General Unified Threshold Model of Survival (GUTS-3S) Version 1.0

User Manual

Authors

Judith Klein

Dr. Udo Hommen

Fraunhofer Institute for Molecular Biology and Applied Ecology (IME)

Auf dem Aberg 1

57392 Schmallenberg, Germany

Date

21 June 2018

45 pages

Content

1. Introduction

This manual describes version 1.0 of the user friendly computer program GUTS-3S, which stands for "General Unified Threshold Model of Survival – 3 Steps". The program is developed by Udo Hommen and Judith Klein. The program is originally based on the Jager et al. 2011.

GUTS-3S is a user-friendly stand-alone software covering all necessary steps for its use in risk assessment, i.e., as a tier 2 tool according to the aquatic guidance document. The program is part of the ring test (Ashauer & Jager 2018).

The implementation can be used for three steps of analysis:

1. Calibration of the substance and species specific parameters using results of ecotoxicological test(s),
2. Verification of the calibrated model based on results of additional test(s), and
3. Application, prediction of effects (survival over time) for exposure scenarios not tested e.g. concentration profiles of focus scenarios, or LC50 values for a given time period of constant exposure.

GUTS-3S predicts survival of an aquatic organism in dependence of a concentration of an active substance.

2. Installing GUTS-3S

GUTS-3S is provided as an installation package to create an executable local program file plus some example data sets. Open the file setup.exe and follow the instructions to install GUTS-3S. Please select first the language for the setup and decide on a few details of the installation:

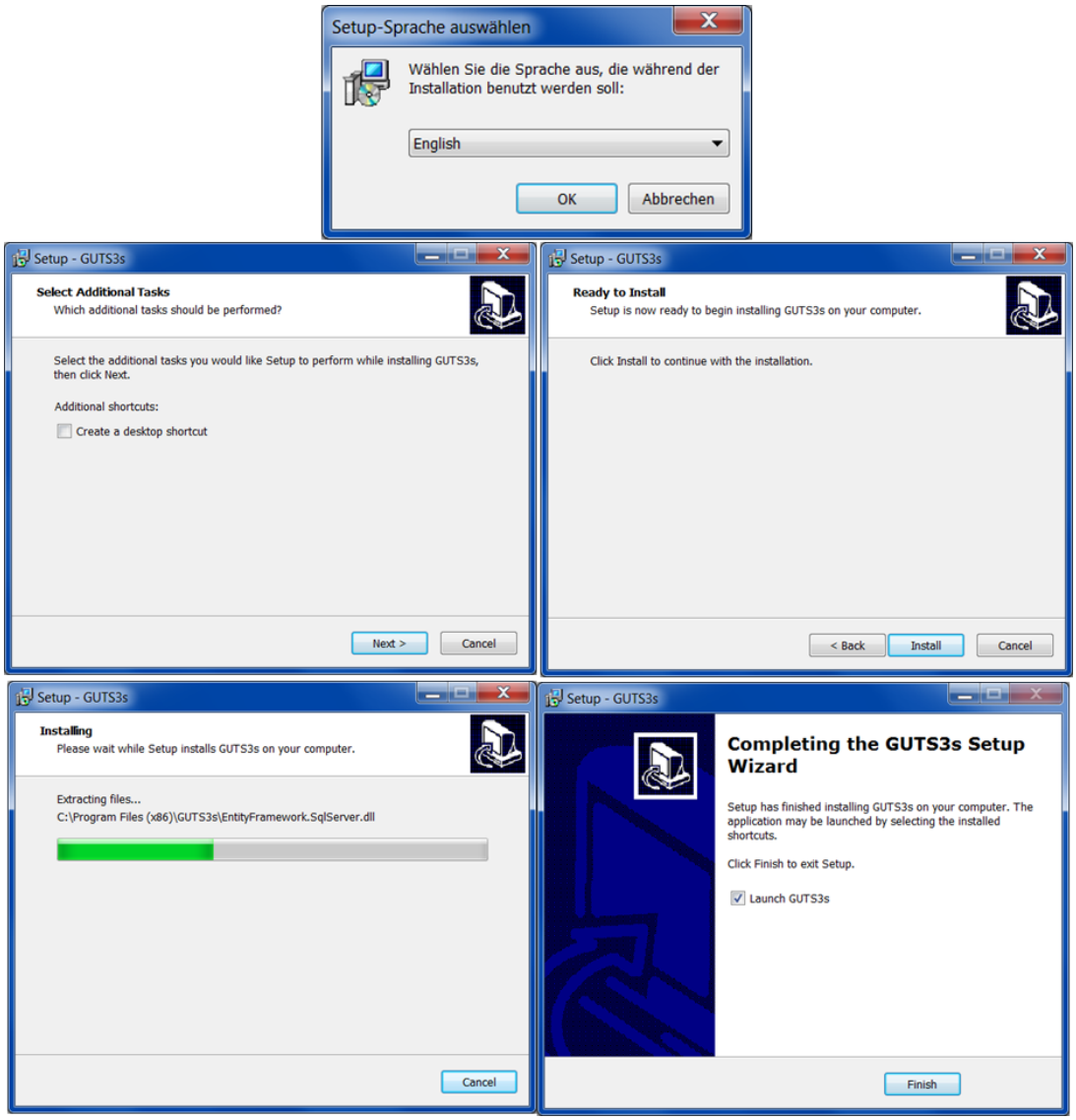


Figure 1: Installing the program

At the end of the setup, you can decide to launch GUTS-3S directly of or exit the program (Figure 1).

3. Working with GUTS-3S

The program offers three different procedures:

1. Calibration of the substance and species specific parameters using results of ecotoxicological test(s),
2. Verification of the calibrated model based on results of additional test(s), and
3. Prediction of effects (survival over time) for exposure scenarios not tested e.g. concentration profiles of focus scenarios, or LC50 values for a given time period of constant exposure.

In Figure 2, the start screen of the program is presented.

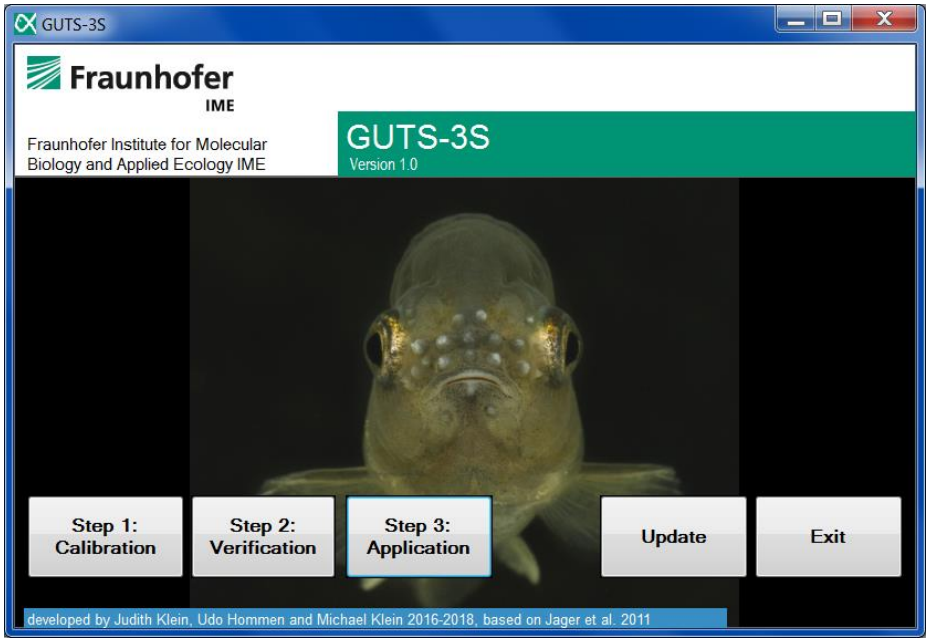


Figure 2: Start up window of the program GUTS-3S

3.1. Calibration

The aim of the calibration is to find the substance and species specific TK-TD parameters by minimizing the deviations between experimental data sets and model outputs.

The experimental data have to be provided as Excel files. It is possible to use more than one data set

as input for the calibration

Before the calibration can start, different settings have to be made, which are described in the following screenshots. By clicking at the calibration button, the following calibration form appears (Figure 3).

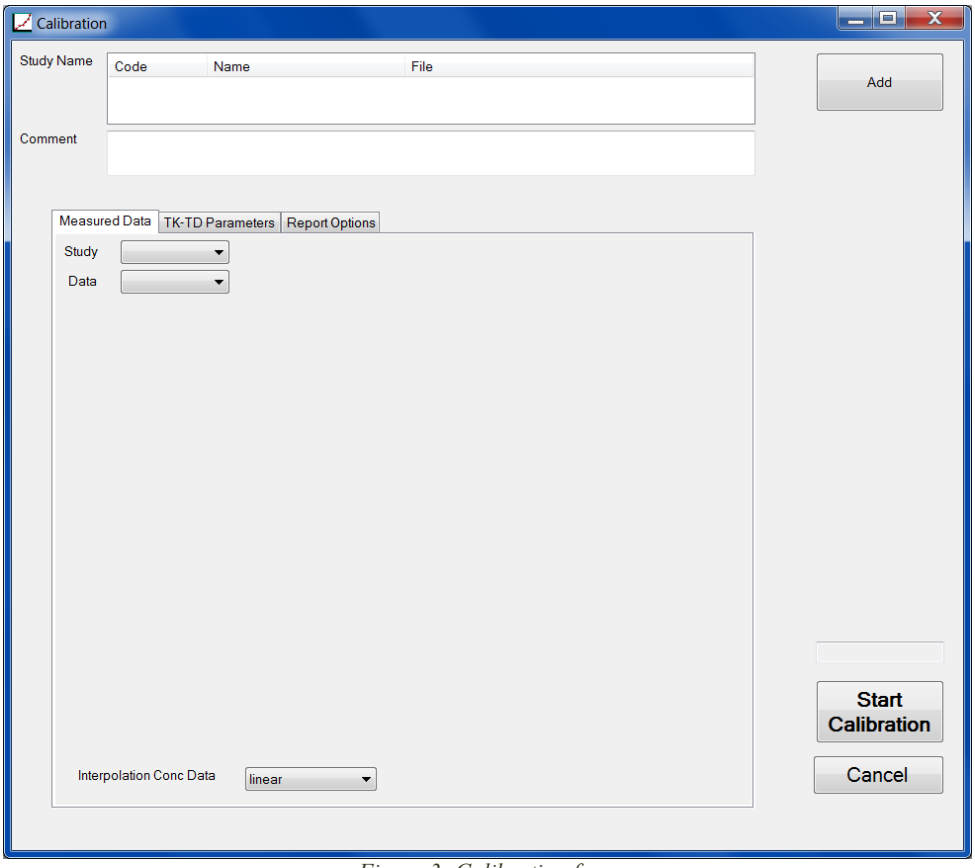


Figure 3: Calibration form

At first, we have to add experimental data. We can do so by clicking the add button and adding an Excel file including the data (Figure 3).

We add as example data sheet the data used in Ducrot et al. 2016. It is possible to add further studies by clicking at the add button. In this case, we only add one study by clicking at “add” and by choosing one study excel file.

The excel file is selected by browsing files on your computer. In Figure 35, an example of such an experimental data file is presented. The aim of the calibration is to find values for the TK-TD

parameters.

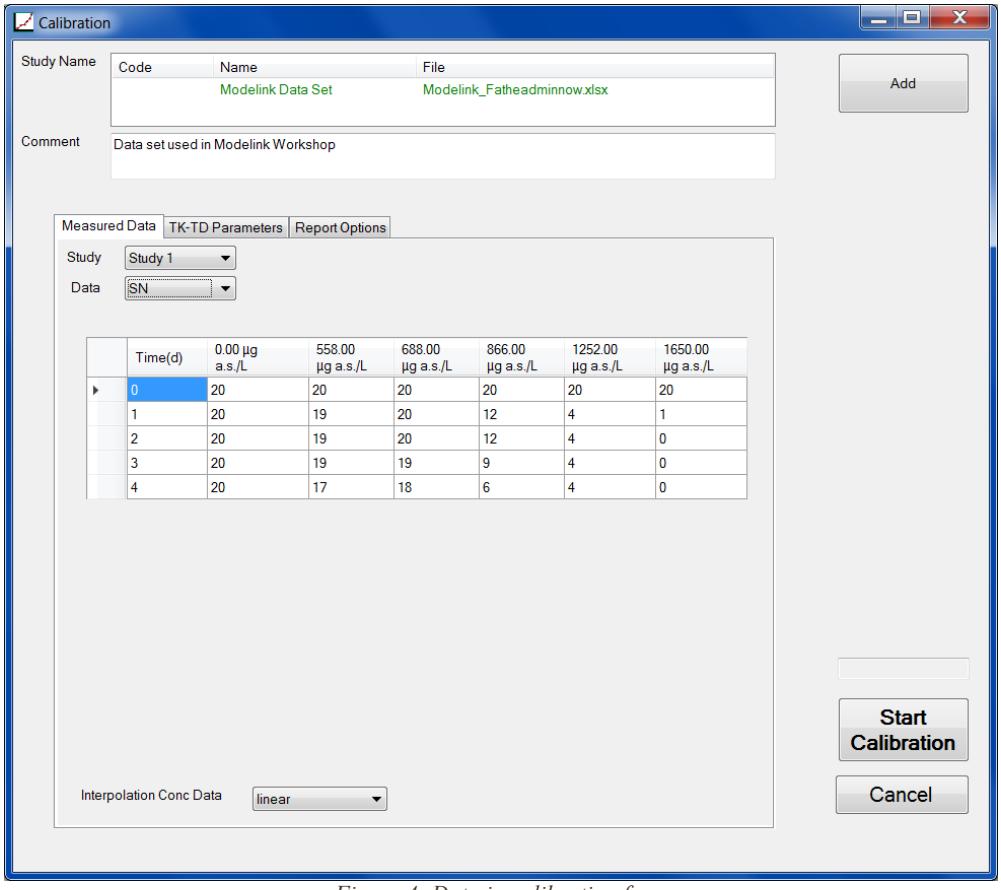


Figure 4: Data in calibration form

The program enters the information of the study file (environmental data and measured experimental data) in the program's surface. In the tab "measured data", one can see the external concentration, possibly measured internal concentrations and number of survivors in time (Figure 4).

The experimental data sets usually contain measurements on survived individuals in time at different exposure levels. In the example of fish, for instance, experimental data according to the OECD guideline 203 about the acute toxicity test can be used (OECD 1992).

The next tab "TK-TD Parameters", Figure 5, is about the substance and species specific TK-TD parameters and the calibration settings, e.g. which parameters should be calibrated, what are initial values and ranges for the parameters and what should be the objective function (e.g. the measure to be optimized during calibration). Furthermore, the user has to decide whether to use the SD or IT model.

In our example case (Ducrot et al. 2016), there is only the survival data and no additional information on the toxicokinetic processes given, such that we have only one dominant rate constant determining the speed of uptake and elimination of the active substance. Thus, we have in total to find values for three parameters with respect to each model approach.

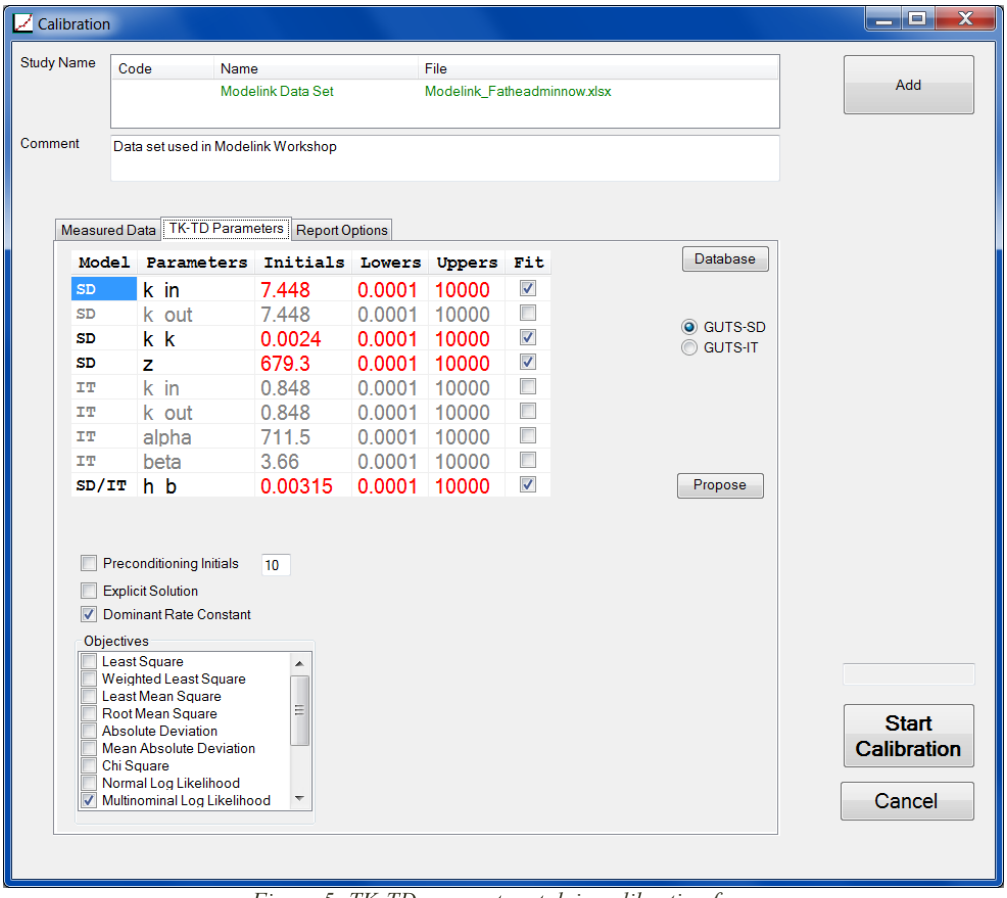


Figure 5: TK-TD parameters tab in calibration form

If you are not sure, which values are realistic, you can click the "Propose" button and the program makes suggestions for the TK-TD values (Figure 6).

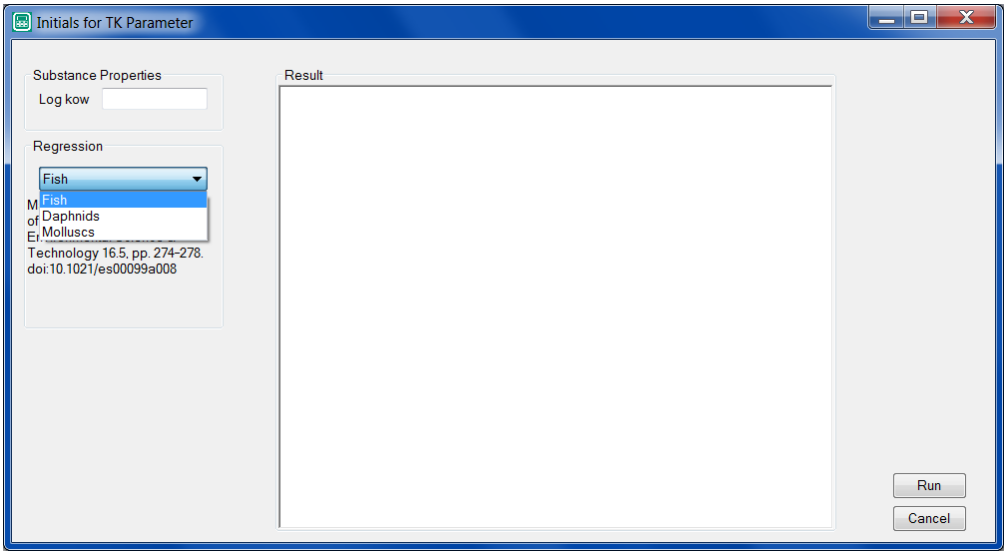


Figure 6: Initial values for TK parameters

The program suggests values for the TK parameters based on the bioconcentration factor (BCF). A bioconcentration factor expresses the tendency of a compound to accumulate in the organism (EFSA PPR Panel 2013). We have found the following regression equations in literature, where the calculations with regard to kin are based on Hawker and Des Connell 1986.

$$\log(BCF) = a \cdot \log(K_{ow}) + b$$

$$\log(\frac{1}{k_{out}}) = x \cdot \log(K_{ow}) + y$$

$$\log(k_{in}) = (a - x) \cdot \log(K_{ow}) + (b - y)$$

The first two regression curves are based on experimental data, where the third equation for k_{in} is obtained by dividing the BCF regression by the $1/k_{out}$ regression.

Fish:

$$BCF = 0.048 \cdot K_{ow} \text{ or} \qquad \text{(Mackay 1982)}$$

$$\log(BCF) = \log(K_{ow}) - 1.32 \qquad \text{(Mackay 1982)}$$

$$\log(\frac{1}{k_{out}}) = 0.663 \cdot \log(K_{ow}) - 0.947 \qquad \text{(Mackay 1982)}$$

$$\log(k_{in}) = 0.337 \cdot \log(K_{ow}) - 0.373$$

Zebra fish (Petersen and Kristensen 1998):

$$\log(k_{out}) = -0.41 \cdot \log(K_{ow}) + 1.47$$
$$\log(k_{in}) = 1.98 \cdot \log(K_{ow}) + 0.147$$

Daphnids:

$$\log(BCF) = 0.989 \cdot \log(K_{ow}) - 1.3155 \qquad \text{(Hawker and Des Connell 1986)}$$

$$\log(\frac{1}{k_{out}}) = 0.507 \cdot \log(K_{ow}) - 2.053 \qquad \text{(Hawker and Des Connell 1986)}$$

$$\log(k_{in}) = 0.391 \cdot \log(K_{ow}) + 0.7375$$

Molluscs

$$\log(BCF) = 0.844 \cdot \log(K_{ow}) - 1.235 \qquad \text{(Hawker and Des Connell 1986)}$$

$$\log(\frac{1}{k_{out}}) = 0.54 \cdot \log(K_{ow}) - 0.983 \qquad \text{(Hawker and Des Connell 1986)}$$

$$\log(k_{in}) = 0.304 \cdot \log(K_{ow}) - 0.252$$

With regard to this rather statistic regression curves, they may be right in many cases. Though the regression equations are not necessary correct for each species substance combination, however, they can serve as initial values. There exist also more complex regression curves based on other substance specific structure related properties (QSAR) e.g. Kaiser 2012.

As the dominant rate constant k_d is just a renaming of k_{out}, a possible initial value for k_d may be then the by regression found value regarding the elimination rate k_{out}.

For the toxicodynamic parameters we relate the parameters directly with the external concentration.

For the threshold parameter z, we use the same parameter value as for α. We assume a constant exposure during the test period $C_{ext}(t) = K$. For the killing rate k_k, we fit the experimental data to

$$S(t) = \exp(-h_b \cdot t) \qquad K \le z$$
$$S(t) = \exp(k_k \cdot z \cdot t - k_k \cdot \mathrm{K} \cdot t \ - h_b \cdot t) \qquad K > z$$

Fitting the log logistic curve to the external concentration $C_{ext}(t) = K$ as well as survival data derived at the last time point t_n.

$$S(t) = 1 - (\frac{1}{1 + \alpha^{\beta} \cdot K^{-\beta}})$$

With respect to the background mortality h_b we propose to fit the equation $S(t) = \exp(-h_b \cdot t)$ (control survival) to the experimental control survival probability data to get an adequate initial value.

In addition to that you can save TK-TD parameters and load parameter values derived in other

calibration runs by clicking the "Database" button. A similar window opens as concerning the growth parameters containing a database of all saved TK-TD parameter sets.

The preconditioning check box is to test several initial values to have greater possibility to find a good optimization solution. Instead of one initial parameter value set, the selected number of sets are tested with the parameters chosen randomly out of their ranges.

In case of constant exposure, it is possible to use the explicit model solution instead of discretizing the model system.

In the objectives box, several measures for the deviation between model results and data are available. Often in literature a multinomial log likelihood function is used (Ashauer 2010; Ducrot et al. 2016; Jager et al. 2011). Below a list of all implemented objective functions is presented (Table 1: Implemented objective functions for calibration).

Let $N \in \mathbb{N}$ be the total number of data points, $O \in \mathbb{R}_+^N$ be the observations and $C \in \mathbb{R}_+^N$ be the calculations. Furthermore, let $\bar{O}, \bar{C} \in \mathbb{R}_+$ be the means of the respective data.

Table 1: Implemented objective functions for calibration

Objective	Formula
Least Square	$\|O-C\|_2^2 = \sum_{i=1}^{N} (O_i - C_i)^2$
Weighted Least Square	Let $w_i \in \mathbb{R}_+$, $\sum_{i=1}^{N} w_i \cdot (O_i - C_i)^2$
Least Mean Square	$\sum_{i=1}^{N} \frac{1}{N} (O_i - C_i)^2$
Root Mean Square	$\sqrt{\sum_{i=1}^{N} \frac{1}{N} (O_i - C_i)^2}$
Absolute Deviation	$\|O-C\| = \sum_{i=1}^{N} \|O_i - C_i\|$
Mean Absolute Deviation	$\|O-C\| = \sum_{i=1}^{N} \frac{1}{N} \|O_i - C_i\|$

Chi Square	$\chi^2 = \sum_{i=1}^{N} \frac{(C_i - O_i)^2}{O_i}$
Normal Log Likelihood	$\ln(L) = \frac{N}{2} \cdot \ln\left(\sum_{i=1}^{N} (O_i - C_i)^2 \right)$
Multinomial Log Likelihood	

In the last tab “Report Options“, the design of the output can be specified (Figure 7).

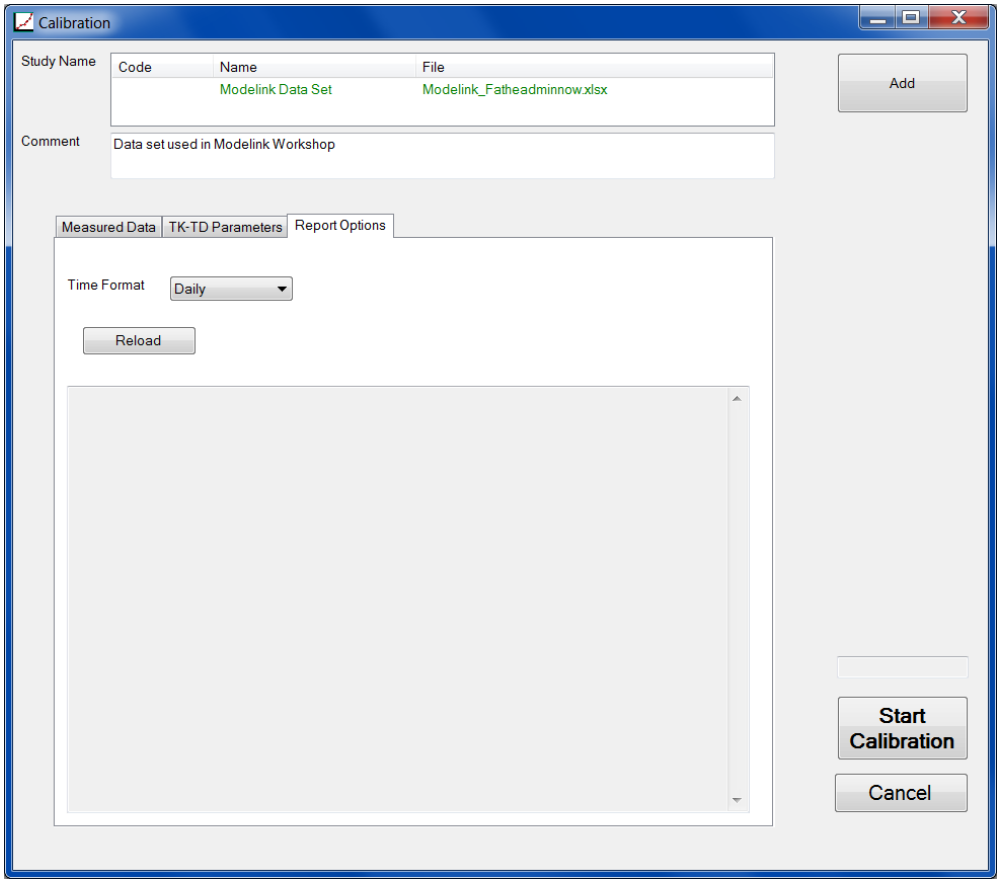


Figure 7: Report option of calibration

Clicking at Start Calibration starts the calibration procedure. The progress of the calibration is shown by the progress bar at the bottom. It is possible to stop the calibration by pressing the "Stop" button (Figure 8). Clicking at "Cancel" leads back to the initial start screen of the program and the calibration is cancelled.

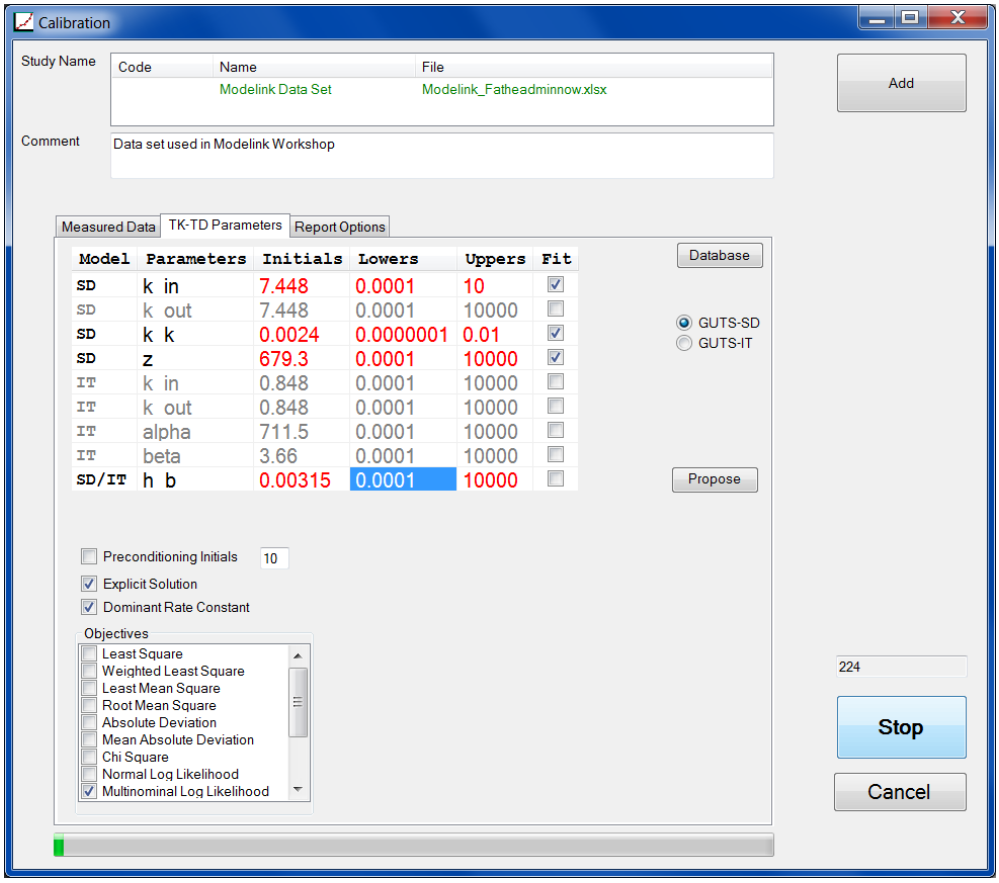


Figure 8: Calibration started

Above the "Stop" button the number of iterations of the algorithm finding the TK-TD parameters is given. Here, the screenshot was taken at iteration 224.

If the solver finds an optimal solution, a new window opens to show the results, in the form of a plot of the observed and modelled survival probability over time (Figure 9).

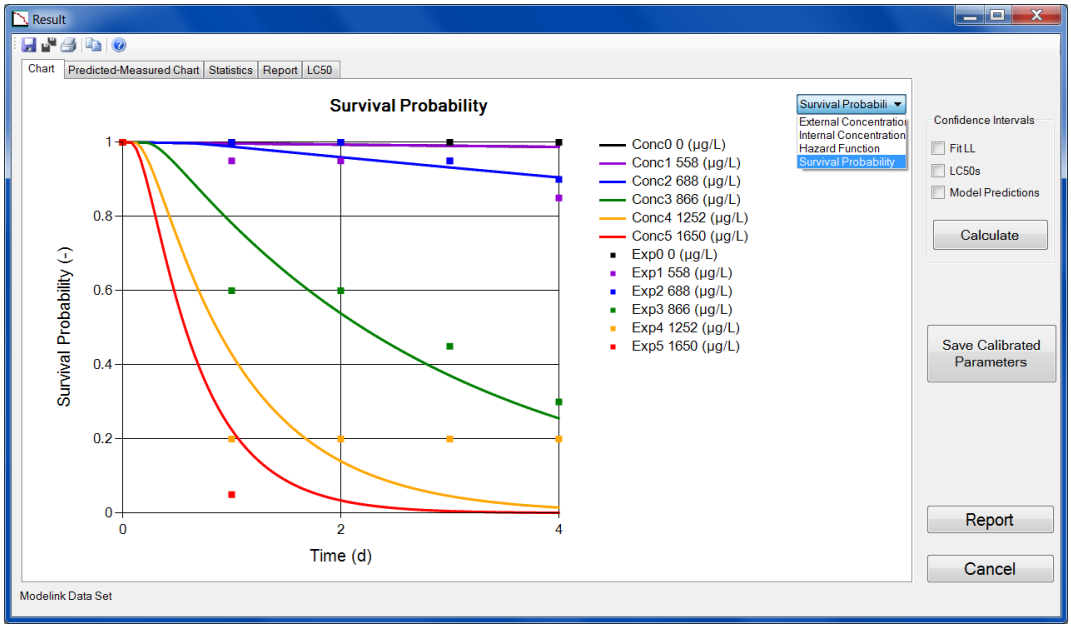


Figure 9: Calibration chart of calibration result form

Additionally the user can see the external concentration, internal concentration and hazard function (in case of SD) or distribution function (in case of IT).

If the user is content with the calibration result, it is possible to calculate confidence intervals. In the group box on the right hand side different settings and options can be made.

The calculation of confidence intervals of parameters is done via profiling likelihood. In the program, two different possibilities for their calculation is given: with fitting the likelihood and without fitting the likelihood (Fit LL). The output of the calculation is a text report (txt file).
In general, the calculation is very time consuming. Thus, by clicking at "Calculate" a form appears showing the process of the calculation (Figure 10).

For more information, see Jager (2016); Meeker & Escobar (1995); Moerbeek et al. (2004).
Calculating the confidence interval of the optimal parameter p^* is finding the intersection of the approximate chi square distribution function $2 \cdot (L_{best} - L(p))$ and a critical chi square value $\chi^2_{df,1-\alpha}$ in an interval $[\underline{p}, p^*]$ respective$[p^*, \overline{p}]$:

$$2 \cdot (L_{best} - L(p) - \chi^2_{df,1-\alpha}) \approx 0$$

The approximate chi square distribution function is calculated based on the optimal likelihood value L_{best} and the current likelihood value $L(p) p \in [\underline{p}, p^*]$ respectively $p \in [p^*, \overline{p}]$.

The critical chi square value is based on the number of fix parameters (1). Using an alpha of 0.05 leads to a critical value of 3.841.

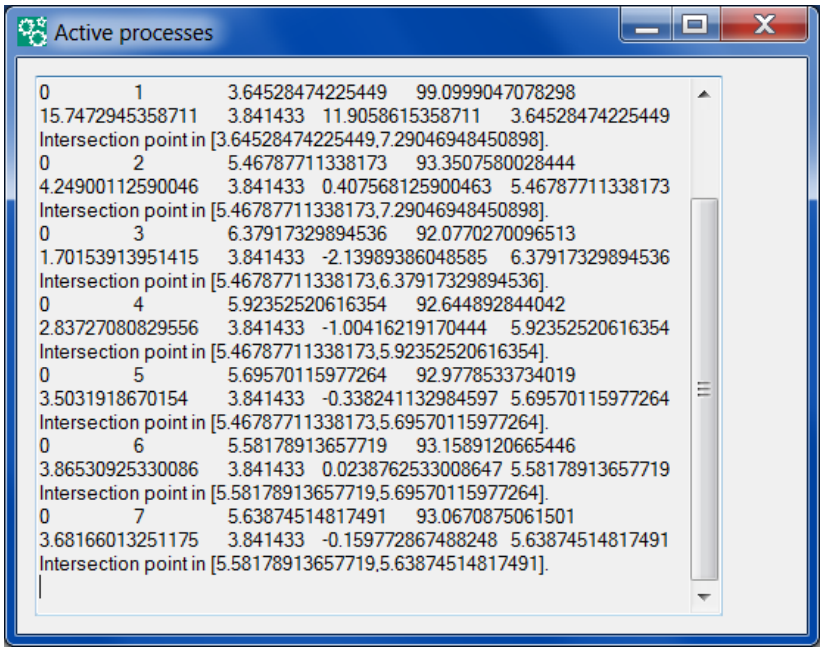


Figure 10: Active processes form appearing by clicking at "Calculate Confidence Intervals"

A report und the result of the numeric estimation of confidence intervals is stored in a text file.

For this example, we obtain the following confidence intervals:

Table 2: Estimated confidence intervals

Parameter	Value	Lower Bound	Upper Bound
k_in	7.290469	5.588909	10
k_k	0.001931	0.001416	0.002517
z	674.442985	646.080488	686.929521

We obtain the following results by checking LC50s and model predictions: During the numeric estimation of confidence intervals all internal concentration values, hazard function respective distribution function values, and survival probability values are stored. Calculating of these values

the 95% and 5% profiles leads to prediction band (Figure 11, Figure 12 and Figure 13).

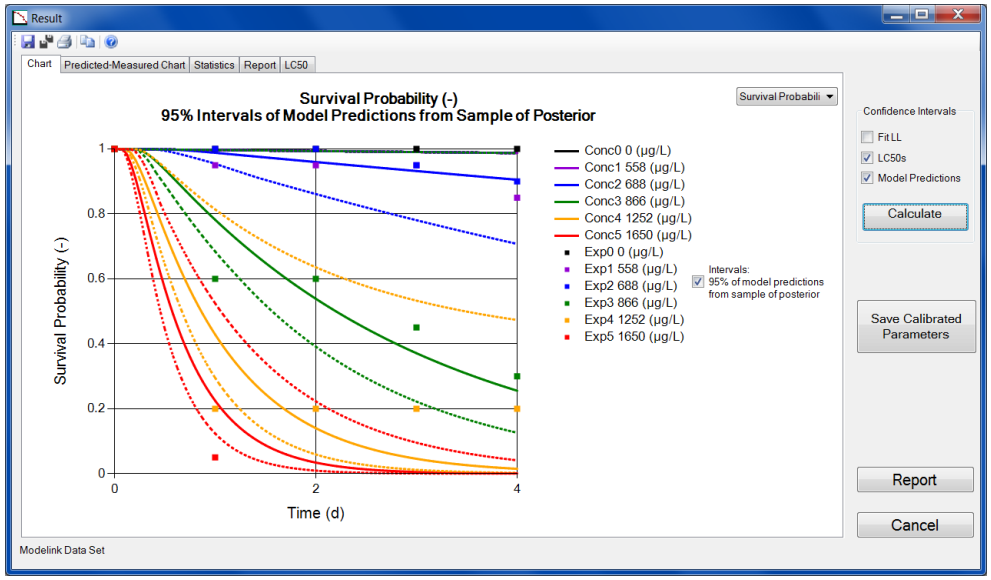


Figure 11: 95% of model survival predictions from sample of posterior

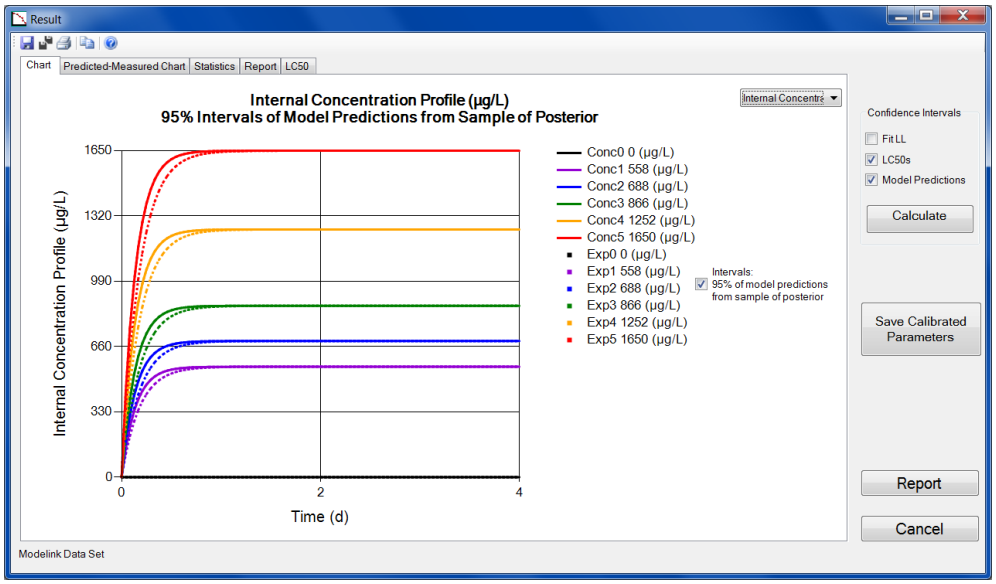


Figure 12: 95% of predicted internal concentration from sample of posterior

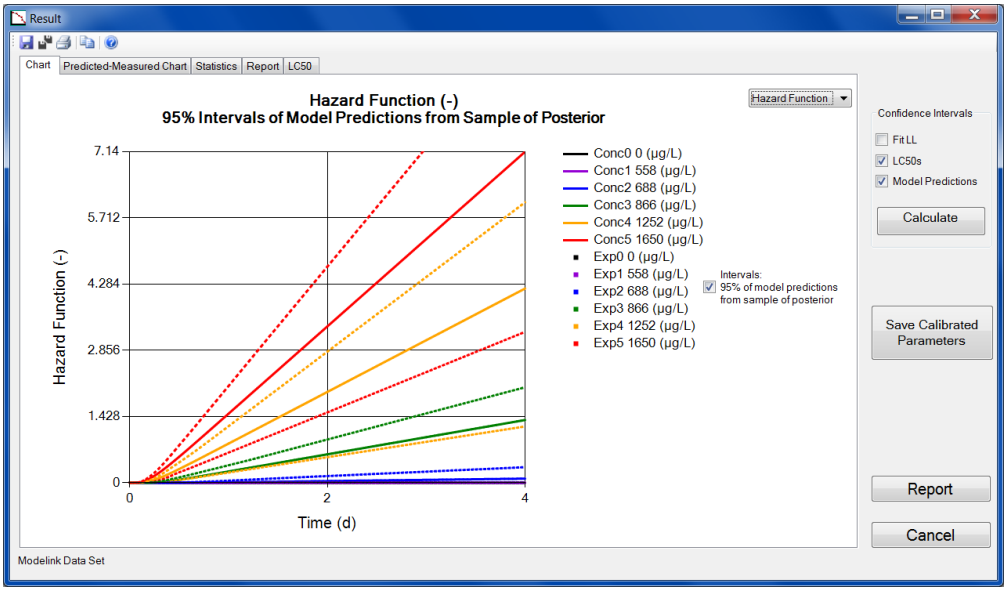


Figure 13: 95% of predicted hazard from sample of posterior

In Figure 14, the predicted LC50 in time with the respecting 95 % intervals is shown. The calculation is based on the samples stored during the numeric estimation of confidence intervals. For each stored survival probability data series the LC50 is calculated in time. Afterwards, the 5% and 95% quantiles are calculated (95% interval).

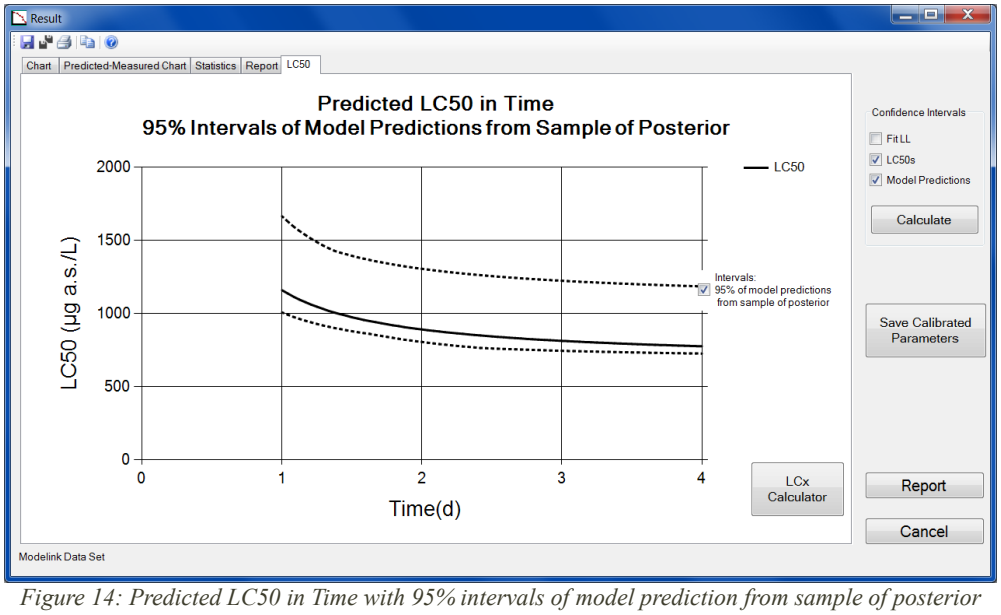


Figure 14: Predicted LC50 in Time with 95% intervals of model prediction from sample of posterior

The second tab "Predicted-Measured Chart", Figure 15, contains a figure of prediction (x-axis) and measured data (y-axis) including initial values and control data.

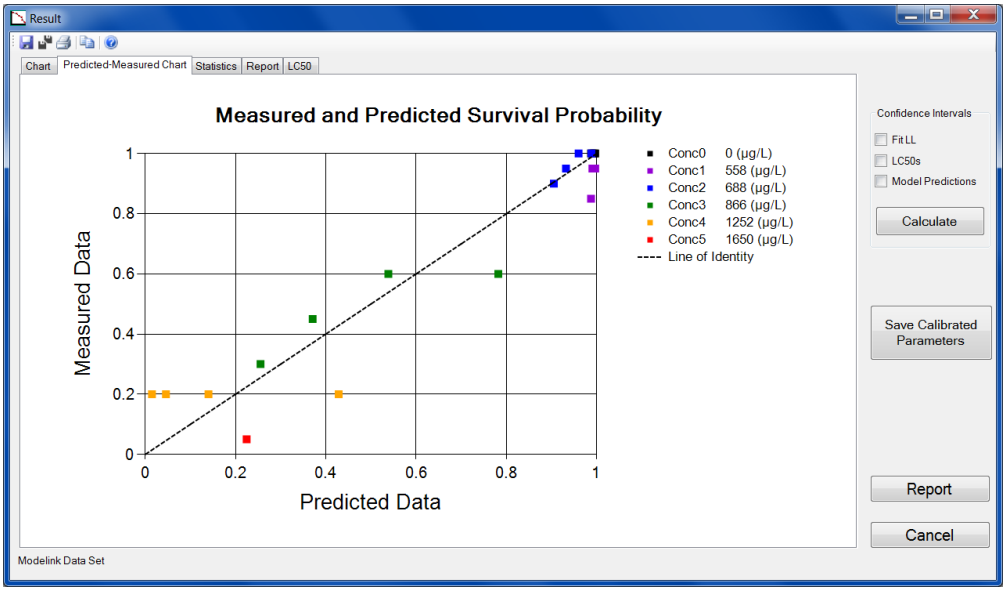


Figure 15: Predicted-Measured Chart of the calibration result form

The third tab, Figure 16, presents the statistical results based on number of survivor or survival probability to provide additional information on the goodness of fit. For these statistical calculations, the initial values (t=0) are not considered. Furthermore, we do not consider the control data for the total study statistics as these are not affected by the TK-TD parameters.

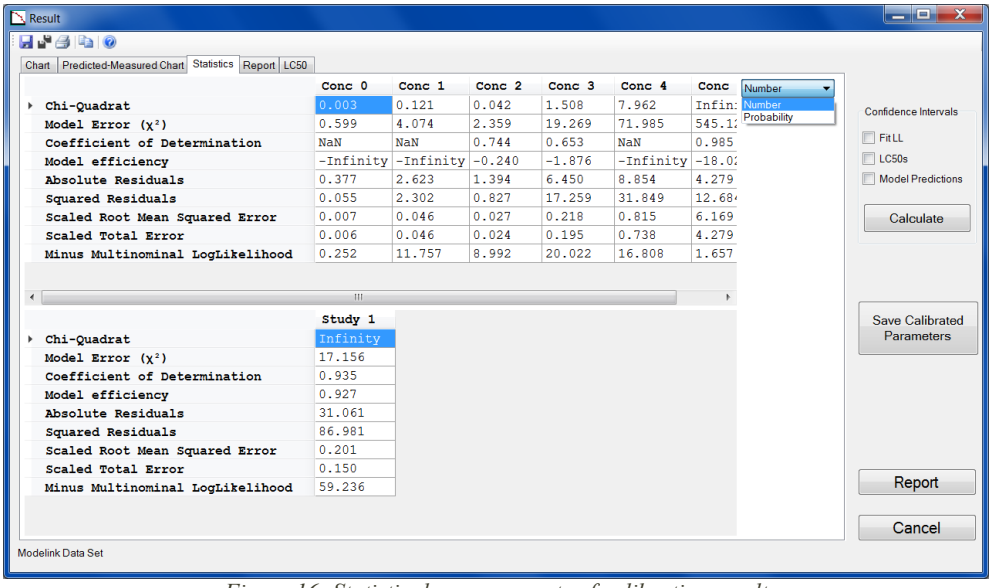


Figure 16: Statistical measurements of calibration result

We perform the statistics for both endpoints, number of survivors and survival probability. Due to the properties of the statistical measurements only the residuals are different, the other statistical measurements have the same values.

However, an overview of the used statistics is given in Table 3. Again, let $N \in \mathbb{N}$ be the total number of data, $O \in \mathbb{R}_+^N$ the experimental observation and $C \in \mathbb{R}_+^N$ the corresponding calculated model prediction. The means are represented by $\bar{O} \in \mathbb{R}_+$ respectively $\bar{C} \in \mathbb{R}_+$.

Table 3: Statistical measurements to compare the correspondence of model prediction and experimental data

Statistics	Formula
Chi-Quadrat	$\chi^2 = \sum_{i=1}^{N} \frac{(C_i - O_i)^2}{O_i}$
Model error	The number $m \in \mathbb{N}$ denotes the degrees of freedom (number of measurements minus number of model parameters) and let be $\alpha \in (0,1)$.

	Let χ^2_{tab} be the tabulated $\chi^2_{m,\alpha}$. Usually a value of $\alpha = 0.05$ is chosen. $\epsilon = 100 \cdot \frac{1}{\overline{O}} \cdot \sqrt{\frac{1}{\chi^2_{tab}} \cdot \sum_{i=1}^{N} (C_i - O_i)^2}$
Coefficient of Determination	$r^2 = \left(\frac{\sum_{i=1}^{N} (O_i - \overline{O})(C_i - \overline{C})}{\sqrt{\sum_{i=1}^{N} (O_i - \overline{O})^2 \cdot \sum_{i=1}^{N} (C_i - \overline{C})^2}} \right)^2$
Model efficiency	$EF = 1 - \frac{\sum_{i=1}^{N} (C_i - O_i)^2}{\sum_{i=1}^{N} (O_i - \overline{O})^2}$
Absolute Residuals	$AR = \sum_{i=1}^{N} \lvert C_i - O_i \rvert$
Squared Residuals	$SR = \sum_{i=1}^{N} (C_i - O_i)^2$
Scaled Root Mean Squared Error	$SRMSE = \frac{1}{\overline{O}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (C_i - O_i)^2}$
Scaled Total Error	$STE = \frac{\sum_{i=1}^{N} \lvert C_i - O_i \rvert}{\sum_{i=1}^{N} O_i}$

Regarding the statistical result in Figure 16, one can see that for example for the control not all statistical measurements deliver a numeric result.

We consider this example in more detail. The used data for the statistics is presented in Table 4.

Table 4: Experimental data of control

time (d)	O	C
1	20	19.937
2	20	19.874
3	20	19.812
4	20	19.75

The coefficient of determination is not a number (NAN) because all observation data has the same values. Thus, we are dividing 0 by 0. The same yields for model efficiency.

The fourth tab contains a written report containing all necessary information on the input, settings of the calibration and the model prediction (Figure 17).

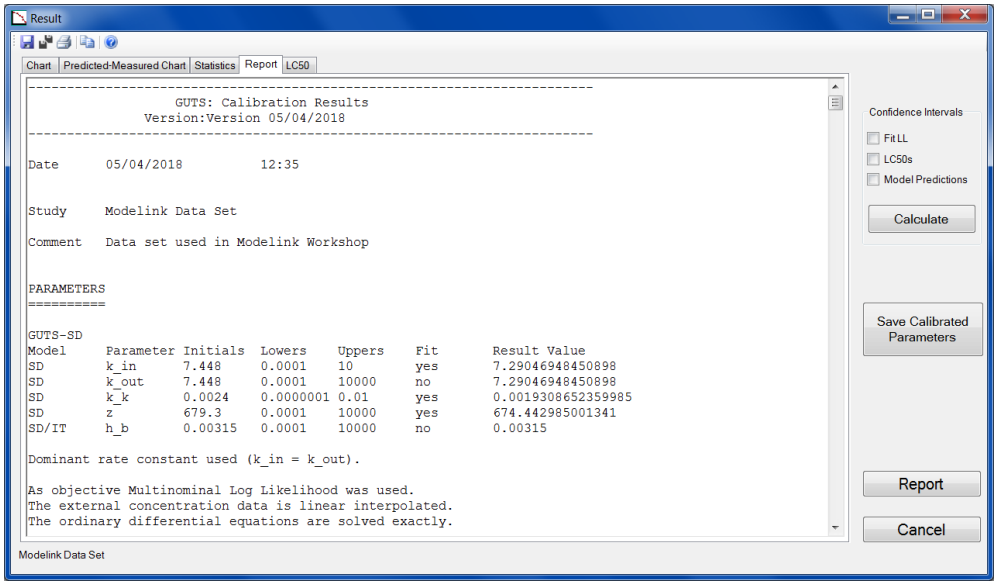


Figure 17: Text report of calibration

It is possible to copy the results, separately or all simultaneously, or to save and print them using the symbols at the top. The file format is dependent on the result file: the charts are saved as jpeg, the statistics and text report in txt.

LC Calculator

Clicking at “LCx Calculator” opens the form in Figure 18 to calculate the concentration leading to x % effect.

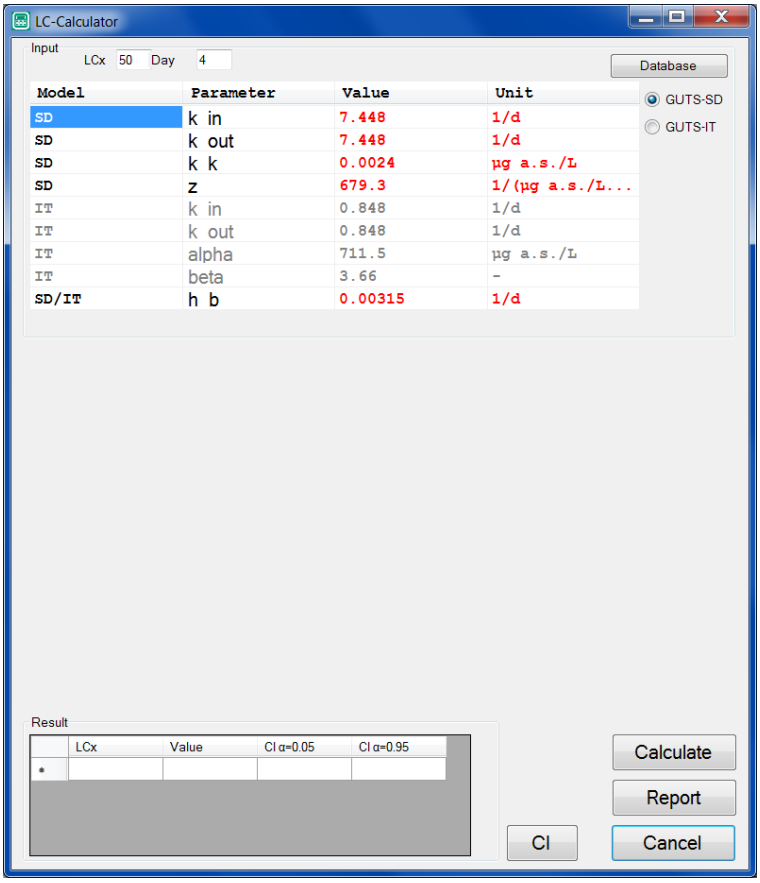


Figure 18: Calculation of LCx

Clicking at the button “Calculate” starts the LCx calculation procedure. As input the program needs the current or different TK-TD parameters, the fraction effect $x \in [0,100]$ and the time span (here, Day =4).

As initial guess for the concentration leading to x % effect, we use 100 times the highest concentration, thus in our example case 165000.

For the calculation of LCx, we need the current predicted survival probability of treatment and control, as well as the fraction effect (e.g. x=50). We use the same function as Tjalling Jager in the BYOM matlab function calc_lc50.m.

$$crit = \frac{S(t)}{S_0(t)} - (1 - \frac{x}{100})$$

For each time point until "Day" (content of Textbox entered by user), we calculate the concentration via intermediate theorem leading to a critical value of zero. The found concentration is the LCx value. This result value for the chosen "Day" is entered into the table and the change of LCx values in time is presented in the chart until the chosen "Day" (Figure 19). By clicking with the right mouse click at the chart, the user can copy to clipboard or save or print the chart.

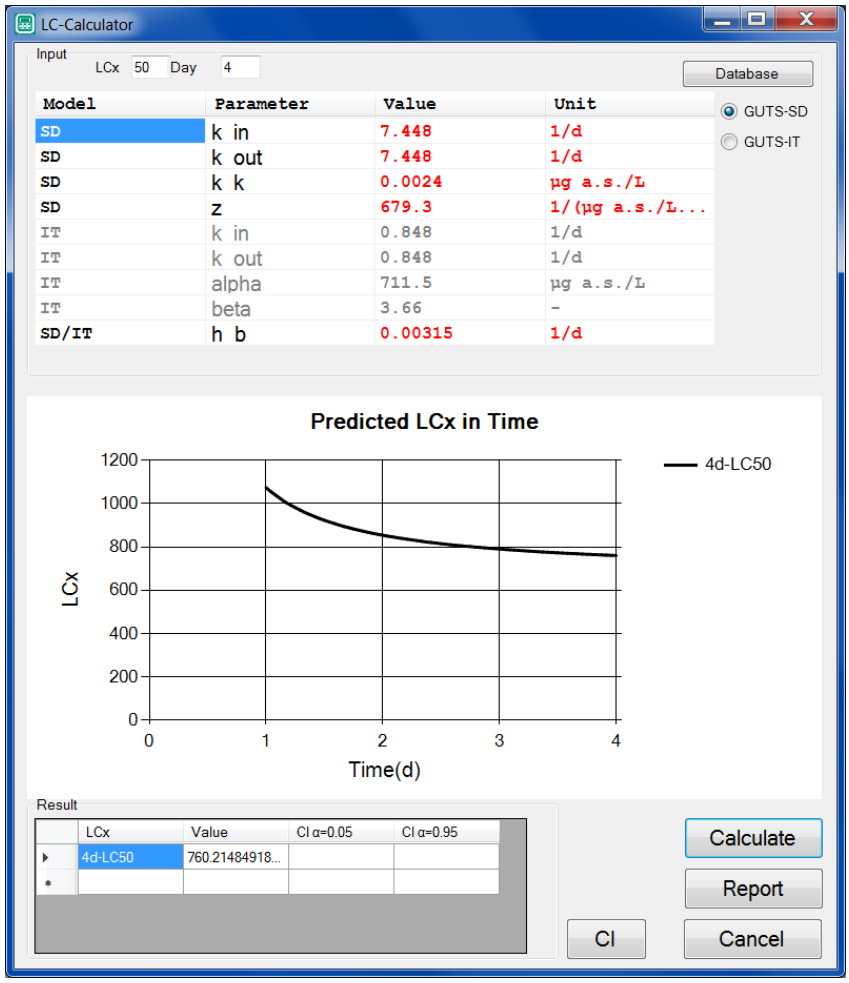


Figure 19: LCx Calculator Form after a first run of the calculation of the 4d-LC50

Clicking at report saves and opens the LCx report. The report contains the values of TK-TD parameters, and the concentrations leading to x % effect in time t until entered "Day". In Figure 20, an LCx report for the 4d-LC50 is given using above chosen experimental data set and parameters of Ducrot et al. 2016.

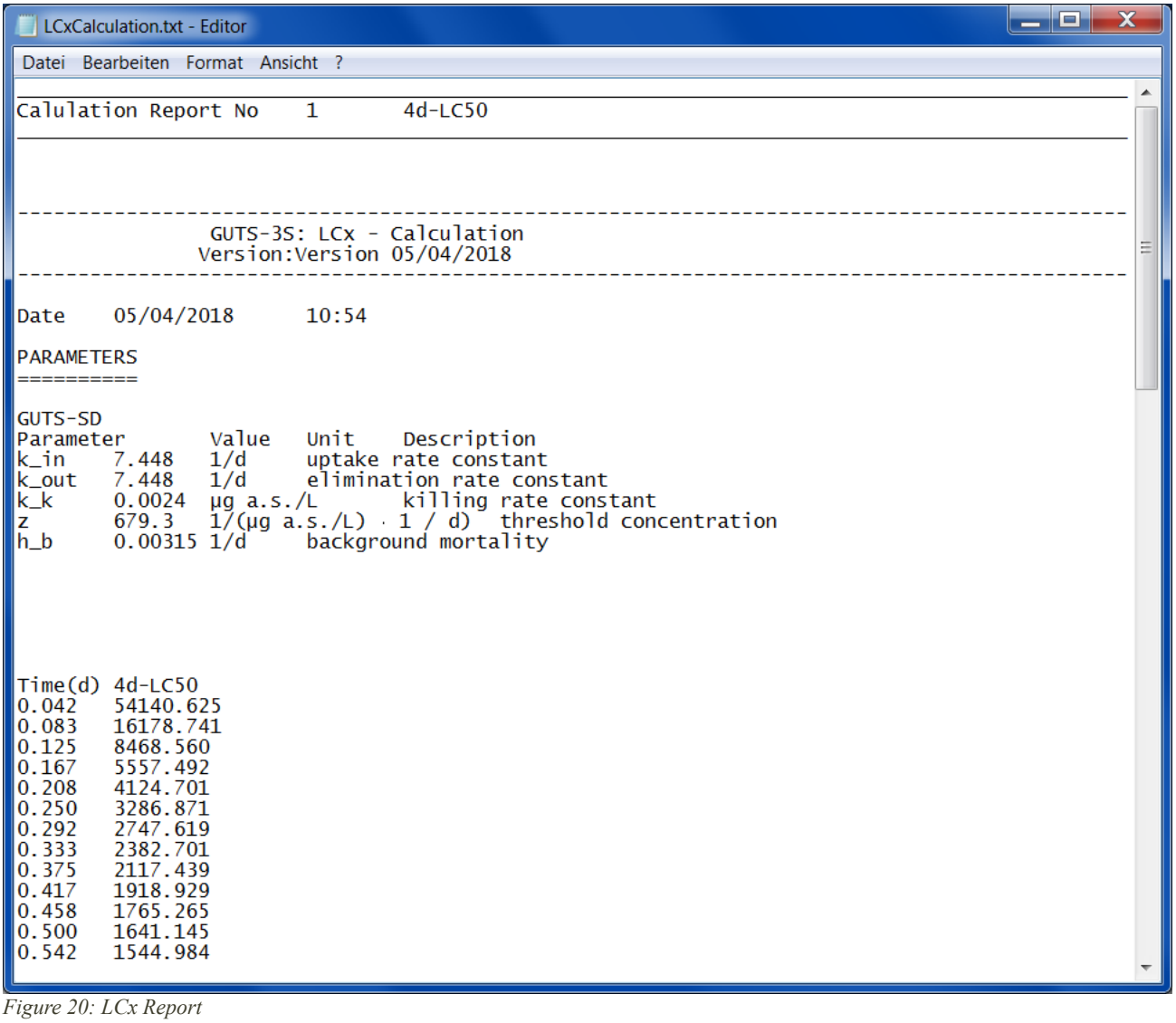


Figure 20: LCx Report

Click at "Cancel" to return to the calibration result form.

Clicking at "CI" calculates the confidence interval for α equal to 0.05 and 0.95. The calculation of confidence intervals of LC50 is done using Monte Carlo Simulation. For this, the user selects an

upper and lower bound for each parameter and a method to generate random samples. For each random generated parameter set, the LC50 is calculated. At last, the program calculates the minimum, maximum, expected value, variance, standard deviation, median, 5th percentile and 95th percentile based on all calculated scenarios.
Possible bounds are, for instance, the confidence intervals of parameters.
In a report, all parameter sets as well as the corresponding LC50 value is given.

Let $\lfloor x \rfloor$ the smallest integer being smaller or equal to x. Assuming we have ordered data of sample size n: $(x_1, x_2, \cdots, x_n)$ with $x_1 \leq x_2 \leq \cdots \leq x_n$. For any number $p \in (0,1)$ we can calculate the p-quantile by

$$x_p = \begin{cases} \frac{1}{2} \cdot (x_{n \cdot p} + x_{n \cdot p+1}) & \text{if } n \cdot p \text{ is integer} \\ x_{\lfloor n \cdot p+1 \rfloor} & \text{else} \end{cases}.$$

3.2. Verification

The aim of this procedure is to verify the TK-TD model by comparing model prediction to experimental studies (not used for calibration).

Input: experimental data, parameter values
Settings: number of fitted parameters (for statistics)

The verification form has the same structure as the calibration form; the only part that is different is the TK-TD tab. We can manually enter the parameter we want to validate or choose a parameter set from the TK-TD database. By clicking at “Start Verification”, the programs calculates the model prediction using the entered TK-TD parameters and compares them to the given experimental data set (Figure 21).

The statistics is similar to the calibration statistics in Table 3.

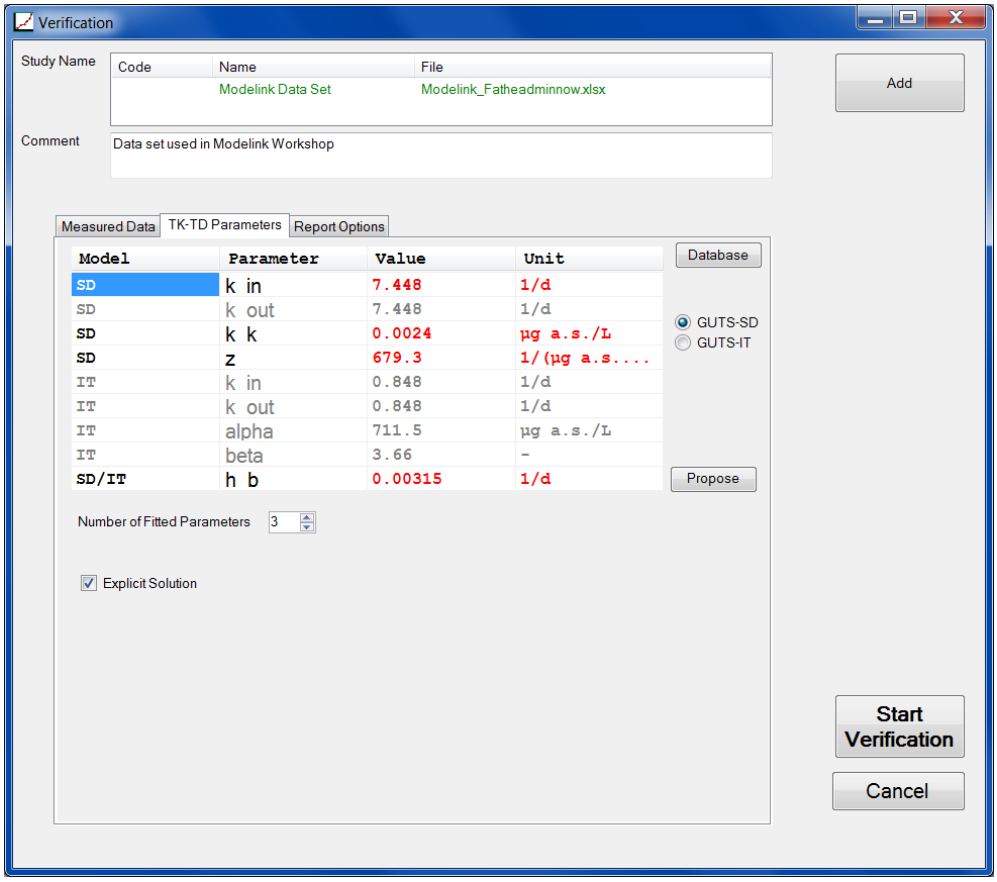


Figure 21: TK-TD parameters tab in verification form

As result, we obtain the time series of exposure, internal concentration, hazard (GUTS-SD) or distribution function (GUTS-IT) and survival probability. It can also be selected which study (if more than one was selected before) should be shown, if all or only specific concentrations should be plotted and if the experimental data should be shown as symbols (Figure 22).

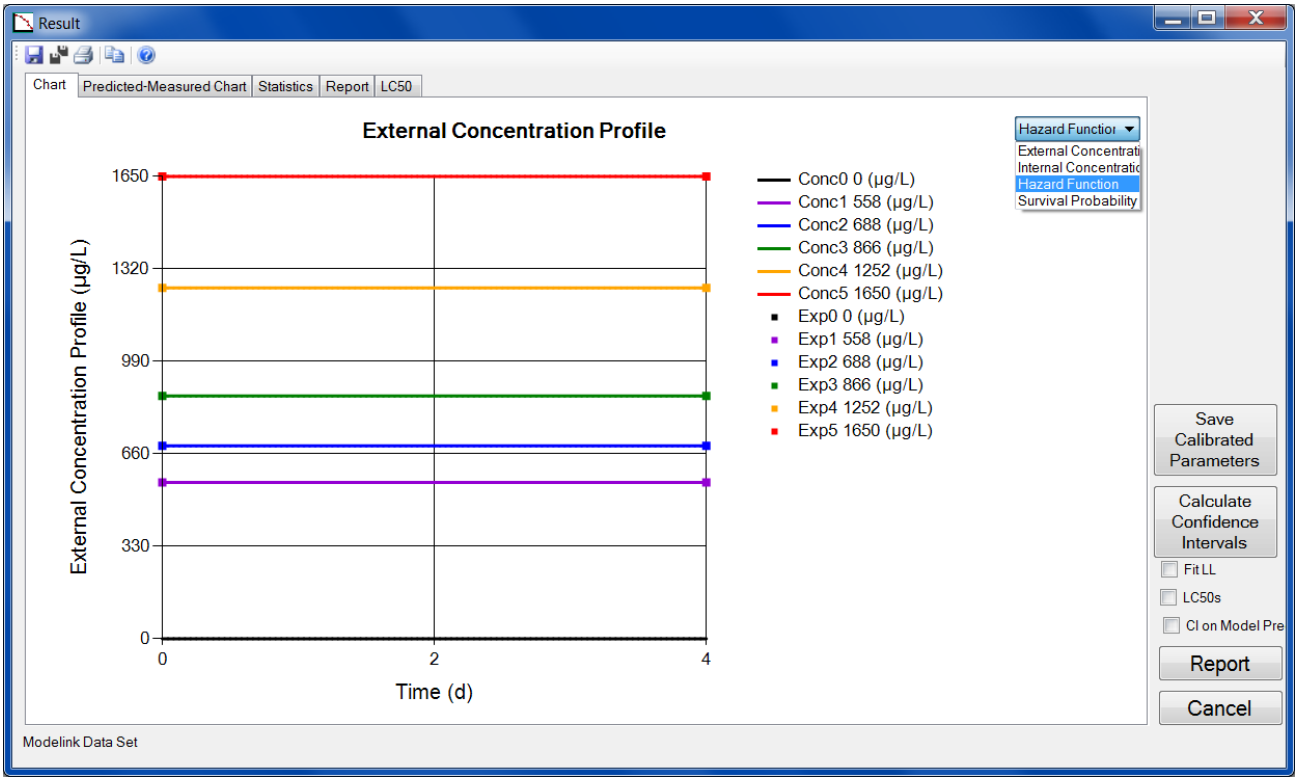


Figure 22: Exposure profile of study in verification result form

In Figure 23, the verification result with respect to the number of fronds can be seen. Experimental data (points) as well as model predictions (line) is presented in arithmetic scale.

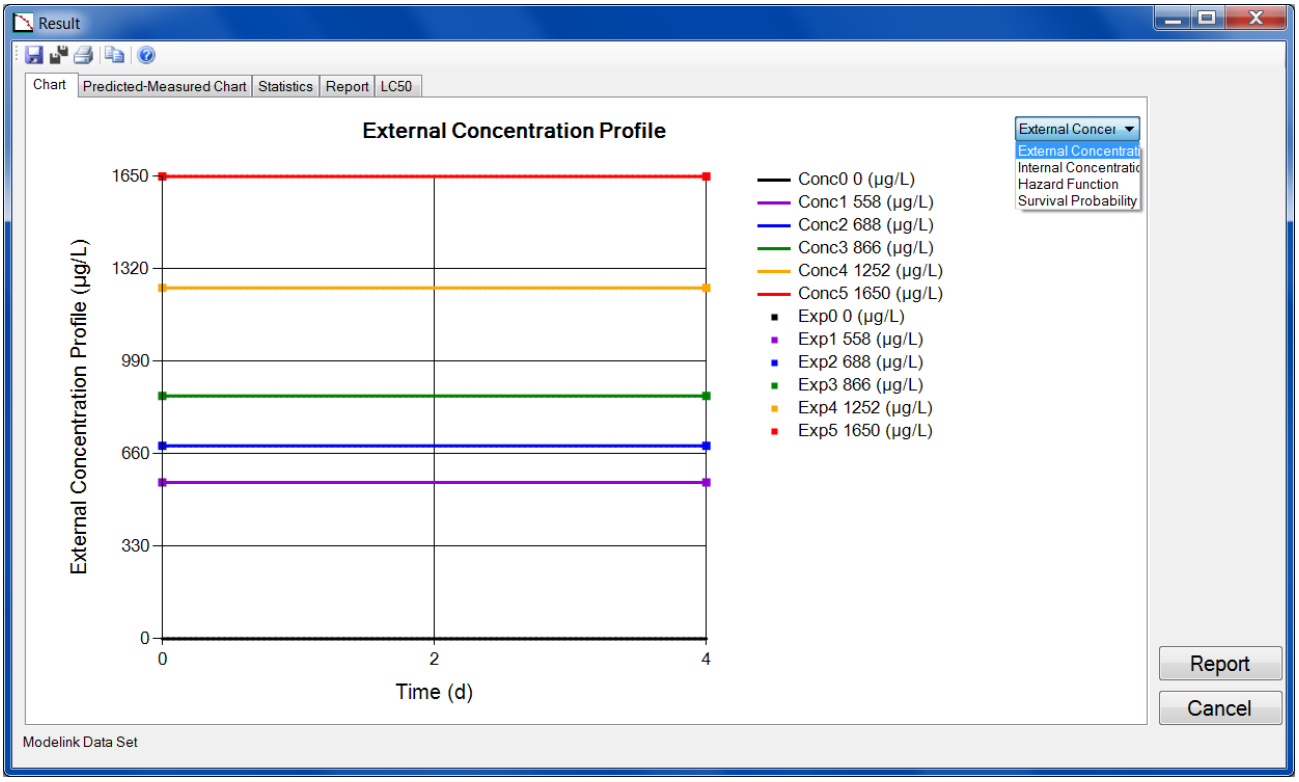


Figure 23: Survival probability of study in verification result form

It is possible to copy the results separately or all simultaneously. Furthermore, one can save or print the results. Again, the file format is jpeg for the plots, and txt for statistics and the text report.

The results contain a similar statistical assessment and report as the calibration statistics.

3.3. Application

Having found suitable TK-TD parameters in calibration and validation, in "Application" the user can simulate the survival probability of a species in dependence of different exposure scenarios. Before starting the calculation procedures, the user has to choose several input parameters such as

- the exposure of substance (tab "measured data", Figure 24),
- the TK-TD parameters and
- the output variables (tab "options").

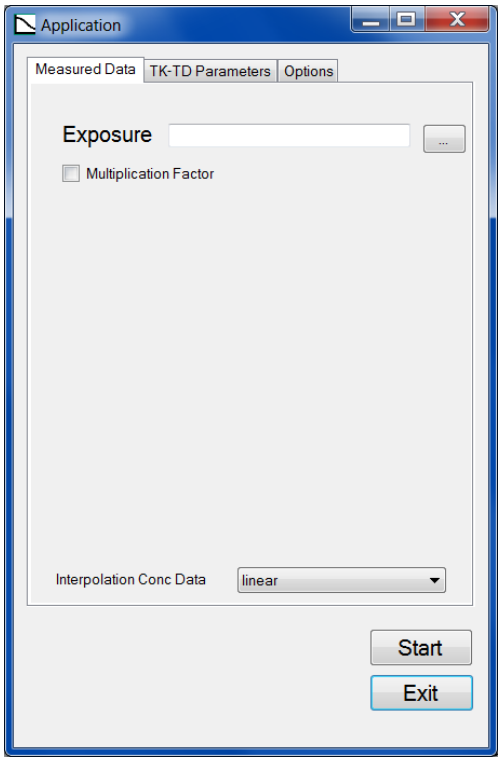


Figure 24: Initial application form: Measured Data Tab

Exposure has to be specified. It can be defined to be constant or they can be read in from txt files (Figure 25).

By clicking at multiplication factor, the program multiplies the exposure pattern by a factor to be entered by the user. Furthermore, the user can decide whether the external concentration shall be interpolated exponentially or linearly.

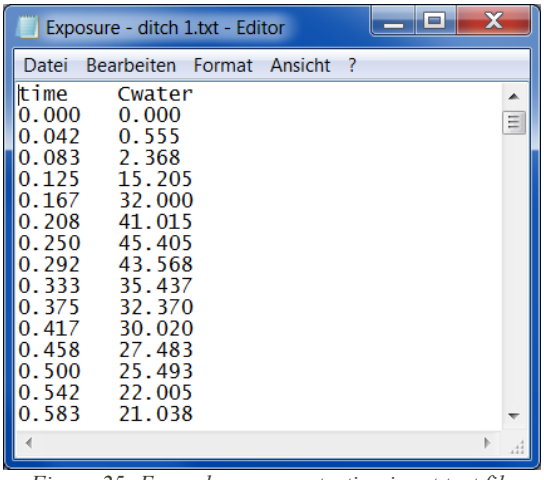


Figure 25: Exemplary concentration input text file

In the next tab "TK-TD Parameters", the user can enter the parameter values as well as supplementary information as the name of the substance, species and reference of the parameter values. These information are obligatory.

The user can enter manually parameter values or choose a parameter data set by clicking at data base. The data base contains all by the user saved TK-TD parameter data sets. We choose as an example the TK-TD parameter set of Ducrot et al. 2016 (Figure 26).

Furthermore, the user can decide whether to use the reduced GUTS model with only one TK parameter (dominant rate constant). In case of constant exposure, it is wise to check "explicit solution": The program uses the explicit solution function of the ordinary differential system, thus there is no discretization of the model system necessary and the calculation time is shorter. However, this is only possible in case of constant exposure.

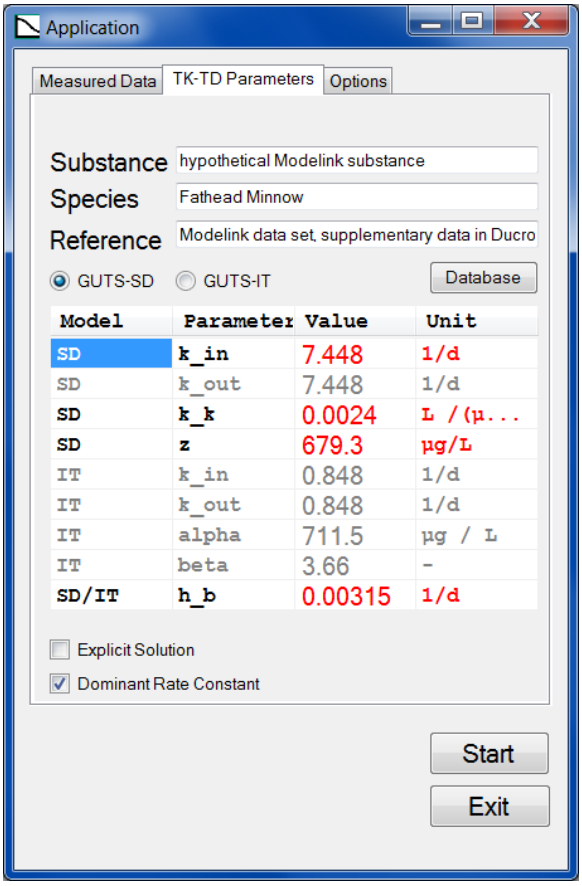


Figure 26: Application form: TK-TD Parameters Tab

In particular using a long time period leads to a longer calculation time. The maximum time period is 485 days. The calculation of LC50 is done similar as above and is in particular for long time periods time consuming.

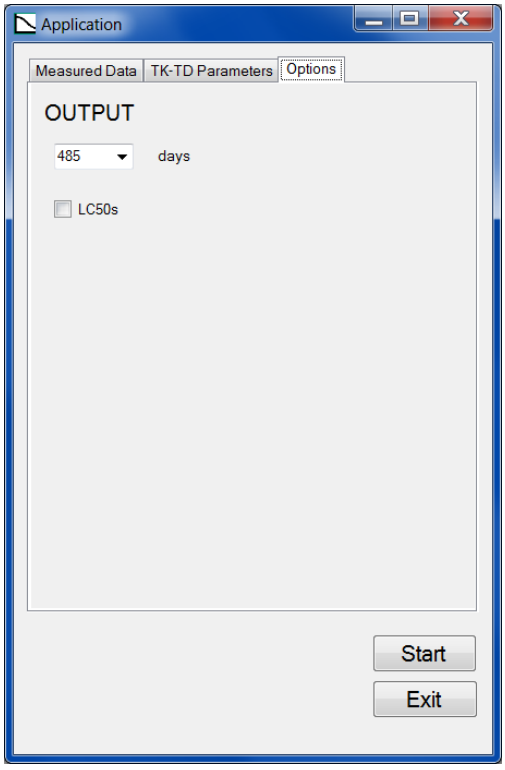


Figure 27: Application form: Options Tab

After the selections have been made, the "Start" button can be clicked to start the simulations.

The Application Result Form gives an overview of the entered information and the result of the application procedure. Therefore, the program provides different diagrams: the external concentration profile (input, Figure 28), the predicted internal concentration over time (Figure 29), the predicted hazard (GUTS-SD, Figure 30) or distribution function (GUTS-IT), as well as survival probability (output, Figure 31). In Figure 30, one can see that there is no mortality due to the exposure. The survival probability decrease only due to the background mortality (Figure 31).

Furthermore, the user can save, print or copy the diagrams. Additionally the user can create an output file (txt format), which can e.g. be imported in Excel for further evaluation of the results.

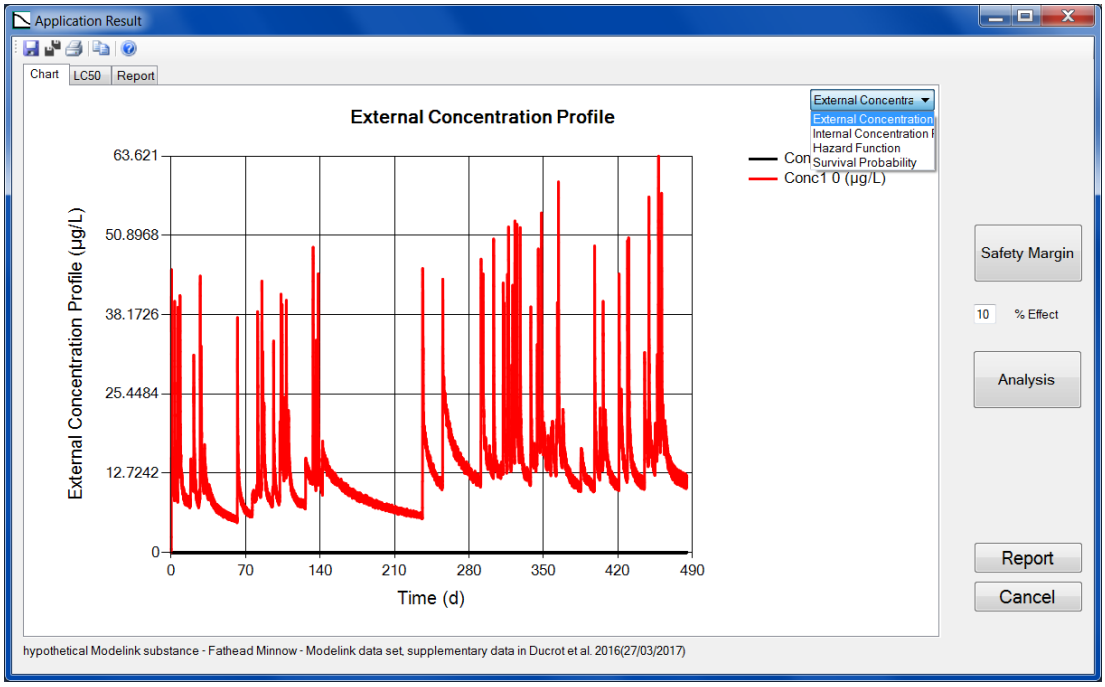


Figure 28: External concentration profile in the application result form

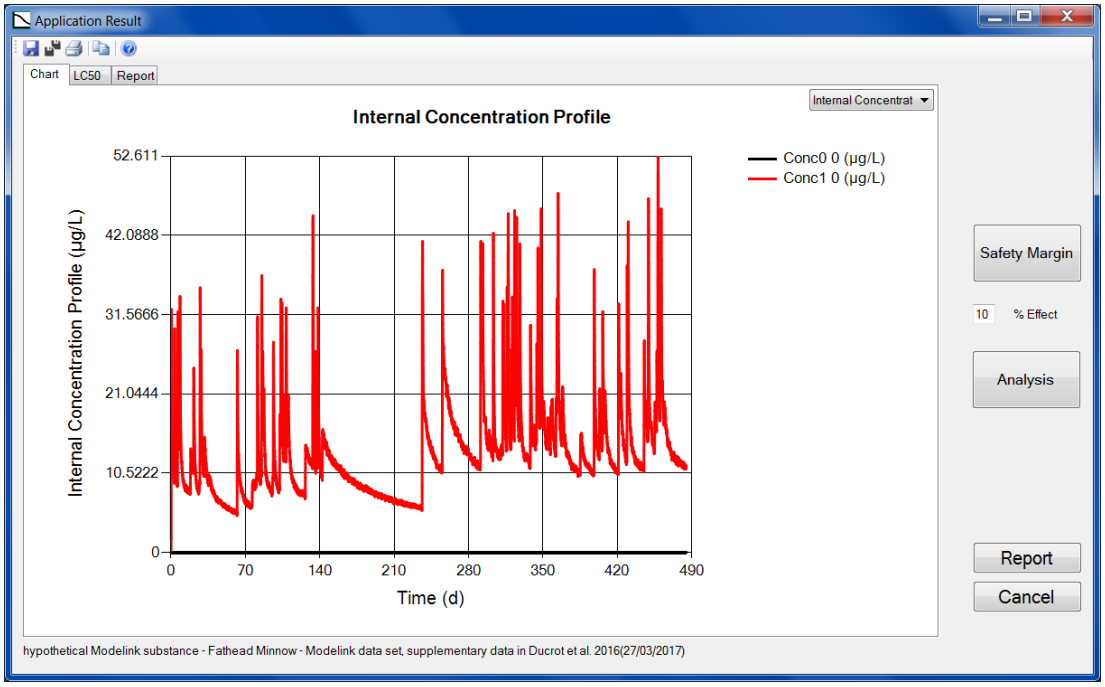


Figure 29: Internal concentration profile in the application result form

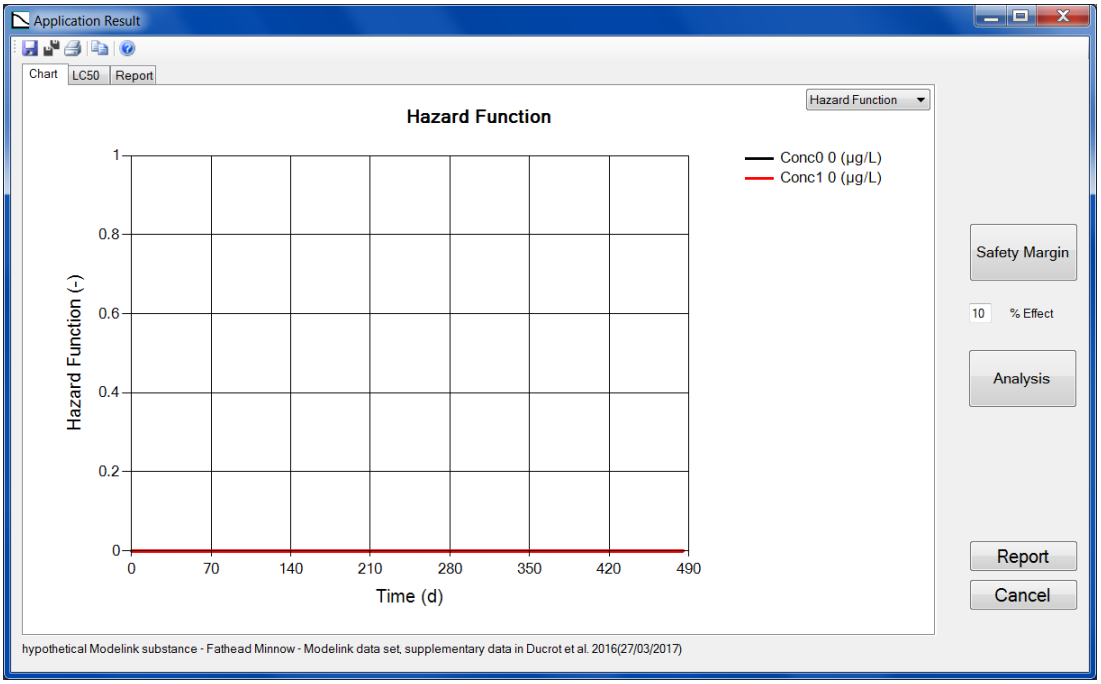


Figure 30: Hazard function in time

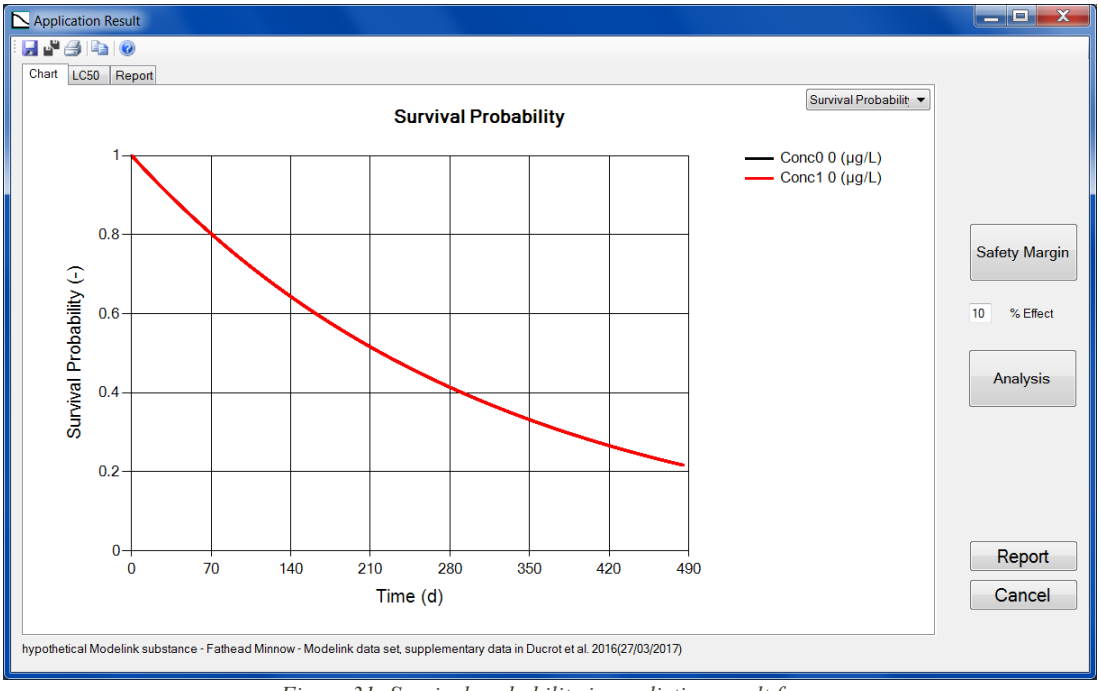


Figure 31: Survival probability in prediction result form

We repeat our calculation. However, we set the background mortality to zero to observe only the effect of the active substance. As result, we obtain a survival probability of 100% for the total time period of 485 days (Figure 32).

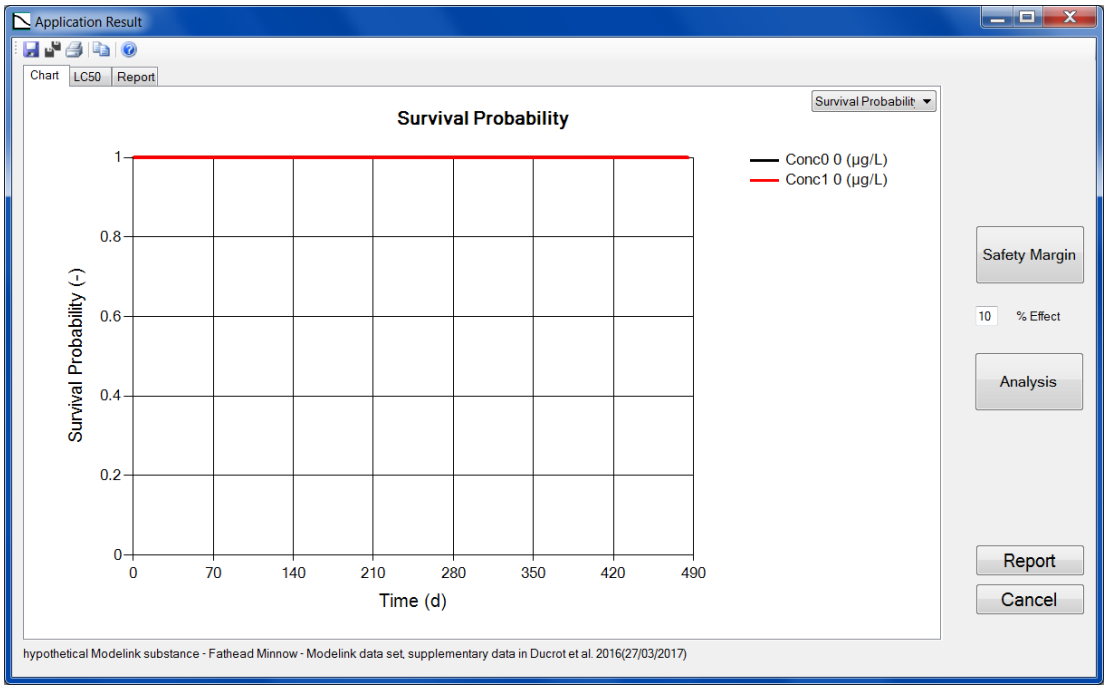


Figure 32: Survival probability with background mortality equal to zero

The button “Safety Margin” calculates the multiplication factor that yields x percentage effect, answering the question by which factor do we have to multiply the exposure profile to obtain x percentage effect on the survival probability.

We want to know by which factor we have to multiply the exposure profile to obtain a mortality of 10%. This can be done by clicking at the button “Safety Margin”.

The calculation of safety margin is done by using the Intermediate value theorem. The program searches in an interval $[0,100 \cdot z]$ for GUTS-SD respectively $[0,100 \cdot \alpha]$ for GUTS-IT. The result is only given in a label on the right side next to the safety margin button and the diagram shows the graphical result. In this particular case, we obtain a safety margin of 15.247 (Figure 33).

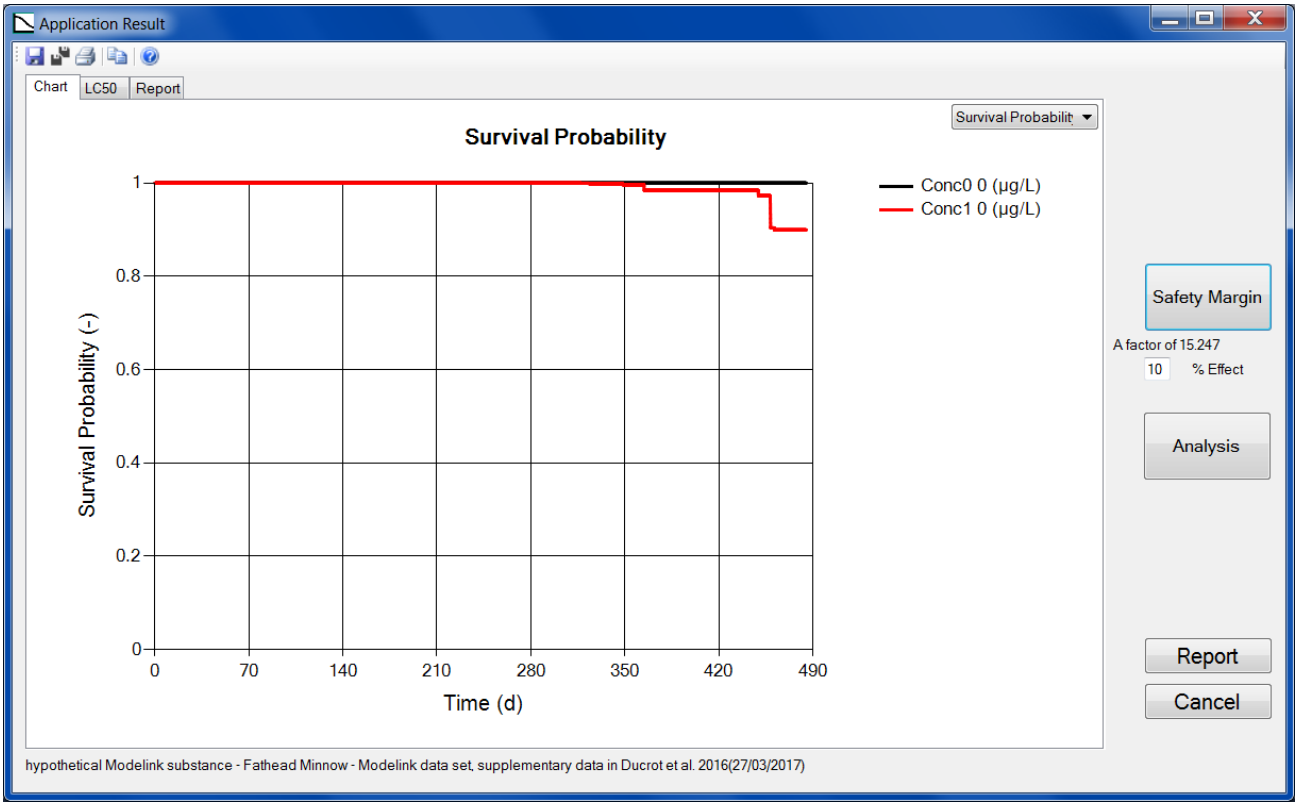


Figure 33: Application result form after clicking at the button "Safety Margin"

Sensitivity analysis is a systematically analysis of the effect of relative changes in parameter value on the model output. For this, only one parameter is changed in systematic manner whereas the other parameters are fixed to their fitted values. It is possible to calculate sensitivity coefficients by dividing the scaled parameter change by the scaled model output (EFSA PPR Panel 2014).
Local sensitivity analysis is when the value of a parameter is varied in a small area around the proper parameter value. Global sensitivity analysis is the consideration of the total range of the parameter (Pianosi et al. 2016). In Grimm et al. 2014 "global sensitivity analysis" is when several or all parameters are varied over their whole ranges.

In principle both forms, sensitivity and uncertainty analysis, are the same. They differ in their "Standard Analysis", namely for the sensitivity analysis we propose to choose the chosen parameters to vary one at a time in equidistant steps. For uncertainty analysis, we vary all parameters at one time using a certain distribution.

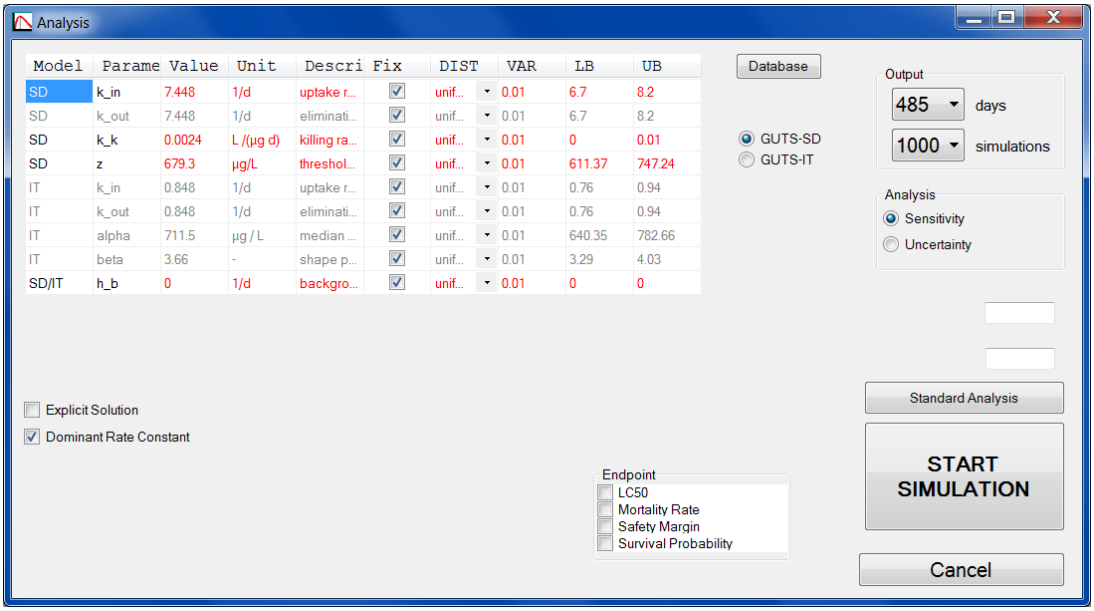


Figure 34: Sensitivity analysis form

The following possibilities are included. One can set a parameter value, growth as well as TK-TD parameter fixed to a constant value or variable in each simulation.
If a parameter is variable, the user can choose between lognormal, normal, uniform, equidistant, triangular and Marsaglia. Furthermore, the user can decide between several endpoints: LC50, mortality rate, safety margin and survival probability.

Distribution	Algorithm
Marsaglia	Let $u_1, u_2 \approx U(0,1)$ be stochastically independently uniformly distributed random variables. Set $v_i := 2 \cdot u_i - 1$ for $i = 1,2$ as long as it yields that $w := v_1^2 + v_2^2 < 1$. Then the random numbers $z_1 := v_1 \sqrt{-2\log(w)/w}$ and $z_2 := v_2 \sqrt{-2\log(w)/w}$ are normally distributed.
Triangular	Let $a, b \in \mathbb{R}_+$ with $a < b$ the range of the distributed values. Furthermore, let $c \in (a,b)$ be the peak value. Let $u \in U(0,1)$ be a uniformly distributed random variable. If $u < \frac{c-a}{b-a}$ then $z := a + \sqrt{(b-a)\cdot(c-a)\cdot u}$ else $z := b + \sqrt{(b-a)\cdot(c-a)\cdot(1-u)}$
Uniform	Let $a, b \in \mathbb{R}_+$ with $a < b$ the range of the distributed values. Let $u \in U(0,1)$ be a uniformly distributed random variable. Set $z := (b-a)\cdot u + a$ to obtain a uniformly distributed random variable $z \in [a,b]$.
Normal	Let $x, y \in \mathbb{R}$ and $b \in \mathbb{R}_+$ the border. Let $\mu \in \mathbb{R}_+$ be the expected value and σ^2 the variance. Let $u_1, u_2 \approx U(0,1)$ be uniformly distributed random variables. Do $u_1, u_2 \approx U(0,1)$ $x = 2 \cdot b \cdot u_1 + (\mu - b)$ and

	$y = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2 \cdot \sigma^2} \right)$ until $u_2 < y$. Then is $x \in \mathrm{N}\left(\mu, \sigma^2\right)$ a normally distributed random variable (with corresponding probability density function value $y \in \mathbb{R}$).
Equidistant	Let $a, b \in \mathbb{R}_+$ with $a < b$ the range of the distributed values. Furthermore, let $N \in \mathbb{N}$ be the total number of simulations. For each simulation $i = 1, \cdots, N$, we calculate the number $z_i \in [a, b]$ with $$z_i = (a + (b-a)) \cdot \frac{i}{N-1}.$$
Lognormal	Let $\mu \in \mathbb{R}_+$ be the expected value and σ^2 the variance. Furthermore, let $x \in \mathrm{N}(0,1)$ be a normally distributed random variable. Set $$z := \sigma^2 \cdot x + log(\mu)$$ to obtain a log normally distributed random variable $z \in \mathbb{R}$.

Appendix

Structure of an Experimental Data File

This is an Excel file summarizing the data from one experimental test, shown in Figure 4 and Figure 21. The structure is fixed.

In B1, the name of the study can be entered. In B2, the code of the study can be entered and in B3 the species and in B4-B5 possible comments can be inserted by the user. All this information is obligatory.

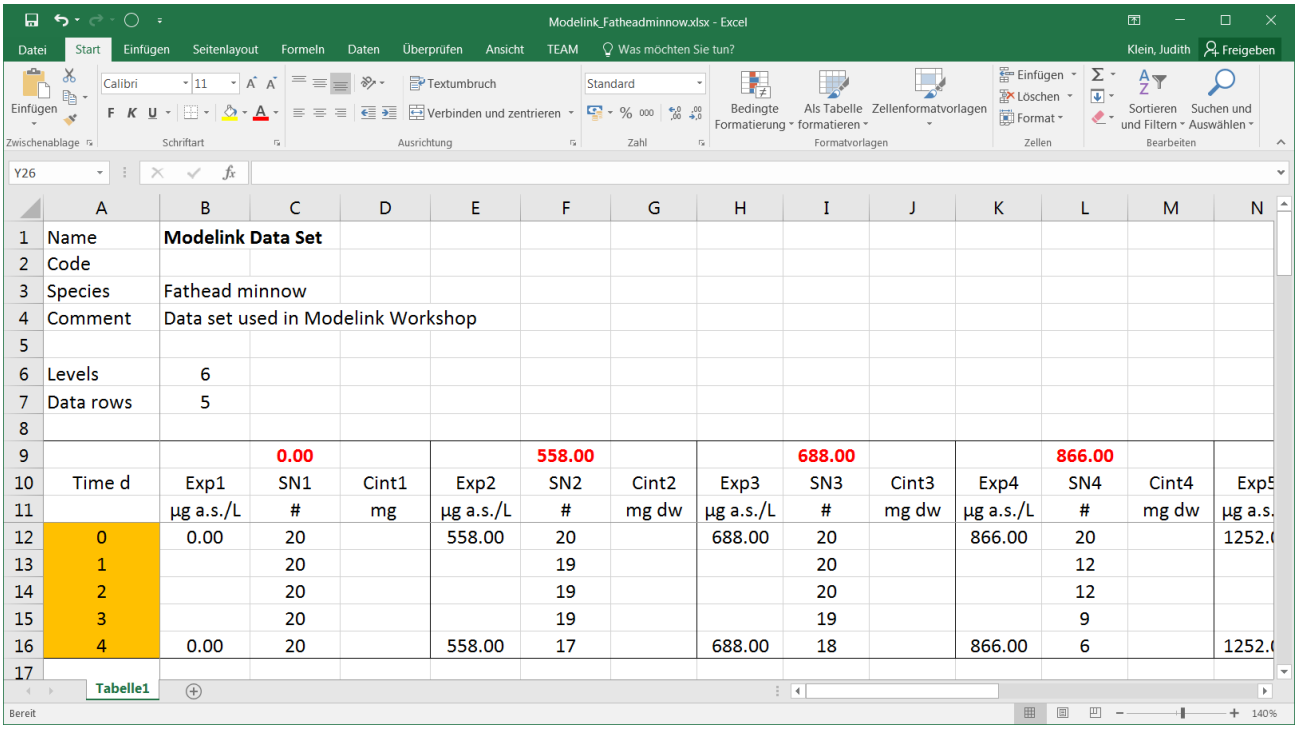
B6 (number of exposure levels including control) and B7 (number of entered data rows) are necessary inputs. In this case, Figure 35, the number of entered data is from row 12 to 16 and thus, altogether 5 rows.

In the first column the user can enter the time data in days. In this case, we have a test that is in total 4 days long with an exposure period of 4 days (orange). The entered time points indicate the time point when the number of survivors was counted.

After this column, the experimental data concerning the treatments is added. Here, we start with the control data (B-D). The first column (B) is the measured concentration in water. Additionally in C9 the nominal concentration is given. The second column (C) contains the counted number of survivors, and the third column (I) contains measurements on internal concentration, if available. This structure is the same for all treatments.

For the time points for which no external concentration data is available, the program interpolates linearly or exponentially.

Below or to the right from the table with the input data, summary statistics, diagrams or additional data can be inserted but will not be used by GUTS-3S.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Name	**Modelink Data Set**												
2	Code													
3	Species	Fathead minnow												
4	Comment	Data set used in Modelink Workshop												
5														
6	Levels	6												
7	Data rows	5												
8														
9			0.00			558.00			688.00			866.00		
10	Time d	Exp1	SN1	Cint1	Exp2	SN2	Cint2	Exp3	SN3	Cint3	Exp4	SN4	Cint4	Exp5
11		µg a.s./L	#	mg	µg a.s./L	#	mg dw	µg a.s./L	#	mg dw	µg a.s./L	#	mg dw	µg a.s.
12	0	0.00	20		558.00	20		688.00	20		866.00	20		1252.(
13	1		20			19			20			12		
14	2		20			19			20			12		
15	3		20			19			19			9		
16	4	0.00	20		558.00	17		688.00	18		866.00	6		1252.(
17														

Figure 35: Example Input (Excel file), Data based on Ducrot et al. 2016

References

Ashauer, Roman (2010). "Toxicokinetic–toxicodynamic modelling in an individual based context-Consequences of parameter variability". In: Ecological Modelling 221.9, pp. 1325–1328. doi:10.1016/j.ecolmodel.2010.01.015

Ducrot, Virginie, Roman Ashauer, Agnieszka J. Bednarska, Silvia Hinarejos, Pernille Thorbek, and Gabriel Weyman (2016). "Using toxicokinetic-toxicodynamic modeling as an acute risk assessment refinement approach in vertebrate ecological risk assessment". In: Integrated environmental assessment and management 12.1, pp. 32–45. doi:10.1002/ieam.1641

EFSA PPR Panel (2013). "Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters". In: EFSA Journal 11.7, p. 3290. doi:10.2903/j.efsa.2013.3290

Hawker, Darryl W. and W. Des Connell (1986). "Bioconcentration of lipophilic compounds by some aquatic organisms". In: Ecotoxicology and Environmental Safety 11.2, pp. 184–197. doi:10.1016/0147-6513(86)90063-1

Jager, Tjalling, Carlo Albert, Thomas G. Preuss, and Roman Ashauer (2011). "General Unified Threshold Model of Survival - a Toxicokinetic-Toxicodynamic Framework for Ecotoxicology". In: Environmental Science & Technology 45.7, pp. 2529–2540. doi:10.1021/es103092a

Jager (2016). Differential Equations and Likelihood Functions, a refresher. http://www.debtox. info /downloads/coursemat/refresher.pdf (visited on 02/01/2018)

Kaiser, K. L. (2012). QSAR in Environmental Toxicology - II: Proceedings of the 2nd International Workshop on QSAR in Environmental Toxicology, held at McMaster University, Hamilton, Ontario, Canada, June 9–13, 1986. Springer Netherlands. ISBN 9789400939370

Mackay, Donald (1982). "Correlation of bioconcentration factors". In: Environmental Science & Technology 16.5, pp. 274–278. doi:10.1021/es00099a008

Meeker, William Q. and Luis A. Escobar (1995). "Teaching about Approximate Confidence Regions Based on Maximum Likelihood Estimation". In: The American Statistician 49.1, p. 48. doi:10.2307/2684811

Moerbeek, Mirjam, Aldert H. Piersma, and Wout Slob (2004). "A comparison of three methods for calculating confidence intervals for the benchmark dose". In: Risk analysis: an official publication of the Society for Risk Analysis 24.1, pp. 31–40. doi:10.1111/j.0272-4332.2004.00409.x

OECD (1992). “Test No. 203: Fish, Acute Toxicity Test”. In: OECD Guidelines for the Testing of Chemicals Section 2. doi:10.1787/9789264069961-en

Petersen, Gitte I. and Preben Kristensen (1998). “Bioaccumulation of lipophilic substances in fish early life stages”. In: Environmental Toxicology and Chemistry 17.7, pp. 1385–1395. doi:10.1002/etc.5620170724

Pianosi, Francesca, Keith Beven, Jim Freer, Jim W. Hall, Jonathan Rougier, David B. Stephenson, and Thorsten Wagener (2016). “Sensitivity analysis of environmental models: A systematic review with practical workflow”. In: Environmental Modelling & Software 79, pp. 214–232. doi:10.1016/j.envsoft.2016.02.008

Abbreviations

AR	Absolute Residuals
BCF	Bioconcentration Factor
Conc	Concentration
EF	Model Efficiency
Exp	Experimental Data
FOCUS	FOrum for international Coordination of pesticide fate models and their USe
GUTS-3S	General Unified Threshold Model of Survival – 3 Steps
IME	Fraunhofer Institute for Molecular Biology and Applied Ecology
IT	Individual Tolerance
LC50	median lethal concentration, that concentration killing 50% of test species
ODE	Ordinary Differential Equation
OECD	Organisation for Economic Co-operation and Development
PEC	Predicted Environmental Concentration
PPR	Panel on Plant Protection Products and their Residues
QSAR	Quantitative Structure-Activity Relationship
SD	Stochastic Death
SR	Squared residuals
SRMSE	Squared Root Mean Squared Error
STE	Scaled Total Error
TD	Toxicodynamic
TK	Toxicokinetic