z1 = c(0, 18.3, 29.6, 46.3, 55.1, 65.7, 39.1, 36, 15.3, 5.6, 1.1, 1.6, 0.6,
+          0.5 * LOD, NA, NA, NA),
+   Z2 = c(0, NA, 0.5 * LOD, 2.6, 3.8, 15.3, 37.2, 31.7, 35.6, 14.5, 0.8, 2.1,
+          1.9, 0.5 * LOD, NA, NA, NA),
+   Z3 = c(0, NA, NA, NA, NA, 0.5 * LOD, 9.2, 13.1, 22.3, 28.4, 32.5, 25.2, 17.2,
+          4.8, 4.5, 2.8, 4.4))
R> FOCUS_2006_Z_mkin <- mkin_wide_to_long(FOCUS_2006_Z)
```

The next step is to set up the models used for the kinetic analysis. As the simultaneous fit of parent and the first metabolite is usually straightforward, Step 1 (SFO for parent only) is skipped here. We start with the model 2a, with formation and decline of metabolite Z1 and the pathway from parent directly to sink included (default in mkin).

```
R> Z.2a <- mkinmod(Z = list(type = "SFO", to = "Z1"),
+                  Z1 = list(type = "SFO"))
R> Z.2a.ff <- mkinmod(Z = list(type = "SFO", to = "Z1"),
+                  Z1 = list(type = "SFO"), use_of_ff = "max")
R> m.Z.2a <- mkinfit(Z.2a, FOCUS_2006_Z_mkin)
```

```
Model cost at call  1 :  30553.16
Model cost at call  4 :  30553.16
Model cost at call  5 :  30553.16
Model cost at call  7 :  9763.132
Model cost at call  8 :  9763.132
Model cost at call  12 :  7280.781
Model cost at call  13 :  7280.781
Model cost at call  17 :  1167.09
Model cost at call  18 :  1167.09
```

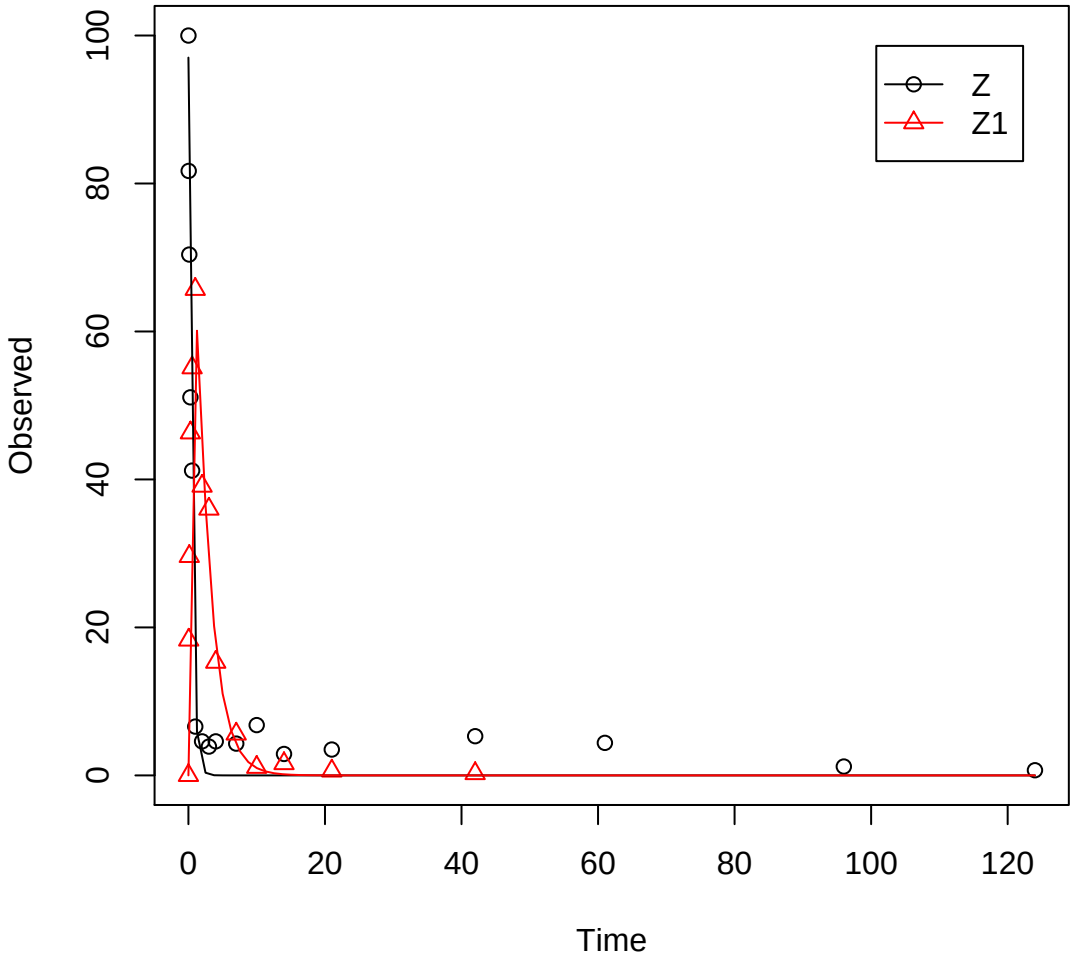
```
Model cost at call  22 :  772.958
Model cost at call  23 :  772.958
Model cost at call  27 :  700.265
Model cost at call  28 :  700.265
Model cost at call  32 :  692.3945
Model cost at call  33 :  692.3945
Model cost at call  37 :  692.3839
Model cost at call  38 :  692.3839
Model cost at call  42 :  692.3838
Model cost at call  47 :  692.3838
```

```
R> summary(m.Z.2a, data = FALSE)
```

```
mkin version:    0.9.12
R version:       2.15.2
Date of fit:     Mon Feb 18 09:52:45 2013
Date of summary: Mon Feb 18 09:52:45 2013

Equations:
[1] d_Z = - k_Z_sink * Z - k_Z_Z1 * Z    d_Z1 = + k_Z_Z1 * Z - k_Z1_sink * Z1

Starting values for optimised parameters:
          initial   type transformed
Z_0         100.0  state  100.000000
k_Z_sink      0.1 deparm   -2.302585
k_Z_Z1        0.1 deparm   -2.302585
k_Z1_sink     0.1 deparm   -2.302585

Fixed parameter values:
   value  type
Z1     0 state

Optimised, transformed parameters:
          Estimate Std. Error
Z_0        97.0149         NA
k_Z_sink  -36.2915         NA
k_Z_Z1      0.8047         NA
k_Z1_sink  -0.7296         NA

Backtransformed parameters:
          Estimate
Z_0         97.015
k_Z_sink     0.000
k_Z_Z1       2.236
k_Z1_sink    0.482

Residual standard error: 5.064 on 27 degrees of freedom

Chi2 error levels in percent:
```

```
          err.min n.optim df
All data   18.20       4 27
Z          18.56       4 13
Z1         15.75       1 13

Estimated disappearance times:
    DT50   DT90
Z  0.255 0.8471
Z1 1.438 4.7759

Estimated formation fractions:
               ff
Z_sink  6.375e-17
Z_Z1    8.226e-01
Z1_sink 1.000e+00
```

```
R> plot(m.Z.2a)
```



As obvious from the summary, the kinetic rate constant from parent compound Z to sink is negligible. Accordingly, the exact magnitude of the fitted parameter `log k_Z_sink`

If the same model is fitted using

```
R> Z.2b <- mkinmod(Z = list(type = "SFO", to = "Z1", sink = FALSE),
+                  Z1 = list(type = "SFO"))
R> Z.2b.ff <- mkinmod(Z = list(type = "SFO", to = "Z1", sink = FALSE), # needs warning
+                  Z1 = list(type = "SFO"), use_of_ff = "max")
R>
```

Accordingly, the exact magnitude of the fitted parameter (ilr transformed) is not defined and the covariance matrix is not returned.