Fitting dose-response curves from bioassays and toxicity testing

by Johannes Ranke

In attempts to quantify detrimental effects of chemical substances on biological organisms, the responses of these organisms to different doses[1] are often observed in a quantitative way. Examples of the observed variables (endpoints of toxicity) are the number of surviving individuals, the length of wheat seedlings after being exposed to different concentrations of the chemical substance for a defined time interval after germination, the luminescence of bacteria, the ability of cell cultures to reduce a specific dye, the number of viable offspring and many others.

These observed variables have in common that a reference magnitude for healthy and viable organisms can be defined (normalised response level $r = 1$), and that the magnitude of the variable (response) is limited by a zero response ($r = 0$) where the maximum of the effect is observed. In between, there is either a continuum of possible responses, or, if only the presence/absence of a specific quality of n individual organisms is observed (commonly life/death), then only $1/n$ response levels are possible.

Such dose-response relationships can generally be expressed as

$$r = f(d, \vec{p}) + \epsilon \qquad (1)$$

where r is the normalised response at dose d, $f(d, \vec{p})$ is the model function with parameter vector $\vec{p}$, and ϵ is the error variable describing the variability in the observations not explainable by the model function $f(d, \vec{p})$.

This article shows, how different model functions $f(d, \vec{p})$ can be conveniently fitted to such dose-response data using the R package **drfit**, yielding the vector of parameters $\vec{p}$ that gives the description of the data with the least residual error. The fitting can be carried out for many substances with a single call to the main function `drfit`.

The main results that the user will probably be interested in are the doses, at which a response of 50 % relative to healthy control organisms is to be expected (termed ED_{50}). These are listed together with the standard error of the fitted `nlsModel`.

The R package **drfit** internally uses the R function `nls` from the recommended **stats** package for nonlinear least squares modelling. **drfit** provides a function to produce different types of plots of the data and the fitted curves. Optionally, the raw data can be kept in an external database connected by **RODBC**. This has proven to be useful if the data of a large number of

dose-response experiments have to be evaluated as e.g. in bioassays based on microtiter plates.

Recently, the R package **drc** containing similar functionalities to **drfit** has been uploaded to CRAN. Unfortunately, I have noticed the existence of this package only during the preparation of this article, after having maintained **drfit** on CRAN for almost one and a half years. Maybe in the future it will be possible to join forces.

Collecting dose-response data

The `drfit` function expects the dose-response data as a data frame containing at least a factor called 'substance', a vector called 'unit' containing the unit used for the dose, a column 'response' with the response values of the test system normalised between 0 and 1 and a column 'dose' with the numeric dose values. An example of such data can be easily obtained

```
> library(drfit)
> data(antifoul)
> antifoul[1:12,
+ c('substance','dose','unit','response')]
      substance        dose  unit      response
1  Zn Pyrithion 5.0000000    muM   0.051453400
2  Zn Pyrithion 2.5000000    muM   0.023387900
3  Zn Pyrithion 1.2500000    muM   0.047444000
4  Zn Pyrithion 0.6250000    muM   0.091546900
5  Zn Pyrithion 0.3125000    muM   0.071500200
6  Zn Pyrithion 0.1562500    muM   0.873371000
7  Zn Pyrithion 0.0781250    muM   1.005680000
8  Zn Pyrithion 0.0390625    muM   0.893418000
9  Zn Pyrithion 0.0195312    muM   1.041760000
10 Zn Pyrithion 5.0000000    muM  -0.000668226
11 Zn Pyrithion 2.5000000    muM   0.067490800
12 Zn Pyrithion 1.2500000    muM   0.079518900
```

Normalisation of the response data is not done within the **drfit** package. It can either be carried out with a typical spreadsheet file, with some extra lines of R code, or by an external procedure, while/before the data is read into a database.

If the data is collected and normalised using MS Excel, it can be easily transferred to R by saving it in CSV format, and reading it in using the R function `read.csv2`. If OpenOffice.org Calc is being used, and the default values are used for exporting the data in CSV format, the function `read.csv` is very helpful.

Figure 1 shows a possible spreadsheet layout for capturing dose-response data including the observed

---

[1] The term dose is used here in a generalised way, referring to doses in the strict sense like mg oral intake per kg body weight as well as to measured concentrations in aquatic toxicity tests or nominal concentrations in cell culture assays.

endpoint (number of fronds in this case) and the response value calculated by dividing the the number of fronds by the mean value of the appropriate control experiments.

Figure 1: Data structure for a typical toxicity test in OpenOffice Calc. Note that the response column is calculated (see text).

If the spreadsheet from Figure 1 (provided in recent versions of drfit > 0.04-54) would be exported by writing a CSV file, this file could be read processed by something like

```
> d <- read.csv('sampledata.csv',skip=2,dec=',')
```

depending on the path to the CSV file, the number of lines before the column headings and the decimal separator used.

Fitting the curves

A quick result for a compatible dataframe can usually be obtained by a simple call to `drfit`

```
> r <- drfit(antifoul)
```

The contents of the dataframe r containing the results of the fitting procedure are shown and explained in Figure 2.

Plotting the data

Once the drfit function has been successfully called and the result assigned a name (`r` in this case), it is fairly easy to create dose-response plots for the fitted data. The following example produces a single graph (`overlay=TRUE`) with the fitted dose-response curves and raw data with standard deviation bars for all substances and fitted models in dataframes `d` and `r` using color (`bw=FALSE`) in the vector based EPS graphics format as shown in Figure 3

```
> d <- antifoul
> drplot(r,d,overlay=T,bw=F,postscript=T)
Created File:  ./drplot.eps
```
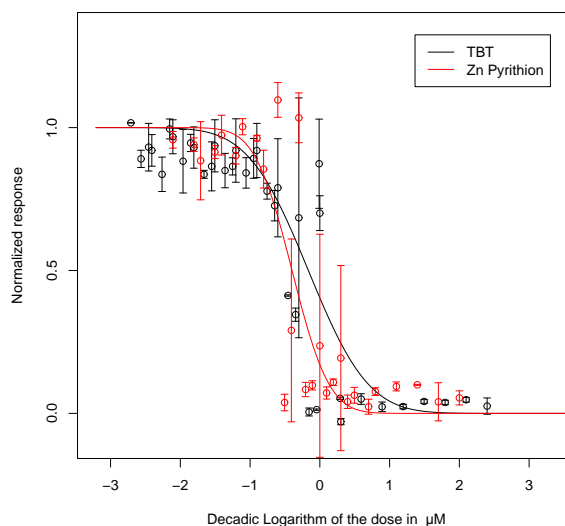

Figure 3: Output of the `drplot` function for the sample data from the package.

The interested reader is encouraged to explore the more refined details of the **drfit** package by studying the help files.

```
\bibliography{drfit-Rnews}
```

Johannes Ranke
Department of Bioorganic Chemistry
UFT Center for Environmental Research and Technology
University of Bremen `jranke@uni-bremen.de`

```
print(r,digits=2)
     Substance ndl n  lld lhd  mtype logED50   std unit sigma     a    b
1          TBT  38 4 -2.7 2.4 probit   -0.16 0.053  muM  0.19 -0.16 0.68
2 Zn Pyrithion  27 3 -2.1 2.0 probit   -0.40 0.054  muM  0.23 -0.40 0.42
```

Figure 2: Contents of the dataframe containing the results from the fitting procedure for the example data from the package