

## Neue statistische Verfahren

### Hampel Ausreißer Test

#### ECx-Bestimmung metrische Daten mittels nicht-linearer Regression

- 2-3-4 Parameter Normal, Sigmoid (Bruce Versteeg)
- 2-3-4 Parameter Logistic
- 2-3-4 Parameter Weibull
- Gewichtungen: relativ, Poisson, nach Variabilität
- Optimierungsmethoden: Levenberg-Marquardt, Downhill-Simplex
- Vertrauensbereichsberechnung: Monte Carlo Simulation, Boots-Trap-Verfahren

#### ECx-Bestimmung quantale Daten:

- Spearman-Kärber / Trimmed Spearman Kärber

#### Multiple Tests zur NOEC-Bestimmung:

- Stepdown Jonckheere-Terpstra
- Bonferroni Median Test
- Stepdown Cochran Armitage

#### Tests auf Monotonie

linear+quadratic contrasts

(Rao-Scott-) Cochran Armitage Trend-Test

Jonckheere-terpstra-Trend-Test

Tarone-Test auf **Extrabinomiale Varianz**

## Neue und aktualisierte Workbooks und Masterbooks

Eine vollständige Liste *aller* verfügbaren Biotest-Datenvorlagen finden Sie im Handbuch.

### Neue Biotest-Workbooks

OECD 205 Avian Dietary Toxicity Test

OECD 206 Avian Reproduction Test

OECD 213 Honeybees Acute Oral Toxicity Test

OECD 214 Honeybees Acute Contact Toxicity Test

OECD 223 Avian Acute Oral Toxicity Test

OECD 235 Chironomus Acute Test

OECD 236 Fish Embryo Acute Toxicity Test (FET)

OECD 237 Honeybees Larval Toxicity Test

OECD 238 Myriophyllum spicatum Sediment-Free Toxicity Test

OECD 239 Myriophyllum Water-Sediment Toxicity Test

IOBC Typhlodromus

IOBC Aphidius

ISO 10872 Caenorhabditis

ISO 10706 Daphnia magna Reproduction

ISO 20665 Ceriodaphnia

ISO 16191 Myriophyllum aquaticum

DIN EN ISO 15088 Eggs of Danio rerio  
DIN EN ISO 17512-1 Earthworm Avoidance  
DIN EN ISO 17512-2 Collembola Avoidance

OPPTS (=OCSP) 850.5400 Alga Toxicity

### **Aktualisierte Biotest-Workbooks** (Anpassung an neuere Guidelines)

OECD 209 Activated Sludge Respiration Inhibition Test (2010)  
OECD 210 Fish Early Life Stage Test (2013)  
OECD 211 Daphnia magna Reproduction Test (2012)

## Änderungen für bestimmte Biotests

### **Algentest Workbooks auf Basis von Fluoreszenzmessungen (OECD 201, DIN 8692)**

Die Kalibrierung der Fluoreszenzmesswerte wurde geändert: die Kurve geht jetzt stets durch den Nullpunkt. Workbookvorlagen aus Version 2.10.05 können deshalb mit 3.0 nicht ausgewertet werden (nur nach Austausch des jeweiligen Kalibrierungssheets) und die Ergebnisse aus 2.10.05 und 3.0 können sich – je nach Kalibrierungsdaten – unterscheiden.

### **OECD 201**

Zusätzliche workbooks für spezielle Test- bzw. Messdesigns:

„OECD201 AlgaeGrowthInhibition(Extinction duo).xls“ Ermöglicht die Eingabe verschiedener Kalibrationskurven für zwei unterschiedliche Extinktionsbereiche

### **OECD 209**

ToxRat enthält zwei verschiedene OECD 209-Workbooks, eine für parallele Messungen und eine für serielle Messungen. Bei Letzterer wird gemäß den Empfehlungen des UBA-Dokuments<sup>1</sup> zur neuen OECD 209-Richtlinie ein paarweiser Vergleich zwischen den Kontrollen, die am Anfang und am Ende der Messreihe gemessen wurden, durchgeführt, um zu prüfen, ob sich die Hintergrundatmung während der Messdauer verändert hat. Eventuelle Unterschiede werden damit dokumentiert, die weiteren Auswertungen erfolgen jedoch auf Basis aller sechs Kontrollreplikate, wenn der Benutzer sich für „poolen“ entscheidet.

### **OECD 210**

Enthält jetzt je ein Dateneingabeblatt für Trockengewicht und Frischgewicht. Frisch- und Trockengewicht können somit parallel ausgewertet werden, alternativ kann ein (oder auch beide) sheet(s) leer bleiben.

### **OECD 211**

Gemäß der aktualisierten Testrichtlinie von 2012 wird anhand eines Trendtests geprüft, ob signifikante Mortalität vorliegt. ToxRat 3.0. berechnet die Anzahl Nachkommen automatisch sowohl auf Basis der eingesetzten Muttertiere (Cumulative offspring of Introduced, COI) als auch auf Basis der Überlebenden (Cumulative Offspring of Survivors, COS). Je nach Ergebnissen muss der Benutzer entscheiden, welchen Endpunkt er berichtet (laut Testrichtlinie: wenn signifikante Mortalität, dann COI, wenn keine signifikante Mortalität dann COS).

---

<sup>1</sup> „Frage und Antwort Dokument des Umweltbundesamtes zur neuen Richtlinie OECD 209 „Activated Sludge, Respiration Inhibition Test (Carbon and Ammonium Oxidation) (22.7.2010), Version vom 2.3.2012“

ToxRat 3.0. kann auch Testdesigns mit beliebigen, unterschiedlichen Replikanzahlen in Kontrolle und treatments auswerten (unbalanced designs).

### OECD 218/219

Neue Einlese-Routine zur Überprüfung der Anzahl Eingesetzter, denn deren Anzahl kann nicht kleiner sein als Anzahl Geschlüpfter. Eine entsprechende Korrektur gab es auch schon in ToxRat 2.10.05, jetzt jedoch ist die durchgeführte Datenkonsolidierung transparenter: ToxRat erstellt zusätzliche Datenblätter und speichert die modifizierte Datei mit spezifischem Namenszusatz ab. Einzelheiten und Konsequenzen für die Handhabung, insbesondere bei der Eingabe neuer Daten, entnehmen Sie bitte dem Handbuch (Kapitel 5.3., S. 81ff).

## Kompatibilität mit früheren Versionen

Alle bestehenden Workbooks und Masterbooks wurden überarbeitet und insbesondere die Angaben im Sheet General Notes, die die Auswertung steuern, wurden vereinheitlicht. Wird ein Workbook aus ToxRat 2.10.05 oder früher mit ToxRat 3.0 eingelesen, so erscheint ein Warnhinweis und Sie werden aufgefordert, stattdessen eine der neuen Vorlagen zu verwenden. Zwar ist ToxRat 3.0 in der Regel abwärtskompatibel, d.h. Sie können die Workbooks aus früheren Versionen auch mit ToxRat 3.0 einlesen und auswerten. Wir empfehlen jedoch dringend, für alle neu einzugebenden Daten nur noch die neuen Vorlagen zu verwenden, da nur für diese alle Funktionen durchgetestet wurden.

Diejenigen Workbooks und Masterbooks, die speziell für ToxRat 3.0 neu entwickelt bzw. aktualisiert wurden (siehe Liste oben), können von früheren ToxRat Versionen nicht eingelesen und ausgewertet werden!

ToxRat 3.0 verlangt – wie frühere Versionen – das Dateiformat xls bzw. xlt. Falls Sie Dateien in MS Excel bearbeiten, stellen Sie bitte sicher, dass Sie in diesem Format abspeichern.

## Handbuch

Das Handbuch wurde vollständig neu erstellt und ist jetzt sowohl in deutscher als auch in englischer Sprache verfügbar. Der Aufbau folgt der praktischen Vorgehensweise bei der Biotestauswertung. Es beschreibt nicht nur die Programmbedienung, sondern erläutert auch die Kriterien zur Testauswahl und erklärt, was die Ergebnisse bedeuten.

## Validierung

Da diverse statistische Verfahren ergänzt oder verändert wurden, erfordert ToxRat 3.0 eine neue Validierung. Sie finden eine vollständige Übersicht von Musterergebnissen aus allen enthaltenen statistischen Methoden in Kapitel 6 des Validierungsdokumentes („Results for Routine Validation of ToxRat Computation“). Diese können Sie anhand der mitgelieferten Dateien validation data 1-6 nachvollziehen.

Zusätzlich zur Routinevalidierung werden alle im Programm verwendeten Algorithmen und Formeln in Kapitel 1-5 des Validierungsdokumentes beschrieben und die Ergebnisse mit Literaturergebnissen oder unabhängigen Berechnungen verglichen. Für die nichtlineare Regression erfolgt dies aktuell am Beispiel des Levenberg-Marquardt-Algorithmus, die detaillierte Darstellung des alternativen Downhill-Simplex-Algorithmus wird noch ergänzt. Musterergebnisse für die Validierung (Kapitel 6, siehe oben) sind jedoch bereits enthalten.

## Installation

ToxRat 3.0. kann unter WIN XP und WIN 7 parallel zu einer bestehenden Version ToxRat 2.10.05 installiert werden.

## Grundsätzliche Änderungen und Ergänzungen gegenüber der ToxRat Version 2.10.05

### Automatische Auswertung („RUN“) auch bei Generic Workbooks

Ab ToxRat 3.0 können Sie auch in den Generic Workbooks eine programmgesteuerte Auswertesequenz aus Vortests und passendem Haupttest für die NOEC- und Limit-Bestimmung mit einem Knopfdruck veranlassen – dazu gibt es die Menüs „Find NOEC“ und „FIND Limit Level“, und in diesen analog zu den Biotest-Workbooks einen RUN-Button.

### Menü-Führung

Sie erreichen das Auswahlfenster für die Methoden zur Hemmwertbestimmung bei Generic Workbooks jetzt direkt über den Menüpunkt „Find Effect Level“ – der frühere (Um-) Weg über „Options“ ist nicht mehr erforderlich. Zum Starten der Berechnung klicken Sie auf den RUN Button im Auswahlfenster.

### Signifikanzniveaus

Bei der Testsequenz für die NOEC-Bestimmung können die Signifikanzniveaus für Vortests und Haupttest jetzt unabhängig voneinander eingestellt werden.

Die Voreinstellung für die Signifikanzniveaus für Vortests auf Normalverteilung und Varianzhomogenität ist jetzt 0,01 (vorher: 0,05), um a) unnötig schwache Haupttests zu vermeiden und um b) der alpha-Kumulation entgegenzuwirken.

### Einstellungen für die Auswertungen („Settings“)

ToxRat 3.0 erzeugt optional bei jeder Auswertung auch eine sogenannte settings-Tabelle. Darin sind alle für die aktuelle Auswertung verwendeten Einstellungen dokumentiert, und es ist dargestellt, ob diese von den programmseitigen Voreinstellungen abweichen. Die settings-Tabelle wird per default als letzte Ergebnistabelle (hinter „Summary“) gezeigt.

Die (Vor-) Einstellungen, die ToxRat für eine Auswertung anwenden soll, sind in sogenannten settings-Dateien gespeichert. Diese sind jetzt Biotestspezifisch organisiert, d.h. es gibt nicht mehr eine Gesamt-Settings-Datei für das ganze Programm, sondern individuelle Settings-Dateien für jeden Workbook-Typ (z.B. set\_OECD210.stp). Wann immer Sie ein Workbook eines bestimmten Typs öffnen, liest ToxRat die Einstellungen aus der zugehörigen Settings-Datei ein. Im „Werkzustand“ sind das die programmseitigen Voreinstellungen, später ggfls. Ihre individuellen Einstellungen. In zukünftigen Updates können Sie somit für Biotests, an denen nichts geändert wurde, Ihre individuellen Settings-Dateien aus früheren Versionen übernehmen.

**Wichtiger Hinweis:** die Umstellung auf Biotestspezifische Settings-Dateien hat zur Folge, dass in ToxRat 3.0 bis auf weiteres **keine individuellen Reportprofile gespeichert** werden können.

### Nicht-lineare Regression

ToxRat 3.0 beinhaltet Methoden zur nicht-linearen Regression ausschließlich für metrische Daten, denn speziell für metrische Daten ist die nicht-lineare Regression eine wichtige Alternative zur linearen Regression, da sie ohne Normalisierung auskommt (Ausnahme: 2-Parameter CDFs). Bei quantalen Daten besteht das Problem der Normalisierung nicht, deshalb ist die lineare Regression bei quantalen Daten grundsätzlich eine gute Wahl.

Da die nicht-lineare Regression in den meisten Guidelines als Methode der Wahl empfohlen wird, wurden die ToxRat Voreinstellungen entsprechend angepasst: default-Einstellung für Hemmwertbestimmung bei metrischen Variablen ist in der Regel die 3-Parameter Nicht-lineare Regression. Angesichts der Vielzahl möglicher Varianten und Einstellungsmöglichkeiten bei der nicht-linearen Regression führen die Standardeinstellungen je nach Datensatz jedoch möglicherweise nicht zu einem erfolgreichen Ergebnis. In diesem Fall müssen Sie verschiedene Varianten ausprobieren, um zu einem sinnvollen Ergebnis zu kommen. Eine Orientierungshilfe zu den Faktoren, die variiert werden können sowie zu den Gütekriterien zur Beurteilung der

Ergebnisse finden Sie im Handbuch. Beachten Sie bitte auch die Beispielergebnisse aus dem Validierungsdokument (Kapitel 6.5): anhand ein und derselben Datei wird hier gezeigt, wie sich die Wahl der Funktion und des Algorithmus auswirken.

Wichtig: Der Downhill-Simplex-Algorithmus und das damit verbundene Bootstrapping Verfahren zur Vertrauensbereichsbestimmung ist sehr rechenintensiv – je nach Datensatz (Anzahl der Variablen und Messzeitpunkte) kann es durchaus mehrere Minuten (!) dauern, bis die Ergebnisse gezeigt werden. Brechen Sie das Programm nicht ab, solange das Aktivitätssymbol gezeigt wird! Tipp: Zum „Ausprobieren“, ob die Methode grundsätzlich geeignet ist, kann die Anzahl der Resamplings deutlich reduziert werden.

### Programmbibliotheken

Um die Größe der ToxRat-Programmdatei zu begrenzen, wurden diverse Biotest-spezifische Textausschnitte in dll-Dateien ausgelagert, welche im Programmverzeichnis gespeichert sind. Auf die dll-Dateien wird während der Auswertungen zugegriffen, hierfür sind (wie bisher auch schon) Leserechte im Programmverzeichnis erforderlich.

### Datentabellen und Ergebnistabellen

Vom Programm bei einer Auswertung erzeugte Datentabellen sind (wie bisher auch) über den Menüpunkt „Tables“ direkt zugänglich. Aus Gründen der Übersichtlichkeit werden Datentabellen jedoch nicht mehr in Form eigener, zusätzlicher Datenblätter in das Workbook eingefügt. Dies ist nun den Ergebnistabellen vorbehalten („Tabs<variablenname>“).

### Zusätzliche Variable möglich für „quantal responses“: #escaped

Die Generic Workbooks Quantal respnses ermöglichen neben „#dead, #survived# und #emerged jetzt auch die Dateneingabe für #escaped. Eine entsprechende Dateivorlage finden Sie im Verzeichnis für Demo-Workbooks und Masterbook.

## Änderungen bei bestehenden statistischen Methoden

### Levene Test

In ToxRat 2.10.05 wurde beim Levene Test mit quadrierten Residuen gerechnet  $(x - x_{\cdot})^2$ . Dies ist eine von mehreren in der Originalveröffentlichung genannten Varianten<sup>2</sup>. SPSS und R verwenden eine andere Variante, nämlich die absolute Abweichung  $|x - x_{\cdot}|$ . Um in Zukunft übereinstimmende Ergebnisse mit SPSS und R zu erzielen, verwendet ToxRat 3.0. ebenfalls die absoluten Abweichungen. Als Konsequenz ergeben sich Unterschiede beim Levene Test zwischen ToxRat 2.10.05 und ToxRat 3.0.

### Lineare Regression - Vertrauensbereichsberechnung

Bei der Linearen Regression für metrische Daten ist jetzt die Berücksichtigung der Kontrollvarianz bei der Vertrauensbereichsberechnung möglich (gemäß Annex 5 OECD 201). Die Vertrauensbereiche werden dadurch in der Regel etwas weiter. Diese Varianzkorrektur ist per default eingeschaltet – damit Sie Ergebnisse aus früheren ToxRat Versionen reproduzieren können, ist sie jedoch optional abschaltbar (Näheres siehe Handbuch).

Als zusätzliches Verfahren für die Vertrauensbereichsberechnung bei linearer Regression ist bei metrischen Daten jetzt das Bootstrapping verfügbar.

---

<sup>2</sup> Levene, H. (1960): Robust tests for equality of variances. In: Contributions to Probability and Statistics. Essays in Honours of H. Hotelling, Stanford, 278-292

Der t-Wert für die Berechnung des 95%-Vertrauensbereichs bei der linearen Regression wird jetzt mittels Funktion auf 14 Nachkommastellen genau bestimmt. In früheren Versionen wurde der auf zwei Stellen gerundete Wert aus der Normalverteilungstabelle (= 1,96) verwendet. Durch die höhere Genauigkeit können sich die Vertrauensbereichsgrenzen aus ToxRat 3.0 leicht von denen aus früheren Versionen unterscheiden.

### **Lineare Regression, Algorithmus: Maximum Likelihood Regression - Anzahl RUNS, weights**

Zwecks höherer Genauigkeit auch bei Datensätzen mit schlechter Dosis-Wirkungsbeziehung wurde das Abbruchkriterium für die Anzahl der durchgeführten Iterationen verändert. Dies führt je nach Datensatz zu einer anderen Anzahl Durchläufe (RUNS) und damit zu anderen Gewichtungen (weights) für die treatments als in ToxRat 2.10.05 (dargestellt werden immer die weights aus dem letzten RUN). Als Konsequenz können sich auch die F-Werte und die p(F)-Werte unterscheiden.

### **Lineare Regression, Algorithmus: linear weighted regression - weights**

Die Funktion zur Weight-Bestimmung wurde durch eine genauere Funktion ersetzt. Deshalb ergeben sich je nach Datensatz andere Weights als in ToxRat 2.10.05 (dadurch andere F-Werte und andere p(F)-Werte)

### **Lineare Regression - Werte für F und p(F)**

Die Funktion zur Ermittlung von p(F) wurde durch eine genauere Funktion ersetzt; dies kann auch bei unveränderten Werten von F zu etwas veränderten p(F)-Werten gegenüber ToxRat 2.10.05 führen.

### **Berechnung der Minimum Detectable Difference (MDD)**

Die Funktion zur Bestimmung der t-Quantilen für das im Test verwendete Signifikanzniveau ("t(krit)-Werte") wurde ersetzt. Die neue Funktion ist bei Freiheitsgraden kleiner 4 genauer als die Funktion aus früheren ToxRat Versionen. Die resultierenden MDDs aus ToxRat 3.0 sind kleiner als die aus früheren ToxRat-Versionen.

Wichtig: Das betrifft nur die t(krit)-Werte für die Berechnung der MDDs, nicht die Berechnung der Stichproben-t-Werte, welche für die Signifikanzprüfung verwendet werden. Die Stichproben-t-Werte sind dieselben wie bisher. Auch die bisherige Funktion für die Berechnung der p(t)-Werte wurde durch eine Funktion ersetzt, die bei extrem großen t-Werten (d.h. für p-Werte kleiner als 0,001) wesentlich genauere, bzw. überhaupt Wahrscheinlichkeiten liefert (die frühere Funktion lieferte für extrem große t-Werte identische p-Werte oder einfach „Null“).

Bei paarweisen t-tests kann das zu leicht anderen MDD-Werten als in früheren ToxRat Versionen führen, die Ergebnisse der Signifikanzprüfungen sind jedoch unverändert gegenüber früheren ToxRat Versionen.

Bei multiplen t-Tests mit Bonferroni-Korrektur kann es ebenfalls zu geringfügig anderen MDD-Werten kommen als in früheren ToxRat Versionen, nämlich dann, wenn in einem Datensatz mit mehreren p(t)-Werten < 0,001 wegen der jetzt höheren Genauigkeit eine andere Sortierung bei der Bonferroni-Korrektur erfolgt und damit einer Behandlung ein anderer alpha(i)-Wert zugeordnet wird, welcher Grundlage für die Bestimmung des individuellen t(krit) und damit der MDD-Berechnung ist.

Wichtig: Auch dies hat keine Auswirkung auf die Ergebnisse der Signifikanzprüfungen mit einem Signifikanzniveau von 0,05 oder 0,01, da die zur Debatte stehenden p-Werte stets kleiner sind als 0,001.

### **Williams Test, Bildung der Likelihood Mean Werte (LhM)**

Die Likelihood-mean-Bildung erfolgt (wie in 2.10.05) ohne Einbezug der Kontrolle, wird jedoch ab ToxRat 3.0 in Abhängigkeit von der gewählten Testrichtung vorgenommen.

Für übliche Testdatensätze (abnehmende Messwerte mit steigender Konzentration, Fragestellung einseitig kleiner - bzw. zunehmende Messwerte mit steigender Konzentration, Fragestellung einseitig größer) hat dies keine Konsequenzen auf die Ergebnisse.

**Ableitung einer NOEC**

Bei heterogener Signifikanzlage erfolgt der Ausschrieb „expert judgement required“, d.h es wird programmseitig kein NOEC-Wert vorgeschlagen.

**Vergleich von Varianzen mittels F-Test beim paarweisen und multiplen t-Test**

Der F-Test wurde bisher einseitig durchgeführt. Um nicht unnötig früh Varianzhomogenität abzulehnen, wird der F-Test zum paarweisen Vergleich zweier Varianzen jetzt zweiseitig durchgeführt.

## Korrektur von Bugs

Es wurden diverse kleinere Funktionsstörungen korrigiert, die im Folgenden nicht im Einzelnen aufgeführt werden. Gelistet werden lediglich verbesserte Bugs mit Einfluss auf die Ergebnisse:

**OECD 201**

Wenn der Wert für Yield zufällig genau der Wert „-1“ ergab, wurde das als „missing value“ interpretiert, statt als Wert → verbessert

**OECD 208**

Validitätskriterien Seedlings Emergence und Seedlings Survival bei bestimmten Datenkonstellationen fälschlicherweise mit 0% angegeben, ist korrigiert

**OECD 218/219**

In ToxRat 2.10.05 wurde fälschlicherweise der Wert „-0,5 Tage“ als Startalter für die eingesetzten Larven verwendet statt „0,5 Tage“. Die Entwicklungsraten wurden dadurch überschätzt. Da jedes gesetzte, einheitliche Startalter willkürlich ist, und auch der Tag, der als Schlüpftag protokolliert wird, nur eine Genauigkeit von +/- 24 Stunden hat, sind die Entwicklungsraten ohnehin nur Näherungswerte.

In ToxRat 2.10.05 wurde stets mit demselben Startalter der eingesetzten Larven gerechnet (-0,5, siehe oben), unabhängig davon, welcher Wert im Fenster „Options / OECD 218/219“ angegeben war. Das Startalter der eingesetzten Larven ist jetzt editierbar, d.h. der im Fenster eingetragene Wert wird auch verwendet. Voreinstellung ist +0,5 Tage.

**Lineare Regression, p(F)**

In ToxRat 2.10.05 wurde für F-Werte *unterhalb von 1* ein fehlerhafter p(F)-Wert ausgegeben: statt p(F) wurde „1 minus p(F)“ ausgegeben. Der Fehler ist korrigiert.

Der Fehler führte dazu, dass z.B. für F=0,9 nicht das (korrekte) p(F) von 0,42 (42%) ausgegeben wurde, sondern  $1-0,42 = 0,58$  (58%).

F-Werte kleiner als 1 kommen nur dann vor, wenn keinerlei Dosis-Wirkungsbeziehung vorliegt und unabhängig vom Testdesign (Freiheitsgraden) ist bei allen F-Werten kleiner 1 sowohl p(F) als auch  $1-p(F)$  stets weitaus größer als das Signifikanzniveau von 5%, siehe Beispiel oben. Der Fehler hat deshalb keine Auswirkung auf die Schlussfolgerungen aus der linearen Regression mit ToxRat 2.10.05.