To T-Test or Not to T-Test, That Is the Question

Johannes Ranke[1], René Lehmann[2]

[1]Wissenschaftlicher Berater - johannes.ranke@jrwb.de
[2]Umweltbundesamt Dessau, Germany

Introduction

In the report of the FOCUS workgroup on degradation kinetics (FOCUS, 2014), the t-test for significant difference of parameter estimates from zero is listed among the recommended methods and introduced with the statement (p. 96):

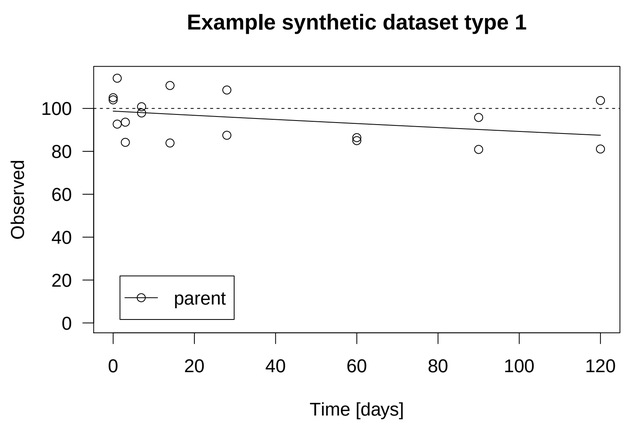
'If the parameters are normally distributed, ... '

In linear regression, when assuming normally distributed errors in y, parameter estimators are normally distributed as well. However, in the case of fitting degradation models, the model functions are nonlinear and parameter estimators can not be assumed to be normally distributed.
Therefore, this application of the t-test is questionable.

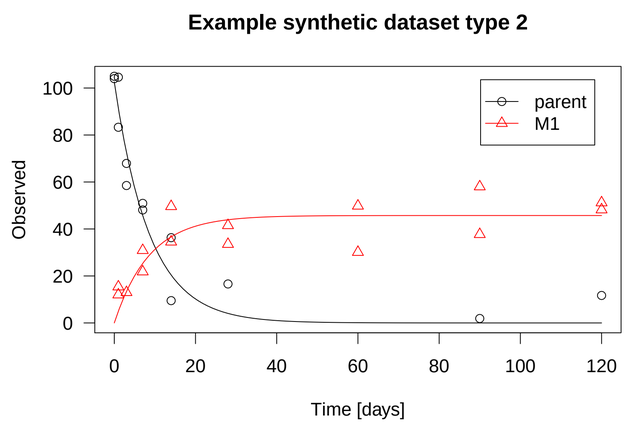
The purpose of this poster is to discuss reasons for using best-fit parameters for modelling instead of ignoring parameter estimates when the t-test fails to show significant difference from zero.

Materials and Methods

Generation of synthetic datasets, fitting of degradation models and plotting of results was done using R and the R package `mkin` (Ranke, 2015).
Datasets were generated with an initial concentration of the parent compound of 100. In **dataset type 1**, the parent compound degrades with simple first order (SFO) kinetics and a half-life of 1000 days. In **dataset type 2**, metabolite M1 degrades with SFO kinetics and the same half-life of 1000 days.
In both types of synthetic datasets, a homoscedastic error term, drawn from a normal distribution with an absolute standard deviation of 10 was added. Resulting values below the assumed LOD of 0.1 were set to `NA`.
1000 instances of each dataset type were generated and the corresponding kinetic models were fitted to each of them.



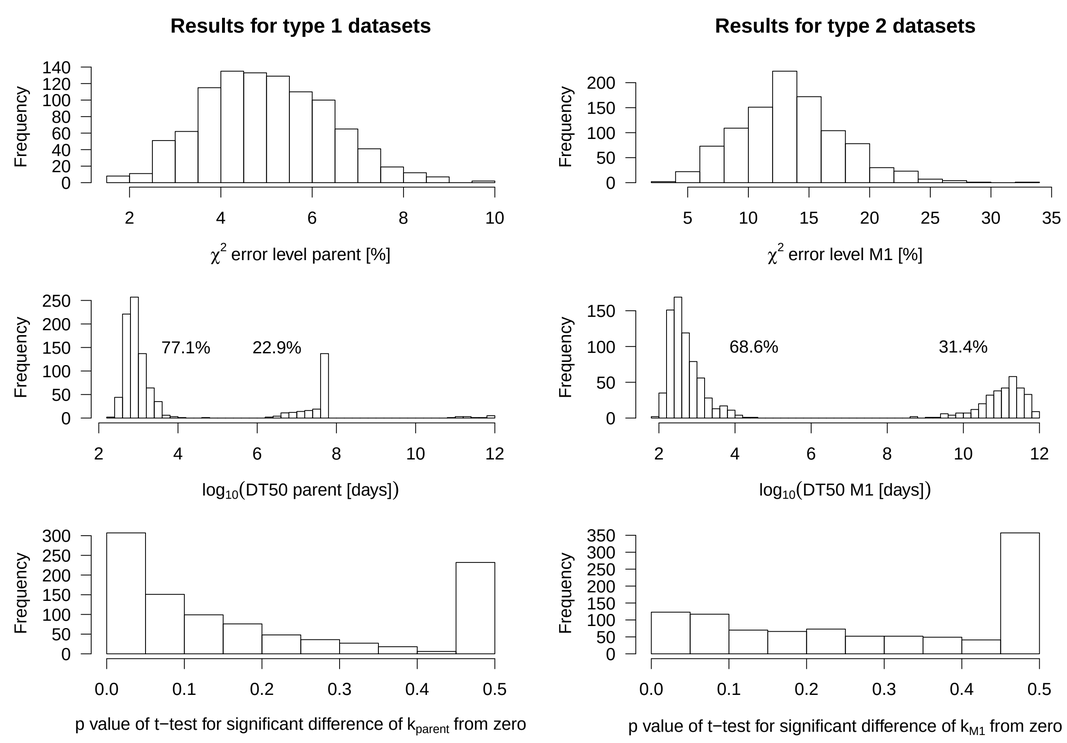
In dataset type 2, metabolite M1 reaches a level of about 50. This means that the relative error of M1 towards later time points is about 20%.



The degradation models were fitted with internal transformation of rate parameters (Bates and Watts, 1988, Ranke and Lehmann, 2012).

Results and Discussions

Type 1 datasets. The histogram of χ^2 error levels for the parent compound obtained by fitting the SFO model to the type 1 datasets shows relative error levels generally smaller than the error term used for data generation which amounts to a relative error of about 10%. This is because part of the variability around the supposed model is compensated by fitting the parameters.
The histogram for the parent DT50 on the log scale shows a bimodal distribution, where 77.1% of the results appear to approximately follow a lognormal distribution around the input DT50 of 10^3 days, a second mode around 10^8 days caused by numerical accuracy with the cases where degradation is masked by the random disturbance terms, and some (1.4%) extreme values $> 10^{10}$ days. The histogram of the p values shows that it is smaller than 0.05 in only 30% of the cases.
Type 2 datasets. The histogram of χ^2 error levels for metabolite M1 shows relative errors around 15%, which was aimed at in order to be around the border of acceptability proposed by the FOCUS workgroup.
The distribution of the fitted DT50 values of metabolite M1 is also bimodal. 68.6% appear to roughly follow a lognormal distribution around the input DT50. In 31.4% of the cases, degradation of M1 is masked by the random error term. The corresponding very low DT50 values form a second mode whose location and shape depends on the convergence tolerance used in the numerical fitting routine. Here, less than 15% of the datasets (cf. the height of the leftmost bar in the histogram) pass the t-test for significant difference from zero.



Conclusion

It is shown above that for small degradation rates, the distribution of the estimated DT50 values is bimodal, with a significant fraction of the estimated rate constants being very small and limited only by numerical aspects. It is also illustrated that in many cases where the parent or a metabolite degrade slowly, the recommended t-test does not show significant degradation. While discarding such datasets may introduce a bias towards faster degradation, it appears to be conservative to use best-fit values.

References

FOCUS (2014) Generic Guidance for Estimating Persistence and Degradation Kinetics from Environmental Fate Studies on Pesticides in EU Registration. Version 1.1
Ranke J (2015) mkin: Routines for Fitting Kinetic Models with One or More State Variables to Chemical Degradation Data. R package version 0.9-40 http://kinfit.r-forge.r-project.org/mkin_static
Bates DM, Watts DG (1988) Nonlinear regression analysis and its applications. Wiley-Interscience
Ranke J and Lehmann R (2012) Parameter reliability in kinetic evaluation of environmental metabolism data - Assessment and the influence of model specification. SETAC World 20-24 May 2012, Berlin, Germany