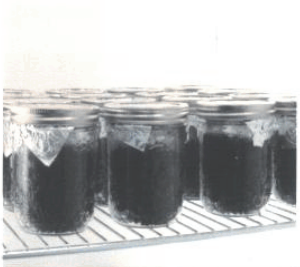
EPS 1/RM/43 – June 2004 (with June 2007 amendments)
Method Development and Applications Section
Environmental Technology Centre
Environment Canada



ENVIRONMENTAL
PROTECTION
SERIES

Biological Test Method: Tests for Toxicity of Contaminated Soil to Earthworms *(Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris)*

Environment Canada Environnement Canada

Canada

Biological Test Method:
Tests for Toxicity of Contaminated Soil to Earthworms (*Eisenia andrei,* *Eisenia fetida,* or *Lumbricus terrestris*)

Method Development and Applications Section
Environmental Technology Centre
Environment Canada
Ottawa, Ontario

Report EPS 1/RM/43
June 2004 (with June 2007 amendments)

Library and Archives Canada Cataloguing in Publication Data

Main entry under title:

Biological test method. Tests for toxicity of contaminated soil to earthworms (*Eisenia andrei, Eisenia fetida*, or *Lumbricus terrestris*)

(Report ; EPS 1/RM/43)
Issued also in French under title: Méthode d'essai biologique. Essais pour déterminer la toxicité de sols contaminés pour les vers de terre *Eisenia andrei, Eisenia fetida*, ou *Lumbricus terrestris*.
Includes bibliographical references.
ISBN 0-660-19366-3
Cat. No. En49-7/1-43E

1. Earthworms, Effect of pollution on.
2. Soil pollution – Measurement.
3. Earthworms – Toxicology.
4. Toxicity testing – In vitro – Methodology.
I. Environmental Technology Centre (Canada)
II. Canada. Environment Canada.
III. Series : Report (Canada. Environment Canada) ; EPS 1/RM/43.

QH90.8T68B56 2004 628.1'61 C2004-980325-5

ISBN 0-660-19366-3
Cat. No. En49-7/1-43E

Readers' Comments

Comments regarding the content of this report should be addressed to:

Richard Scroggins
Chief, Biological Methods Division
Environmental Technology Centre
Environment Canada
335 River Road
Ottawa, Ontario
K1A 0H3

Cette publication est aussi disponible en français. Pour l'obtenir, s'adresser à:

Publications de la Protection de l'environnement
Environnement Canada
Ottawa (Ontario)
K1A 0H3

Review Notice

This report has been reviewed by the staff of the Environmental Technology Advancement Directorate, Environment Canada, and approved for publication. Mention of trade names or commercial products does not constitute endorsement by Environment Canada for use. Other products of similar value are available.

Abstract

This document provides detailed procedures, conditions, and guidance for preparing for and conducting each of three discrete biological test methods for measuring soil toxicity using earthworms (Eisenia andrei, Eisenia fetida, *or* Lumbricus terrestris). *The test methods described herein are as follows:*

(1) an acute (14-day) lethality test, using adult or sub-adult earthworms;
(2) an acute (48 or 72 h) test for avoidance of contaminated soil by adult earthworms; and
(3) a 56-day (or, in certain instances, 63-day) test for effects of prolonged exposure of adult earthworms (cultured E. andrei *only) and their progeny to one or more samples or concentrations of contaminated or potentially contaminated soil.*

Each test method is conducted as a static (i.e., no renewal) test, using one or more samples of contaminated or potentially contaminated soil or one or more concentrations of chemical(s) or chemical product(s) spiked in negative control (or other) soil. Worms are fed (cooked oatmeal) only during the 56-day test for effects of prolonged exposure to the sample(s) of test soil.

The acute lethality test is performed as a 14-day test using 500-mL glass jars containing a measured wet weight equivalent to a volume of ~350 mL of test soil. Worms of one of the three species are used with five adult worms (if E. andrei/fetida*) or three adult worms (if* L. terrestris*) being used per jar. This test uses ≥5 replicates/treatment if a single-concentration test, ≥3 replicates/treatment if a multi-concentration test with* E. andrei or E. fetida, *and ≥5 replicates/treatment if a multi-concentration test with* L. terrestris.

The acute avoidance test is performed as a 48-h (if E. *andrei or* E. fetida*) or 72-h (if* L. terrestris*) sublethal test using a series of circular test units constructed of stainless steel or Plexiglas. Each test unit has a circular central chamber devoid of substrate (where 10 adult worms of the same species are initially placed) with holes leading to each of six pie-shaped, interconnected test compartments. Three of the test compartments contain aliquots of the same sample (or concentration) of test material, and three (in alternating positions) contain aliquots of clean (i.e., negative control or reference) soil. The number in each of the test compartments is determined following a test period of 48 or 72 h for the worms in each test unit to distribute themselves in clean or test soil. This test uses ≥5 replicated test units/treatment if a single-concentration test, and normally one or two test units/concentration if a multi-concentration test.*

The test for effects of prolonged exposure to contaminated or potentially contaminated soil uses laboratory-cultured E. andrei only. *This test is initiated by placing two adult worms in each of a series of 500-mL glass jars (10 replicate jars/treatment) containing a measured wet weight equivalent to ~350 mL of test or clean (negative control or reference) soil. Following a 28-day exposure, the survival rate for the replicate groups of adult worms in each treatment is determined, and they are discarded. The test is continued for an additional 28 days with their progeny. At the end of the 56-day test period, the number of live juvenile worms produced in each replicate and treatment is determined and the treatment means compared. Additionally, the dry weight of individual juvenile worms surviving at test end is determined for each replicate, and the treatment means compared.*

General or universal conditions and procedures are outlined for test preparation and performance. Additional conditions and procedures are stipulated that are specific to the intended use of each test. Each of these three biological test methods is suitable for measuring and assessing the toxicity of samples of field-collected soil, biosolids, sludge, or similar particulate material; or of natural or artificial soil spiked (mixed) in the laboratory with test chemical(s) or chemical product(s). Instructions and requirements are included on test facilities, sample collection, handling and storing samples, culturing and/or holding and acclimating test organisms, preparing soil or spiked-soil mixtures and initiating tests, specific test conditions, appropriate observations and measurements, endpoints and methods of calculation, and the use of a reference toxicant.

Résumé

*Le présent document expose, conseils à l'appui, les modes opératoires détaillés et les conditions suivre pour la préparation et la réalisation de trois méthodes servant à mesurer la toxicité des sols à l'aide de vers de terre (*Eisenia andrei, E. fetida *ou* Lumbricus terrestris*) : (1) essai de « létalité aiguë » (14 jours), avec des vers adultes ; (2) essai rapide (48 ou 72 h) sur le comportement d'évitement du sol contaminé par des vers adultes ; (3) essai (56 j ou, dans certains cas, de 63 j) pour déterminer les effets d'une exposition prolongée de vers adultes (*E. andrei *d'élevage, uniquement) et de leur progéniture à un ou à plusieurs échantillons ou concentrations de sol éventuellement ou effectivement contaminé. Appliquée dans des conditions statiques (sans renouvellement du milieu), chaque méthode emploie un ou plusieurs échantillons de sol effectivement ou éventuellement contaminé ou une ou plusieurs concentrations de substance(s) ou de produit(s) chimique(s) ajoutés à un sol servant de témoin négatif (ou à un autre sol). On ne nourrit les vers que pendant l'essai de 56 j (de flocons d'avoine bouillis).*

L'essai de létalité aiguë (14 j) se déroule dans des bocaux de verre de 500 ml, en présence d'une masse humide mesurée, équivalant à un volume d'environ 350 ml de sol et avec, selon le cas, 5 E. andrei ou fetida *ou 3* L. terrestris*. On compte au moins 5 répétitions par concentration si l'essai porte sur une seule concentration* ***ou*** *emploie* L. terrestris *et au moins 3 répétitions par concentration s'il porte sur plusieurs concentrations* ***et*** *s'il emploie* E. andrei *ou* E. fetida*.*

Effectué en tant qu'essai sublétal, l'essai rapide sur le comportement d'évitement dure 48 h (avec E. andrei *ou* E. fetida*) ou 72 h (avec* L. terrestris*), dans des enceintes circulaires d'acier inoxydable ou de plexiglas, chacune possédant une alvéole centrale circulaire, dépourvue de substrat (où on dépose, au début de l'essai, 10 congénères) et percée d'orifices donnant chacun sur l'un des six compartiments communicants, en forme de pointe de tarte. Trois compartiments renferment des aliquotes du même échantillon (ou de la même concentration) de matière. Ils sont séparés l'un de l'autre par autant de compartiments renfermant des aliquotes de sol propre (c'est-à-dire, de témoin négatif ou de sol de référence). À la fin de l'essai, on détermine le nombre de vers répartis dans les différents compartiments. On emploie au moins 5 enceintes si on utilise une seule concentration et normalement 1 ou 2 enceintes par concentration si les concentrations sont multiples.*

L'essai visant à déterminer les effets de l'exposition prolongée à un sol effectivement ou éventuellement contaminé n'utilise que des E. andrei *élevés en laboratoire. On commence par en placer deux par bocal de verre dans une série de bocaux de 500 ml (10 répétitions par concentration) renfermant une masse humide mesurée, équivalant à environ 350 ml de sol d'essai ou de sol propre (témoin négatif ou sol de référence). Après 28 j, on détermine le taux de survie des groupes exposés à chaque concentration, puis on poursuit l'essai pendant encore 28 j, en ne gardant que la progéniture de ces vers. À la fin des 56 j, on détermine le nombre de jeunes vers produits dans chaque enceinte et à chaque concentration, et on compare les moyennes relatives à chaque concentration. En outre, on détermine le poids sec de chaque jeune ver survivant à l'essai, pour chaque répétition, et on compare les moyennes obtenues pour chaque concentration.*

Pour la préparation et l'exécution des essais, on expose des conditions et des modes opératoires généraux ou universels. S'y ajoutent des conditions ou des modes opératoires propres à la destination de chacun des essais. Chacune des méthodes permet de mesurer et d'évaluer la toxicité : d'échantillons de solides biologiques, de boues ou de sols prélevés sur le terrain ou de matières particulaires semblables ; de sols naturels ou artificiels enrichis, au laboratoire, du ou des produits chimiques ou substances d'essai. Elle est augmentée d'instructions et d'exigences sur les installations expérimentales, le prélèvement, la manipulation et l'entreposage des échantillons, l'élevage et l'acclimatation des vers, la préparation de mélanges de sols ou de sols enrichis et la mise en branle des essais, leurs conditions particulières, les observations et les mesures à faire, les effets mesurés (ou paramètres de mesure), les méthodes de calcul, l'emploi d'un toxique de référence.

Foreword

This is one of a series of ***recommended methods*** *for measuring and assessing the toxic effect(s) on single species of terrestrial or aquatic organisms, caused by their exposure to samples of toxic or potentially toxic substances or materials under controlled and defined laboratory conditions. Recommended methods are those that have been evaluated by Environment Canada (EC), and are favoured:*

- *for use in EC environmental toxicity laboratories;*
- *for testing which is contracted out by Environment Canada or requested from outside agencies or industry;*
- *in the absence of more specific instructions, such as are contained in regulations; and*
- *as a foundation for the provision of very* explicit *instructions as might be required in a regulatory* protocol *or standard* reference method.

The different types of tests included in this series were selected because of their acceptability for the needs of programs for environmental protection and management carried out by Environment Canada. These reports are intended to provide guidance and to facilitate the use of consistent, appropriate, and comprehensive procedures for obtaining data on the toxicity to terrestrial or aquatic life of specific test substances or materials destined for or within the environment. Depending on the biological test method(s) chosen and the environmental compartment of concern, substances or materials to be tested for toxicity could include samples of chemical or chemical product, soil or similar particulate material, sediment or similar particulate material, effluent, elutriate, leachate, or receiving water. Appendix A lists the biological test methods and supporting guidance documents published to date by Environment Canada's Method Development and Applications Section in Ottawa, ON.

Words defined in the Terminology section of this document are italicized when first used in the body of the report according to the definition. Italics are also used as emphasis for these and other words, throughout the report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Chemical Formulae

AES	atomic emission spectrophotometry
ANOVA	analysis of variance
$CaCl_2$	calcium chloride
$CaCO_3$	calcium carbonate
$Ca(OH)_2$	calcium hydroxide
CCME	Canadian Council of Ministers of the Environment
cm	centimetre(s)
CV	coefficient of variation
EC50	median effective concentration
ECx	effective concentration for a (specified) percent effect
ES	effect size
g	gram(s)
h	hour(s)
HCl	hydrochloric acid
HNO_3	nitric acid
HPLC	high pressure liquid chromatography
H_2O	water
H_o	null hypothesis
ICAP	inductively coupled argon plasma
ICp	inhibiting concentration for a (specified) percent effect
KCl	potassium chloride
kg	kilogram(s)
L	litre(s)
LC50	median lethal concentration
LOEC	lowest-observed-effect concentration
LSD	least significant difference
m	metre(s)
M	mole(s) (concentration)
mg	milligram(s)
mL	millilitre(s)
mm	millimetre(s)
mS	millisiemens
MW	molecular weight
n	sample size
nm	nanometre(s)
NOEC	no-observed-effect concentration
OM	organic matter
P	probability
PAHs	polycyclic aromatic hydrocarbons
SD	standard deviation
s	second
sp.	species (singular)
spp.	species (plural)
t	time
TOC	total organic carbon
TM (™)	Trade Mark
v:v	volume-to-volume
WHC	water-holding capacity
wt	weight
wt:wt	weight-to-weight ratio
°C	degree(s) Celsius
α	level of statistical significance (alpha)
µg	microgram(s)
µm	micrometre(s)
µmhos	micromhos
µmol	micromole(s)
>	greater than
<	less than
≥	greater than or equal to
≤	less than or equal to
%	percentage or percent
=	equals
+	plus
–	minus
±	plus or minus
×	times
÷	divided by
/	per; alternatively, “or” (e.g., holding/acclimation)
~	approximately

Terminology

Note: all definitions are given in the context of the procedures in this report, and might not be appropriate in another context.

Grammatical Terms

Must is used to express an absolute requirement.

Should is used to state that the specified condition or procedure is recommended and ought to be met if possible.

May is used to mean "is (are) allowed to".

Can is used to mean "is (are) able to".

Might is used to express the possibility that something could exist or happen.

Technical Terms

Acclimation is physiological adjustment to a particular level of one or more environmental factors such as temperature. The term usually refers to the adjustment to controlled laboratory conditions.

Adult (worm) is an earthworm that is sexually mature and bears an apparent *clitellum*. (See also *clitellum*, *juvenile*, and *sub-adult*.)

Amphimictic refers to reproduction involving the fertilization of an ovum by a sperm.

Anecic refers to species of earthworms that are active in all strata of soil, normally in permanent burrows which open to the soil surface; such species feed on and bury surface litter.

Biomass is the total weight (mass) of a group of animals or plants.

Clitellum is the fleshy "ring" or "saddle" of glandular tissue found on certain mid-body segments of lumbricid earthworms. It is the most visible feature of an adult earthworm, and is prominent only in sexually mature (i.e., adult) individuals. Adolescent or younger worms, which are sexually immature, are distinguished from adults by the absence of a clitellum. The *clitellum* secretes the cocoon into which eggs and sperm are deposited. During copulation, it also exudes mucous which envelopes the anterior ends of the two individuals.

Cocoon is the protective egg case formed by the clitellum of earthworms, from which neonates emerge. (See also *clitellum*.)

Compliance means in accordance with governmental regulations or requirements for issuing a permit.

Conductivity is a numerical expression of the ability of an aqueous solution to carry an electric current. This ability depends on the concentrations of ions in solution, their valence and mobility, and on the solution's temperature. Conductivity is measured at 25 °C, and is reported as micromhos per centimetre (μmhos/cm) or as millisiemens per metre (mS/m); 1 mS/m = 10 μmhos/cm.

Culture, as a noun, means the stock of organisms raised in the laboratory under defined and controlled conditions through one or more generations, to produce healthy test organisms. As a verb, it means to carry out the procedure of raising healthy test organisms from one or more generations, under defined and controlled conditions.

Ecological risk assessment is the process of identifying and quantifying *risks* to nonhuman organisms and determining the acceptability of those risks.

Epigeic refers to litter-dwelling species of earthworms that are active primarily in the detritus and feed primarily on fresh organic material.

Epilobic (prostomium) refers to the type of prostomium in which the tongue of the earthworm partly divides the peristomium. (See also *prostomium* and *peristomium*.)

Genital tumescences refer to areas of modified epidermis (i.e., glandular swellings) on the body of an earthworm without distinct boundaries and through which follicles of genital setae open. (Reynolds, 1977).

Growth is the increase in size or weight as the result of proliferation of new tissues. In this test, it refers to an increase in dry weight.

Hatchling is an earthworm that has recently emerged from a cocoon, and has begun to actively feed and grow. (See also *juvenile*.)

Hormesis is an observed stimulation of performance among organisms, compared to the control organisms, at low concentrations in a toxicity test.

Juvenile (worm) is an earthworm that is sexually immature and lacks an apparent *clitellum*. This classification can include *hatchlings* (e.g., worms within 48 h of emergence from cocoons), as well as all other sexually immature life stages for hatched earthworms up to and including *sub-adults*. (See also *adult, clitellum*, *hatchling*, and *sub-adult*.)

Lumbricid refers to earthworms that are members of the family Lumbricidae, superfamily Lumbricoidea, order Haplotaxida, subclass Oligochaeta, class Clitellata, and phylum Annelida.

Lux is a unit of illumination based on units per square metre. One lux = 0.0929 foot-candles and one foot-candle = 10.76 lux. For conversion of lux to quantal flux [$\mu mol/(m^2 \cdot s)$], the spectral quality of the light source must be known. Light conditions or irradiance are properly described in terms of quantal flux (photon fluence rate) in the photosynthetically effective wavelength range of approximately 400–700 nm. The relationship between quantal flux and lux or foot-candles is highly variable and depends on the light source, the light meter used, the geometrical arrangement, and the possibilities of reflections (see ASTM, 1999a). Approximate conversions between quantal flux and lux, however, are:

- cool-white fluorescent light: 1 lux $\simeq$ 0.014 $\mu mol/(m^2 \cdot s)$;
- full-spectrum fluorescent light (e.g., Vita-Lite® by Duro-Test®): 1 lux $\simeq$ 0.016 $\mu mol/(m^2 \cdot s)$; and
- incandescent light: 1 lux $\simeq$ 0.019 $\mu mol/(m^2 \cdot s)$ (Deitzer, 1994; Sager and McFarlane, 1997).

Monitoring is the routine (e.g., daily, weekly, monthly, quarterly) checking of quality, or collection and reporting of information. In the context of this report, it means either the periodic (routine) checking and measurement of certain biological or soil quality variables, or the collection and testing of samples of soil for toxicity.

Palearctic refers to a zoogeographical region including all of Europe and the former USSR to the Pacific Ocean, Africa north of the Sahara Desert, and Asia north of the Himalaya Mountains.

Papillae refer to protruding dermal structures on the external body of an earthworm. (Reynolds, 1977).

Peristomium refers to the first body segment of an earthworm, which is without setae and contains the mouth. (See also *seta*.)

pH is the negative logarithm of the activity of hydrogen ions in gram equivalents per litre. The pH value expresses the degree or intensity of both acidic and alkaline reactions on a scale from 0 to 14, with 7 representing neutrality, numbers less than 7 indicating increasingly greater acidic reactions, and numbers greater than 7 indicating increasingly basic or alkaline reactions.

Photoperiod is the duration of illumination and darkness within a 24-h day.

Pollution is the addition of a substance or material, or a form of energy such as heat, to some component of the environment, in such an amount as to cause a discernible change that is deleterious to some organism(s) or to some human use of the environment. Some national and international agencies have formal definitions of pollution, which should be honoured in the appropriate contexts.

Pretreatment means treatment of a sample of soil, or portion thereof, before exposure of the test organisms.

Progeny means the young or offspring (i.e., immediate descendants) of a sexually mature (adult) earthworm.

Prostomium refers to the anterior lobe projecting in front of the peristomium and above the mouth. (See also *peristomium*.)

Protocol is an explicit set of procedures for a test, formally agreed upon by the parties involved, and described precisely in a written document.

Quality assurance (QA) is a program within a laboratory, intended to provide precise and accurate results in scientific and technical work. It includes selection of proper procedures, sample collection, selection of limits, evaluation of data, *quality control*, and qualifications and training of personnel.

Quality control (QC) consists of specific actions within the program of quality assurance. It includes standardization, calibration, replication, control samples, and statistical estimates of limits for the data.

Rearing means collecting test organisms from the wild or from a commercial supplier, and maintaining them in the laboratory until they are acclimated to test conditions and are healthy animals of the right age and/or size for toxicity testing.

Reference method refers to a specific protocol for performing a toxicity test, i.e., a biological test method with an explicit set of test procedures and conditions, formally agreed upon by the parties involved and described precisely in a written document. Unlike other multi-purpose (generic) biological test methods published by Environment Canada, the use of a *reference method* is frequently restricted to testing requirements associated with specific regulations.

Remediation is the management of a contaminated site to prevent, minimize, or mitigate damage to human health or the environment. *Remediation* can include both direct physical actions (e.g., removal, destruction, and containment of toxic substances) and institutional controls (e.g., zoning designations or orders).

Risk is the probability or likelihood that an adverse effect will occur.

Risk assessment – see *Ecological risk assessment*.

Seminal vesicles refer to the storage sacs for an earthworm's own sperm, until copulation.

Septa refer here to internal partitions at intersegmental furrows along the outer body of an earthworm. Septa also act as supporting membranes for internal organs. (Reynolds, 1977).

Seta refers to a solid rod or bristle secreted by cells at the internal end of a tubular epidermal ingrowth referred to as the setal follicle. Setae are of various types (e.g., general, genital, or penial; see Reynolds 1977 for further description).

Spermathecae refer to the pouches developed in the septa of an earthworm, which receive sperm from another individual during copulation; the sperm are stored here until the period of cocoon laying (Reynolds, 1977). See also *septa*.

Sub-adult (worm) is a *juvenile* "adolescent" earthworm that is sexually immature and lacks an apparent *clitellum*. For an acute lethality test, which can begin with either sub-adult or *adult* worms, the wet weight of any sub-adult earthworm used to start the test must range within 250–600 mg if *E.* andrei or *E. fetida*, or range within 3–10 g if *L. terrestris*. (See also *adult, clitellum*, and *juvenile*.)

Tanylobic (prostomium) refers to the type of prostomium in which the tongue of the earthworm partly divides the peristomium. (See also *prostomium* and *peristomium*.)

Tubercula pubertatis refers to a glandular swelling on an earthworm, which appears near the ventrolateral margins of the clitellum. It is not always present, and it might be continuous or discontinuous and of varied size and shape (Reynolds, 1977).

Terms for Test Materials or Substances

Artificial soil is a laboratory-formulated soil, prepared to simulate a natural soil using a specific ratio of natural constituents of sand, clay, and peat. *Artificial soil* may be used as a *negative control soil*, and as a diluent to prepare multiple concentrations of *site soil(s)* or *chemical-spiked soil(s)*.

Batch means the total amount of a particular *test soil* (or specific concentration thereof) prepared for each treatment (concentration) in a test. A *batch* is any hydrated *test soil* ready for separation into replicates.

Chemical is, in this report, any element, compound, formulation, or mixture of a substance that might be mixed with, deposited in, or found in association with soil or water.

Chemical-spiked soil is natural or artificial soil (usually *negative control soil, reference soil*, or other *clean* soil) to which one or more chemicals or chemical products have been added, and mixed thoroughly to evenly distribute the substance(s) throughout the soil at a specific concentration to form a *batch* for use in a soil toxicity test. See also *spiked soil*.

Clean soil is soil that does not contain concentrations of any substance(s) or material(s) causing discernible toxic effects on the test organisms.

Concentration means, for these biological test methods, the ratio of the weight of test substance of material to the weight of soil, and is frequently expressed as the weight of test substance or material per kg of dry soil (mg/kg). *Concentration* might also be expressed as a percentage of the test substance or material (e.g., contaminated site soil) per dry weight of soil.

Contaminant is a substance or material that is present in a natural system, or present at increased concentrations, often because of some direct or indirect human activity. The term is frequently applied to substances or materials that are present at concentrations that have the potential to cause adverse biological effects.

Contaminated (soil) means (soil) containing chemical substances or materials at concentrations that pose a known or potential threat to environmental or human health.

Control is a treatment in an investigation or study that duplicates all the conditions and factors that might affect results, except the specific condition being studied. In toxicity tests, the control must duplicate all the conditions of the exposure treatment(s), but must contain no contaminated test material. The *control* is used as a check for the absence of toxicity due to basic test conditions such as temperature, health of test organisms, or effects due to their handling. *Control* is synonymous with *negative control*, unless indicated otherwise.

Control soil – see *negative control soil*.

Definitive (soil toxicity test) means decisive (as opposed to a preliminary, range-finding test). See also *range-finding (test)*.

De-ionized water is water that has been purified by passing it through resin columns or a reverse osmosis system, for the purpose of removing ions such as Ca^{++} and Mg^{++}.

Distilled water is water that has been passed through a distillation apparatus of borosilicate glass or other material, to remove impurities.

Fertility (of soil) refers to the potential of a soil to supply nutrient elements in the amounts, forms, and proportions required for optimal plant growth. Soil fertility is measured directly in terms of the ions and compounds important for plant nutrition. The fundamental components of fertility are the essential nutrients (macronutrients including C, H, O, N, P, K, Ca, Mg, S and micronutrients including Fe, Mn, Mo, B, Cu, Zn, and Cl). Indirectly, soil fertility is measured by demonstrating its productivity (i.e., the capacity of the soil to produce plants that supply man with essential food and fibre; Hausenbuiller, 1985).

Hydration water means water used to hydrate test soils, to create a specific moisture content suitable for the test organisms. The water used for hydration is normally *test water*, and is frequently de-ionized or distilled water. Depending on study design and intent, a surface water or groundwater from the *site* might be used instead of de-ionized or distilled water for the hydration of each test soil (including negative control soil). (See also *test water, de-ionized water,* and *distilled water*.)

Material is the *substance* or substances from which something is made. A *material* would have more or less uniform characteristics. Soil, sediment, or surface water are materials. Usually, the *material* would contain several or many substances.

Moisture content is the percentage of water in a sample of test soil, based on its wet or dry mass. It is determined by measuring both the wet and dry weights of a subsample of the soil. The soil's moisture content is then calculated and expressed on a dry-weight basis, by dividing the mass of water in the subsample (wet mass–dry mass) by the mass of dry soil, and then multiplying by 100. Units for mass (i.e., g or mg) must be the same in each instance.

Negative control (see *control*).

Negative control soil is *clean* soil that does not contain concentrations of one or more contaminants which could affect the survival, reproduction, growth, or behaviour of the test organisms. *Negative control soil* might be natural soil from an uncontaminated site, or artificial (formulated) soil. This soil must contain no added test material or

substance, and must enable acceptable survival and performance of the test organisms during the test. The use of *negative control soil* provides a basis for interpreting data derived from toxicity tests using test soil(s).

Organic matter (OM) in soil consists primarily of plant and animal residues, at different stages of decomposition, including soil humus. The accumulation of OM within soil is a balance between the return or addition of plant and animal residues and their subsequent loss due to the decay of these residues by soil micro-organisms. For most types of soil, the following equation (AESA, 2001) is suitable for estimating the total OM content of soil from *total organic carbon* (TOC) measurements: % OM = % TOC × 1.78. (See also *total organic carbon.*)

Positive control soil is *contaminated* soil that contains concentrations of one or more contaminants that adversely affect the survival, reproduction, growth, or behaviour of the test organisms using the biological test methods defined herein. *Positive control soil* might be used as a reference toxicant to assess the sensitivity of the test organisms at the time the test material or substance is evaluated, and to determine the precision of results obtained by the laboratory for that reference toxicant.

Product is a commercial formulation of one or more chemicals. (See also *chemical*.)

Range-finding (test) means a preliminary soil toxicity test, performed to provide an initial indication of the toxicity of the test material under defined conditions and to assist in choosing the range of concentrations to be used in a definitive multi-concentration test. (See also *definitive (soil toxicity test)*.)

Reference soil is typically *clean* field-collected soil or formulated (artificial) soil, that is selected for use in a particular toxicity test together with a *negative control soil* and one or more samples of *test soil*. The test soil might be either field-collected site soil that is contaminated or potentially so, or chemical-spiked soil. Reference soil used in a test frequently exhibits physicochemical properties (e.g., texture, compactness, total organic carbon content, pH) closely matching those of the test soil sample(s), except that it is free from the source of contamination being assessed. In tests involving samples of site soil, one or more samples of reference soil are often selected from the general location of test soil sampling, and thus might be subject to other sources of contamination aside from the one(s) being studied. Reference soil is used to describe matrix effects in the test, and may also be used as a diluent to prepare concentrations of the test soil. In tests involving chemical-spiked soil, one or more samples of artificial (formulated) soil with differing physicochemical characteristics might be chosen to investigate the influence of certain soil properties (e.g., soil texture, or percent organic matter) on the toxicity of a chemical mixed in each of these soil types. (See also *negative control soil*, *site soil*, *test soil*, *clean*, *artificial soil*, and *chemical-spiked soil.*)

Reference toxicant is a standard chemical used to measure the sensitivity of the test organisms to establish confidence in the toxicity data obtained for a test material or substance. In most instances, a toxicity test with a reference toxicant is performed to assess the sensitivity of the organisms at the time the test material or substance is evaluated, and the precision and reliability of results obtained by the laboratory for that chemical.

Reference toxicity test is a test conducted using a reference toxicant in conjunction with a soil toxicity test, to appraise the sensitivity of the organisms and the precision and reliability of results obtained by the laboratory for that chemical at the time the test material or substance is evaluated. Deviations outside an established normal range indicate that the sensitivity of the test organisms, and the performance and precision of the test, are suspect. A reference toxicity test with earthworms is performed as a *spiked-soil* test, using a standard chemical.

Sampling station means a specific location, within a *site* or sampling unit (depending on the study design), where the sample(s) of field-collected soil are obtained for toxicity tests and associated physicochemical analyses.

Site means a delineated tract of land that is being used or considered as a study area, usually from the perspective of it being contaminated or potentially contaminated by xenobiotics.

Site soil is a field-collected sample of soil, taken from a location thought to be contaminated with one or more chemicals, and intended for use in the toxicity test with earthworms. In some instances, the term includes *reference soil* or *negative control soil* from a site.

Soil is whole, intact material representative of the terrestrial environment, that has had minimal manipulation following collection or formulation. In the natural environment, it is formed by the physical, chemical, and biological weathering of rocks and the decomposition and recycling of nutrients from organic matter originating from plant and animal life. Its physicochemical characteristics are influenced by microbial and invertebrate (including earthworm) activities therein, and by anthropogenic activities.

Solvent control soil is a sample of (usually artificial) soil included in a test involving chemical-spiked soil, in which an organic solvent is required to solubilize the test chemical before mixing it in a measured quantity of *negative control soil*. The amount of solvent used when preparing the solvent control soil must contain the same concentration of solubilizing agent as that present in the highest concentration of the test chemical(s) in the sample of chemical-spiked soil to be tested. This concentration of solvent should not adversely affect the performance of earthworms during the test. Any test that uses an organic solvent when preparing one or more concentrations of chemical-spiked soil must include a *solvent control soil* in the test. (See also *artificial soil*, *negative control soil*, and *chemical-spiked soil*.)

Spiked soil is natural or artificial soil (usually *negative control soil, reference soil*, or other *clean* soil) to which one or more chemicals, chemical products, or other test substances or materials (e.g., a sample of sludge or drilling mud) have been added in the laboratory, and mixed thoroughly to evenly distribute the substance(s) or material(s) throughout the soil at a specific concentration to form a *batch* for use in a soil toxicity test. (See also *chemical-spiked soil* and *spiking*.)

Spiking refers to the addition of a known amount of chemical(s), chemical product(s), or other test substance(s) or material(s) (e.g., a sample of sludge or drilling mud) to a natural or artificial soil. The substance(s) or material(s) is usually added to *negative control soil, reference soil*, or another *clean* soil, but sometimes to a contaminated or potentially contaminated soil. After the addition ("spiking"), the soil is mixed thoroughly. If the added test material is a *site soil*, Environment Canada documents typically do not call this spiking, but instead refer to the manipulation as "dilution" or simply "addition". (See also *chemical-spiked soil* and *spiked soil*.)

Stock solution means a concentrated solution of the substance(s) to be tested, followed by the addition of a measured quantity of this solution to a sample of natural or artificial soil and thorough mixing to prepare a *batch* of *chemical-spiked soil*. To prepare the required strength of the *stock solution*, measured weights or volumes of test chemical(s) or chemical product(s) are added to test water (*de-ionized* or *distilled water, or equivalent*), with or without the inclusion of an organic solvent.

Substance is a particular kind of material having more or less uniform properties. The word *substance* has a narrower scope than *material*, and might refer to a particular chemical (e.g., an element) or chemical product.

Test soil is a sample of field-collected soil or *chemical-spiked soil* to be evaluated for toxicity to earthworms. In some instances, the term also applies to any solid-phase sample or mixture thereof (e.g., negative control soil, positive control soil, reference soil, sludge, drilling mud) used in a soil toxicity test.

Test water is water used to prepare stock solutions, rinse test organisms, or rinse glassware and other apparatus used for culturing or holding and acclimating earthworms and for other purposes associated with the biological test method (e.g., to hydrate samples of test soil). *Test water* must be de-ionized or distilled water or better (e.g., reagent-grade water produced by a system of reverse osmosis, carbon, and ion-exchange cartridges). (See also *hydration water*.)

Texture is defined based on a measurement of the percentage by weight of sand, silt, and clay in the mineral fraction of soils. Classification as to texture confers information on the general character and behaviour of substances in soils,

especially when coupled with information on the structural state and organic matter content of the soil. Soil texture is determined in the laboratory by measuring the particle-size distribution using a two-step procedure whereby the sand particles (coarse fragments) are initially separated by sieving from the silt and clay particles; followed by separation of the silt and clay particles by their sedimentation in water. Textural classification systems typically refer to groupings of soil based on specific ranges in relative quantities of sand, silt, and clay. There are three main textural classes:

(i) coarse texture (sands, loamy sands, sandy loams);
(ii) medium texture (loams, silt loams, silts, very fine sandy loams); and
(iii) fine texture (clays, silty clay loams, sandy clay loams, silty clays, and sandy clays).

Further distinction as to texture (e.g., "sandy clay", "silt loam", "loam") can be made based on classification schemes using the relative amounts of percent sand, percent silt, and percent clay in the soil (Hausenbuiller, 1985; ACECSS, 1987).

Total organic carbon (TOC) refers to the organic carbon content of soil exclusive of carbon from undecayed plant and animal residues. The TOC is determined by dry combustion analysis (ISO, 1995). (See also *organic matter*.)

Water-holding capacity (WHC) refers to the maximum quantity of water that a soil can retain, following complete saturation. It is usually determined gravimetrically, and is generally expressed as the percentage of water (by mass; wt water:wt dry soil) retained in a sample of soil that has been saturated with water.

Statistical and Toxicological Terms

Acute means within a short period of exposure (seconds, minutes, hours, or a few days) in relation to the life span of the test organism.

Acute toxicity is a discernible adverse effect (lethal or sublethal) induced in the test organisms within a short period (usually a few days, and for purposes of this document within 14 days) of exposure to *test soil(s)*.

Battery of toxicity tests is a combination of several toxicity tests, normally using different species of test organisms (e.g., a series of soil toxicity tests using earthworms, plants, or springtails), different biological endpoints (e.g., lethal and various sublethal), and different durations of exposure (e.g., acute and chronic).

Bioassay is a test (= assay) in which the strength or potency of a substance is measured by the response of living organisms. In standard pharmacological usage, a *bioassay* assesses the unknown potency of a given preparation of a drug, compared to the known potency of a standard preparation. *Toxicity test* is a more specific and preferred term for environmental studies.

Chronic means occurring within a relatively long period of exposure (weeks, months, or years), usually a significant portion of the life span of the organism such as 10% or more.

Chronic toxicity implies adverse effects during or after relatively long-term exposures to one or more contaminants, that are related to changes in reproduction, growth, metabolism, ability to survive, or other biological variables (e.g., behaviour) being observed.

Coefficient of Variation (CV) is the standard deviation (SD) of a set of data divided by the mean of the data set, expressed as a percentage. It is calculated according to the following formula: CV (%) = 100 (SD ÷ mean).

EC50 is the *median effective concentration.* That is the concentration (e.g., % or mg/kg) of substance(s) or material(s) in soil that is estimated to cause some defined toxic effect on 50% of the test organisms. In most instances, the EC50 and its 95% confidence limits are statistically derived by analyzing the percentages of organisms affected

(e.g., showing an avoidance response) at various test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 48 h or 72 h). The EC50 describes *quantal* effects, lethal or sublethal, and is not applicable to continuous (i.e., *quantitative*) effects (see *ICp*). Depending on the study objectives, an ECx other than EC50 (e.g., an EC20) might be calculated instead of or in addition to the EC50.

Endpoint means the measurement(s) or value(s) that characterize the results of a test (e.g., LC50, EC50, IC25). It also means the response of the test organisms that is measured (e.g., death, or number of progeny produced).

Environmental toxicology is a branch of *toxicology* with the same general definition; however, the focus is on ecosystems, natural communities, and wild living species, without excluding humans as part of the ecosystems.

Geometric mean is the mean of repeated measurements, calculated on a logarithmic basis. It has the advantage that extreme values do not have as great an influence on the mean as is the case for an arithmetic mean. The *geometric mean* can be calculated as the n^{th} root of the product of the "n" values, and it can also be calculated as the antilogarithm of the mean of the logarithms of the "n" values.

Heteroscedasticity refers herein to data showing heterogeneity of the residuals within a scatter plot (see Figures J.2B and J.2C in Appendix J). This term applies when the variability of the residuals changes significantly with that of the independent variables (i.e., the test concentrations or treatment levels). When performing statistical analyses and assessing residuals (e.g., using Levine's test), for test data demonstrating heteroscedasticity (i.e., non-homogeneity of residuals), there is a significant difference in the variance of residuals across concentrations or treatment levels. (See also *homoscedasticity* and *residual*.)

Homoscedasticity refers herein to data showing homogeneity of the residuals within a scatter plot (see Figure J.2A in Appendix J). This term applies when the variability of the residuals does not change significantly with that of the independent variables (i.e., the test concentrations or treatment levels). When performing statistical analyses and assessing residuals (e.g., using Levine's test), for test data demonstrating homoscedasticity (i.e., homogeneity of residuals), there is no significant difference in the variance of residuals across concentrations or treatment levels. (See also *heteroscedasticity* and *residual*.)

ICp is the *inhibiting concentration for a (specified) percent effect*. It represents a point estimate of the concentration of test substance or material that causes a designated percent inhibition (*p*) compared to the control, in a *quantitative* (continuous) biological measurement such as number of progeny produced or size attained by individuals at the end of the test.

LC50 is the *median lethal concentration*, i.e., the concentration (e.g., % or mg/kg) of substance or material in soil that is estimated to be lethal to 50% of the test organisms. The LC50 and its 95% confidence limits are usually derived by statistical analysis of percent mortalities in five or more test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 7-day LC50 or 14-day LC50). Depending on the study objectives, an LCx other than LC50 (e.g., an LC20) might be calculated instead of or in addition to the LC50.

Lethal means causing death by direct action. Death of test organisms is defined as the cessation of all visible signs of movement or other activity indicating life.

LOEC is the *lowest-observed-effect concentration*. This is the lowest concentration of a test substance or material for which a statistically significant adverse effect on the test organisms was observed, relative to the control.

NOEC is the *no-observed-effect concentration*. This is the highest concentration of a test substance or material at which no statistically significant adverse effect on the test organisms was observed, relative to the control.

Precision refers to the closeness of repeated measurements of the same quantity to each other, i.e., the degree to which data generated from replicate measurements are the same. It describes the degree of certainty around a result, or the tightness of a statistically derived endpoint such as an ICp.

Quantal effects in a toxicity test are those in which each test organism responds or does not respond. For example, an animal might respond by dying in or avoiding a contaminated test soil. Generally, quantal effects are expressed as numerical counts or percentages thereof. (See also *quantitative*.)

Quantitative effects in a toxicity test are those in which the measured effect is continuously variable on a numerical scale. Examples would be number of progeny produced, or dry weight of progeny at test end. Generally, quantitative effects are determined and expressed as measurements. (See also *quantal*.)

Replicate (*treatment*, *test chamber*, or *test unit*) refers to a single test chamber containing a prescribed number of organisms in either one concentration of the test material or substance, or in the control or reference treatment(s). A *replicate* of a treatment must be an independent test unit; therefore, any transfer of organisms or test material from one test chamber to another would invalidate a statistical analysis based on the replication.

Replicate samples are field-replicated samples of soil collected from the same sampling station, to provide an estimate of the sampling error or to improve the precision of estimation. A single soil sample from a sampling station is treated as one replicate. Additional samples are considered to be additional *replicate samples* when they are treated identically but stored in separate sample containers (i.e., not composited).

Residual, in the context of Section 6.4.2.1 and Appendix J, refers to the difference between the predicted estimate (based on the model) and the actual value observed, as determined by subtracting the former from the latter. (See also *heteroscedasticity* and *homoscedasticity*.)

Static describes a toxicity test in which the *test soil* (nor any chemical or chemical product therein) is not renewed or replaced during the test.

Sublethal (toxicity) means detrimental to the organism, but below the concentration or level of contamination that directly causes death within the test period.

Sublethal effect is an adverse effect on an organism, below the concentration or level of contamination that directly causes death within the test period.

Toxic means poisonous. A toxic chemical or material can cause adverse effects on living organisms, if present in sufficient amount at the right location. *Toxic* is an adjective or adverb, and should not be used as a noun; whereas *toxicant* is a legitimate noun.

Toxicant is a toxic substance or material.

Toxicity is the inherent potential or capacity of a substance or material to cause adverse effect(s) on living organisms. These effect(s) could be lethal or sublethal.

Toxicity test is a determination of the effect of a substance or material on a group of selected organisms (e.g., *Lumbricus terrestris* or *Eisenia fetida*), under defined conditions. A toxicity test involving samples of *test soil* usually measures (a) the proportions of organisms affected (*quantal*), and/or (b) the degree of effect shown (*quantitative* or graded), after exposure of the test organisms to the whole sample (e.g., undiluted site soil) or specific concentrations thereof.

Toxicology is a branch of science that studies the toxicity of substances, materials, or conditions. There is no limitation on the use of various scientific disciplines, field or laboratory tools, or studies at various levels of

organization, whether molecular, single species, populations, or communities. Applied toxicology would normally have a goal of defining the limits of safety of chemical or other agents. See also *environmental toxicology*.

Treatment refers to a specific *test soil* (e.g., a *site soil*, *reference soil*, or *negative control soil*) from a particular sampling station, or a concentration of chemical-spiked soil (or a mixture of test soil diluted with *clean* soil) prepared in the laboratory. Test soils representing a particular *treatment* are typically replicated in a toxicity test. (See also *replicate* and *replicate samples*.)

Warning chart is a graph used to follow changes over time in the endpoints for a reference toxicant. Date of the test is on the horizontal axis and the effect-concentration is plotted on the vertical logarithmic scale.

Warning limit is plus or minus two standard deviations, calculated logarithmically, from a historic geometric mean of the endpoints from tests with a reference toxicant.

Acknowledgements

This document was co-authored by D.J. McLeay (McLeay Environmental Ltd., Victoria, BC) and G.L. Stephenson (Aquaterra Environmental, Orton, ON). J.A. Miller (Miller Environmental Sciences Inc., Innisfil, ON) and J.I. Princz (Biological Methods Division, Environmental Technology Centre, EC, Ottawa, ON) are sincerely thanked for their contributions to certain sections of the document. Input from J.B. Sprague (Sprague Associates Ltd., Salt Spring Island, BC) with respect to certain definitions in the Terminology section and the statistical guidance herein is gratefully acknowledged. B.A. Zajdlik (Zajdlik & Associates, Rockwood, ON) is thanked for his input to the statistical guidance on regression analyses. N.C. Feisthauer (Stantec Consulting Ltd., Guelph, ON) and J. McCann (University of Waterloo, Waterloo, ON) contributed much useful technical guidance and specific advice included herein.

R.P. Scroggins (Biological Methods Division, Environmental Technology Centre, EC, Ottawa , ON) acted as Scientific Authority and provided technical input and direction throughout the work. The studies resulting in the biological test methods defined herein were developed by G.L. Stephenson, as part of her doctoral thesis. Additional studies were directed by Ms. Stephenson, and managed and performed by N.C. Feisthauer and co-workers with Stantec Consulting Ltd. (formerly ESG International Inc.) (Guelph, ON). Laboratory assistance was also provided by many undergraduate and graduate students (University of Guelph). R. Pandey of Guelph Chemical Laboratories Ltd. (Guelph, ON) and V. Marsielle-Kerslake of Analytical Services (University of Guelph, Guelph, ON) are thanked for providing assistance on methods for physicochemical analyses.

The interlaboratory studies undertaken to validate the acute lethality test method described herein were coordinated by J. Princz (Biological Methods Division, Environment Canada, Ottawa, ON) and performed by the following participating laboratories: Environment Canada's Pacific Environmental Science Centre (North Vancouver, BC), Environment Canada's Atlantic Environment Science Centre (Moncton, NB), Environment Canada's Soil Toxicology Laboratory at the Environmental Technology Centre (Ottawa, ON), NRC Biotechnology Research Institute (Montreal, PQ), Centre d'Expertise en Analyse Envrionnementale du Québec (Sainte-Foy, PQ), Ontario Ministry of the Environment (Etobicoke, ON), Saskatchewan Research Council (Saskatoon, SK), BC Research Inc. (Vancouver, BC), EVS Consultants Ltd. (North Vancouver, BC), HydroQual Laboratories Ltd. (Calgary, AB), Stantec Consulting Ltd. (formerly ESG International Inc.) (Guelph, ON), Pollutech EnviroQuatics Ltd. (Point Edward, ON), and Bodycôte Essais Matériaux Canada Inc. (Sainte-Foy, PQ). The contribution of Buchanan Environment Ltd. (Fredericton, NB) to discussions leading to these method-validation studies is also acknowledged with thanks. Names of laboratory personnel participating in these tests are provided in the technical report on these studies (EC, 2004a).

The interlaboratory studies performed to validate the method for measuring avoidance described herein were coordinated by J. Princz (Biological Methods Division, Environment Canada, Ottawa, ON) and conducted by the following laboratories: Environment Canada's Atlantic Environmental Science Centre (Moncton, NB), Environment Canada's Soil Toxicology Laboratory at the Environmental Technology Centre (Ottawa, ON), Pollutec EnviroQuatics Ltd. (Point Edward, ON), and Stantec Consulting Ltd. (Guelph, ON). Names of laboratory personnel participating in this series of test are included in the technical report on these studies (EC, 2004a).

We gratefully acknowledge the many useful comments provided by each member of Environment Canada's Scientific Advisory Group (SAG) responsible for the document's initial and final reviews, as follows: C. Bastien (Centre d'expertise en analyse environnementale du Québec, Ministère de l'Environnement, Saint-Foy, PQ), C. Edwards (Department of Entomology, Ohio State University, Columbus, OH), R. Kuperman (Geo-Centers, Inc., Aberdeen Proving Ground, MD), R. Lanno (Department of Entomology, Ohio State University, Columbus, OH), F. Riepert (Biologische Bundesantalt fur Land- und Forstwirtschaft, Berlin, Germany), J. Römbke (ECT Oekotoxikologie GmbH, Flörsheim am Main, Germany), G. Sunahara (Biotechnology Research Institute, National Research Council, Montreal, PQ), G. van Aggelen (Pacific Environmental Science Centre, EC, North Vancouver, BC), C.A.M. van Gestel (Institute of Ecological Science, Vrije Universiteit, Amsterdam), and S. Visser (Department of Biological Sciences, University of Calgary, Calgary, AB). Contact information for each SAG member is provided in Appendix

D. Support and input from various members of the Inter-Governmental Environmental Toxicity Group (see Appendix B) is also acknowledged.

In addition to the SAG members who reviewed this biological test methods document, the following people also reviewed the initial and/or final drafts and provided useful comments: K. Becker-van Slooten (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland); S. Campiche (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland); W. Diehl (Department of Biological Sciences, Mississippi State University, MS); J. Filser (Zentrum für Umweltforschung und Umwelttechnologie, University of Bremen, Bremen, Germany); P. Hankard (CEH Monks Wood, Abbots Ripton, Huntington, UK); J. Hatcher (HydroQual Laboratories Ltd., Calgary, Alberta); M. Hughes (University of Northern BC, Prince George, BC); K. Hund-Rinke (Fraunhofer Institut fuer Molekularbiologie und Angewandte Oekologie, Fraunhofer-IME, Schmallenberg, Germany); L. Kapustka (Ecological Planning and Toxicology Inc., Corvallis, OR); G. Linder (United States Geological Survey Field Station, Brooks, OR), J. McCann (Department of Biology, University of Waterloo, Waterloo, ON); J. Miller (Miller Environmental Sciences Inc., Innisfil, ON), J. Princz (Biological Methods Division, Environment Canada, Ottawa, ON); J.M.L. Rodrigues (Departmento de Biologia, Universidade de Aveiro, Portugal); M. Rutherford (University of Northern BC, Prince George, BC); M. Schaefer (Zentrum für Umweltforschung und Umwelttechnologie, University of Bremen, Bremen, Germany); J. Scott-Fordsmand (Department of Terrestrial Ecology, National Environmental Research Institute, Silkeborg, Denmark; J.P. Sousa (Instituto do Ambiente e Vida, Universidade de Coimbra, Coimbra, Portugal); J.B. Sprague (Sprague Associates Ltd., Salt Spring Island, BC); B.-J. Unis (HydroQual Laboratories Ltd., Calgary, AB); and M. Warne (Environment Protection Authority, Sydney, NSW, Australia).

Introduction

1.1 Background

The *Method Development and Applications Section* (*MDAS*) of Environment Canada is responsible for the development, standardization, and publication (see Appendix A) of a series of biological test methods for measuring and assessing the toxic effect(s) on single species of terrestrial or aquatic organisms, caused by their exposure to samples of test *materials* or *substances* under controlled and defined laboratory conditions. In 1994, MDAS commenced a multi-year program to research, develop, validate, and publish a number of standardized biological test methods for measuring the *toxicity* of samples of contaminated or potentially contaminated *soil*, using appropriate species of terrestrial test organisms. At the outset, it was decided by Environment Canada that the new biological test methods must be applicable to soil conditions typically found in Canadian environments, and that the selected test organisms must be representative of the species of terrestrial invertebrates and plants inhabiting soil ecosystems in Canada. The initial phase of this multi-year program involved a comprehensive review of existing biological test methods used in North America and elsewhere to evaluate the toxicity of contaminated soils to plants and soil invertebrates. The resulting report recommended that Environment Canada support the development, standardization, and publication of a number of single-species biological test methods for measuring soil toxicity including those using earthworms (Bonnell Environmental Consulting, 1994). This recommendation was endorsed by both the Inter-Governmental Environmental Toxicity Group (IGETG) (Appendix B) and the headquarters and regional offices of Environment Canada (Appendix C).

Since 1994, a series of laboratory studies have been completed under the direction of the MDAS, on the selection of suitable and sensitive test organisms for measuring soil toxicity to meet Canadian regulatory and *monitoring* requirements, and on the development of appropriate biological test methods. A technical report was produced describing the efforts and findings associated with the development of earthworm *toxicity tests* for the assessment of contaminated soils (Aquaterra Environmental, 1998a). Other technical reports generated from these laboratory studies describe a test for mortality and reproductive inhibition of a small soil-dwelling arthropod (springtail species; Aquaterra Environmental, 1998b) and multi-species tests for adverse effects of contaminated soil on the survival, development, and growth of terrestrial plants (Aquaterra Environmental, 1998c).

A number of soil toxicity tests have been coordinated or undertaken by Environment Canada, using various species of earthworms exposed to samples of soil contaminated with pesticides, metals, petrochemical wastes, or prospective *reference toxicants*. These studies (ESP, 1992; Aquaterra Environmental, 1998a; Stephenson *et al.*, 1997, 1998, 1999a, b, 2000a; EC, 2000a, b; Aquaterra Environmental and ESG, 2000; ESG, 2001, 2002; ESG and Aquaterra Environmental, 2002; Stantec and Aquaterra Environmental, 2004) focussed on the development and standardization of biological test methods for determining the *lethal* or *sublethal* toxicity of samples of contaminated soil to earthworms. Based on the results of these studies, Environment Canada proceeded with the preparation of the following three biological test methods for measuring soil toxicity:

- an *acute* (14-day) lethality test, using *Eisenia* sp. (i.e., *E. andrei* or *E. fetida*) or *Lumbricus terrestris*;
- an acute (i.e., 48 h if *Eisenia* spp.; 72 h if *L. terrestris*) sublethal test of avoidance responses, using *E. andrei*, *E. fetida*, or *L. terrestris*; and
- a test for the effects of more prolonged (8-week) exposure of laboratory-cultured *Eisenia andrei* on their survival, reproduction, and *growth*.

A Scientific Advisory Group (see Appendix D) of international experts experienced with the design and implementation of soil toxicity tests using earthworms provided key references which were reviewed and considered as part of this undertaking. These individuals also served actively in providing a critical peer review of two drafts of this methodology document. A larger group of knowledgeable persons

(see Acknowledgements) provided further review comments in response to the final draft preceding this publication. The experience of the international scientific community when performing similar soil toxicity tests using earthworms (see Appendices E, F, and G) was relied on heavily when preparing these three biological test methods.

Detailed procedures and conditions for preparing and performing each of these three biological test methods are defined herein. Universal procedures for preparing and conducting soil toxicity tests using a selected species of earthworm (i.e., *E. andrei, E. fetida,* or *L. terrestris*) are described. Also presented are specific conditions and procedures which are required or recommended when using any of these three biological test methods for evaluating different types of *substances* or *materials* (e.g., samples of field-collected soil or similar particulate waste, or samples of one or more *chemicals* or chemical *products* experimentally mixed into or placed in contact with natural or formulated soil).

The flowchart in Figure 1 illustrates the universal topics covered herein, and lists topics specific to testing samples of field-collected soil, similar particulate waste (e.g., sludge, drilling mud, or dredged material), or soil spiked experimentally with chemical(s) or chemical product(s).

These biological test methods are intended for use in evaluating the lethal and sublethal toxicity of samples of material such as:

(1) field-collected soil that is contaminated or potentially contaminated;

(2) soils under consideration for removal and disposal or *remediation* treatment;

(3) dredged material destined or under consideration for land disposal after dewatering;

(4) industrial or municipal sludge and similar particulate wastes that might be deposited on land; and (5) *clean* or contaminated soil (natural or *artificial*), spiked with one or more chemicals or chemical products (e.g., for risk assessment of new or current-use chemicals).

In formulating these biological test methods, an attempt has been made to balance scientific, practical, and cost considerations, and to ensure that the results will be sufficiently precise for most situations in which they will be applied. It is assumed that the user has a certain degree of familiarity with soil toxicity tests. Explicit instructions that might be required in a regulatory *protocol* are not provided in this report, although it is intended as a guidance document useful for that and other applications.

For guidance on the implementation of these and other biological test methods, and on the interpretation and application of *endpoint* data for soil toxicity, the reader should consult Sections 4.1.2, 5.5, and 5.6.4 in EC (1999).

1.2 Identification, Distribution, and Life History of Eisenia *spp. and* L. terrestris

The species of earthworms to be used for one or more of the biological test methods described herein (i.e., *Eisenia andrei*, *Eisenia fetida*, or *Lumbricus terrestris)* belong to the Lumbricidae family (phylum, Annelida; class, Clitellata; subclass, Oligochaeta; order, Haplotaxida; superfamily, Lumbricoidea; family, Lumbricidae). The lumbricids are not native to Canadian soils, and were most likely introduced from Europe by early settlers (Bonnell, 1994; Fox, 2000). Definitive information regarding the identification, distribution, biology, and life history of lumbricid earthworms including *Eisenia* spp. and *L. terrestris* is found in a number of publications, including: Edwards and Lofty, 1977; Reynolds, 1977; Fender, 1985; Sims and Gerard, 1985; Curry, 1988; Bouché, 1992; Christensen and Mather, 1994; and Edwards and Bohlen, 1996. Lumbricid earthworms are important members of the soil fauna, and are appropriate organisms for use in the assessment of potentially *toxic* soils. Together with other earthworms, they constitute up to 92% of the invertebrate *biomass* of soil and are important in the maintenance of soil structure and nutrient cycling (Edwards and Lofty, 1977; Lee, 1985). Additionally, lumbricid and other earthworms represent a significant component of the diet of many species of birds, small mammals, reptiles, amphibians, and invertebrates (Macdonald, 1983; Cooke *et al.*, 1992). Earthworms can accumulate a variety of organic and inorganic compounds which might (or might not) affect them adversely (Edwards and Bohlen,

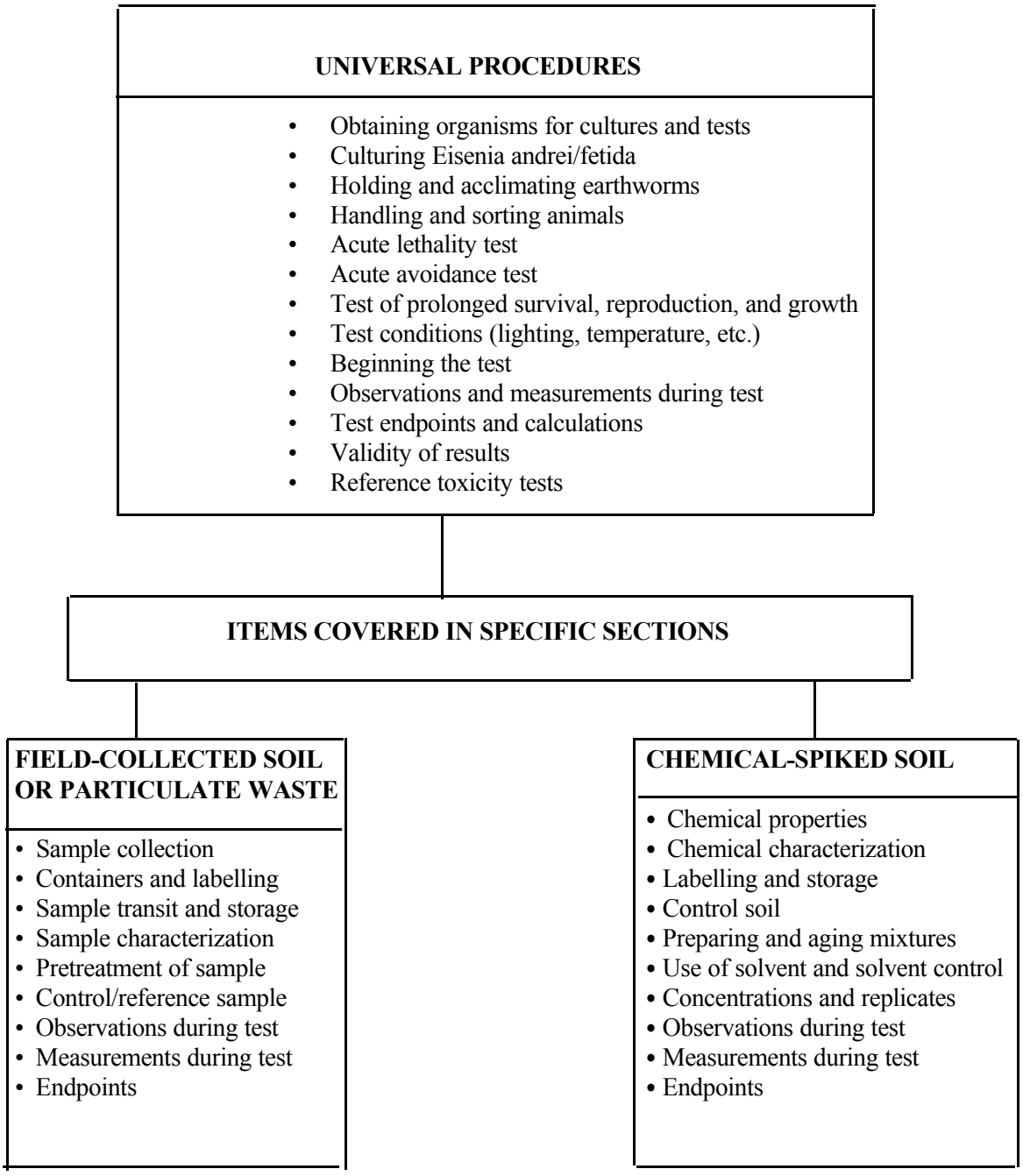


Figure 1 Considerations for Preparing and Performing Soil Toxicity Tests Using Earthworms (*Eisenia* sp. or *Lumbricus terrestris*) and Various Types of Test Materials or Substances

1992). A major change in the abundance of lumbricid earthworms could have serious adverse ecological effects on the entire terrestrial system (ASTM, 1999b).

1.2.1 Eisenia andrei/fetida

Eisenia andrei/fetida are commonly referred to as the *red wiggler, compost worm*, or *manure worm* (Aquaterra Environmental, 1998a). Taxonomists have difficulty distinguishing *E. andrei* from *E. fetida*, and morphological features alone are insufficient to enable this (R. Blakemore, personal communication, VermEcology, Canberra, Australia, 2000; W.J. Diehl, personal communication, Department of Biological Sciences, Mississippi State University, Mississippi State, MS, 2000) . However, a definitive identification can be made based on differing electrophoretic patterns of certain enzymes for these two species[1] (Jaenike, 1982; Øien and Stenersen, 1984; McElroy and Diehl, 2001; McCann, 2004). It now appears that numerous researchers have misidentified *E. andrei* as *E. fetida* (or as *E. fetida andrei*) in much of the literature reported to date, with *E. andrei* being the sibling species found most commonly in North American composts or cultures from commercial suppliers of earthworms of *Eisenia* spp. (W.J. Diehl, personal communication, Department of Biological Sciences, Mississippi State University, Mississippi State, MS, 2000; McCann, 2004).

Historically, *E. fetida* has been referred to by some as a "species complex" (Bouché, 1992; Christensen and Mather, 1994). *E. fetida*/*andrei* has also been described by certain taxonomists as having two morphologically similar subspecies or races (i.e., *E. fetida fetida* which typically has transverse striping or banding on its segments, and *E. fetida andrei* which lacks this and has a variegated reddish colour). This (now outdated) means of classification has been adopted in certain biological test methods for measuring soil toxicity using earthworms (OECD, 1984; ISO, 1993; ASTM, 1999b). However, earthworm taxonomists now classify *E. andrei* and *E. fetida* as distinct species, while recognizing that morphological characteristics including colouration and segmental banding or striping patterns are insufficient to distinguish them with complete confidence (R. Blakemore, personal communication, VermEcology, Canberra, Australia, 2000; W.J. Diehl, personal communication, Department of Biological Sciences, Mississippi State University, Mississippi State, MS, 2000). More recent reports on soil toxicity tests and related methodology documents have referred to *E. fetida* (or *E. fetida fetida*) and *E. andrei* (or *E. fetida andrei*) as distinct species (Sheppard, 1988; ISO, 1991, 1998; OECD, 2000). This approach is supported by the results of recent breeding experiments involving these two species, which found that cocoons were produced when *E. fetida* and *E. andrei* were cross-bred although none of these were viable (Ferreiro *et al.*, 2002). Investigators using *E. andrei* or E. *fetida* as the test organism must indicate which species is used, based on an examination of electrophoretic patterns[1] (or other molecular tests proven to distinguish these two species) together with recognized taxonomic distinctions (to genus). Qualified personnel should confirm this identification (see Sections 2.1 and 7.1.2). To date, the species of *Eisenia* typically used in Canadian laboratories for soil toxicity tests has now been confirmed to be *E. andrei* (W.J. Diehl, personal communication, Department of Biological Sciences, Mississippi State University, Mississippi State, MS, 2000; McCann, 2004), although it was formerly identified as *E. fetida* (e.g., Aquaterrra Environmental, 1998a; Aquaterra Environmental and ESG, 2000; EC, 2000a, b).

Eisenia spp. (i.e., *E. andrei* or *E. fetida*) are commonly found in North American composters and are sold commercially for fish bait (as "red wigglers") and composting (as "compost worms"). *Adult* worms of these species have an average body length of 35–130 mm and an average diameter of 3–5 mm, with between 80 and 110 body segments. Diagnostic characteristics common to both species include an *epilobic prostomium*, first dorsal pore on 4/5 or sometimes 5/6 segments, and a *clitellum* on segments 24–32 (Reynolds, 1977). The *tubercula pubertatis* is

[1] Using starch gel electrophoresis, the pattern of polymorphism in the enzyme phosphoglucomutase is distinctive for these sibling species. *E. andrei* has two alleles (bands) with the slower homozygous band being in greater frequency than the faster homozygous band; whereas *E. fetida* has two different alleles with the faster homozygous band being in greater frequency than the slower homozygous band. The pattern of polymorphism in the enzyme mannose phosphate isomerase is also distinctive. In *E. andrei*, there are two alleles, whereas in *E. fetida* there is only one allele and it is different from the ones in *E. andrei* (McElroy and Diehl, 2001; McCann, 2004).

found on segments 28–30. The *setae* are closely paired with a characteristic variation in patterning that differs from the anterior to the posterior end of the worm. *Genital tumescences* might be present around any of the setae on segments 9–12 of the cylindrical body that can vary in colour from red to dark red, brownish red, or purple, with alternating bands of red-brown pigment and pigmentless yellow inter-segmental areas (Reynolds, 1977). The male pores usually have large glandular *papillae* on segment 15. The *spermathecae* are two pairs with ducts, which open on segments 9/10 and 10/11. Four pairs of *seminal vesicles* are found on segments 9–12.

E. andrei/fetida are native to the Palearctic, and are also found in Europe, North and South America, Asia, Africa, Iceland, and Australasia (Reynolds, 1977). One or both of these gregarious species are generally associated with anthropogenic activities, and are commonly found across North America in gardens, compost, and manure piles (Edwards and Lofty, 1977). Within Canada, *E. andrei/fetida* have been found in the provinces of British Columbia, Alberta, Ontario, Quebec, New Brunswick, Nova Scotia, and Prince Edward Island (Reynolds, 1977; M.J. Clapperton, personal communication, Research Branch, Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge, Alberta, 2000). These sibling species prefer moist soils with high levels of *organic matter*. They are generally restricted to the upper layers of soil and are considered to be *epigeic*, selectively feeding (with little ingestion of soil) on organic material dispersed throughout the soil (Wallwork, 1983). *E. andrei/fetida* are tolerant of a wide range (i.e., 4 to 8) of soil *pH* values (Stephenson, 2002), although they prefer soils with a pH between 7 and 8 (Edwards and Lofty, 1977).

These worms *(E. andrei/fetida)* are thought to have a life span of four to five years, although between one and two years is more common (Reynolds, 1977). *E. andrei/fetida* are obligatorily *amphimictic*, although uniparental reproduction has been reported (Reynolds, 1977, 1995). *E. andrei/fetida* copulates and casts below ground. One or both of these species reproduce rapidly at temperatures ranging within 20–25 °C and can reach sexual maturity within 52 days. Time for completion of a life cycle is appreciably slower at cooler temperatures (e.g., >166 days at 13 °C) (ASTM, 1999b). *Cocoons* are produced at a frequency of one or two, every three or four days; each cocoon can produce as many as six or more *hatchlings*, although one to four offspring per cocoon is more commonly observed (Reinecke and Viljoen, 1991; Reinecke *et al.*, 1992; Edwards and Bohlen, 1996). These characteristics (i.e., rapid rate of cocoon production, large number of offspring, short generation time, rapid maturation time) and the fact that *E. andrei/fetida* can be easily cultured in the laboratory (see Appendix E) make these earthworms the most commonly used test species for earthworm reproduction tests (Aquaterra Environmental, 1998a).

There are few studies on the relative sensitivity of these two related species to samples of contaminated soil. Side-by-side laboratory tests by Ingraldi *et al.* (2004), performed according to Section 4.1.8 herein, showed that *E. andrei* was somewhat more sensitive than *E. fetida* to boric acid in artificial soil, with seven-day LC50s of 3236 mg/kg and 4365 mg/kg, respectively and 95% confidence limits that did not overlap. Similarly, results for 14-day LC50s performed concurrently by these investigators (according to Section 4.1 herein) using each of these two species exposed to a sample of condensate-contaminated soil mixed in an uncontaminated clay loam soil, indicated a somewhat greater sensitivity of *E. andrei*. Comparative 48-h avoidance tests with multiple concentrations of this same condensate-contaminated soil in clean clay loam soil, performed according to Section 4.2, herein, also showed a greater sensitivity of response by *E. andrei* to the contaminated soil (Ingraldi *et al.*, 2004). Accordingly, these two species of earthworms should not be used interchangeably (i.e., as a mixed culture) or without distinction to species, when undertaking soil toxicity tests.

1.2.2 Lumbricus terrestris

Lumbricus terrestris is commonly referred to as the dew worm or night crawler (Aquaterra Environmental, 1998a). Unlike *Eisenia* spp., the taxonomy of this lumbricid earthworm is straightforward and relatively free from controversy. Adult worms of this species have an average body length of 90–300 mm and an average diameter of 6–10 mm, with between 120 and 160 body segments. Diagnostic characteristics include a *tanylobic prostomium*, first dorsal pore on the 7/8 segment, and the clitellum on segments 21–27. The tubercula pubertatis is found on segments 23–36. Setae are enlarged and widely paired in the caudal and cephalic regions, but closely paired and smaller in the

central regions. The body is cylindrical and posteriorly, it is strongly compressed dorsoventrally. It is heavily pigmented brownish-red on the dorsal surface, and yellowish-orange in colour on the ventral surface. The female genital pore is located on segment 13, and is usually accompanied by genital tumescences. Male pores are located on segments 14–16, with large elevated glandular papillae. Three pairs of seminal vesicles are found on segments 9, 11, and 12/13. The spermatheca has two pairs of short ducts opening at segments 9/10 and 10/11 (Reynolds, 1977).

L. terrestris is Palearctic, and has been found in Europe, Iceland, North America, Siberia, South Africa, and Australasia (Reynolds, 1977). Within Canada, *L. terrestris* is found in most provinces including British Columbia, New Brunswick, Newfoundland, Nova Scotia, Ontario, Prince Edward Island, and Quebec; but has yet to be identified in the Yukon or Northwest Territories or Nunivut (Reynolds, 1977; M.J. Clapperton, personal communication, Research Branch, Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge, AB, 2000). *L. terrestris* inhabits various soil types and can tolerate a fairly wide range of soil *pH* (i.e., 4–8) values. It typically frequents meadows, grasslands, pastures, and golf courses, and commonly occupies fields with wheat-corn-soybean rotations (Tomlin, 1995). This species is seldom found in the forested regions of North America (Reynolds, 1977). It is non-gregarious and exhibits behaviour that suggests it might be territorial in terms of food supply (Tomlin *et al.*, 1993). *L. terrestris* is *anecic*, selectively feeding on organic material found at the surface of the soil while burrowing deeply into the soil. The relatively permanent, vertical burrow systems of *L. terrestris* are unique in that they are often lined with pebbles, faecal earth, and organic material drawn into the burrows from the soil surface. Although this species will form middens at the soil surface, it often casts below ground.

L. terrestris copulates at the soil surface. Copulation is year-round in favourable climates, although a summer and winter period of rest might be climatically imposed in northern temperate regions (Reynolds, 1977). This large earthworm, which remains active for most of the year in temperate climates (Edwards and Bohlen, 1992), reaches sexual maturity within 8–16 months and reproduces at a relatively slow rate. Typically, the species produces 3 to 13 cocoons per year with, on average, three *progeny* per cocoon (Wallwork, 1983); growth is optimal at ~10 °C (Edwards and Lofty, 1977). *L. terrestris* is believed to have a life span of ~6.5 years (Fox, 2000). This species cannot be bred easily under laboratory conditions (Edwards and Bohlen, 1992), although field-collected animals can be maintained there in good health for an extended period of time (i.e., many months).

1.3 Historical Use of Earthworms in Toxicity Tests

Earthworms are frequently exposed to toxic chemicals in soil. Besides the myriad of fertilizers, insecticides, herbicides, and fungicides from agricultural and domestic applications, earthworms are sometimes exposed to heavy metals, petroleum hydrocarbons, or other chemicals such as wood preservatives (e.g., pentachlorophenol) or nitroaromatic explosive compounds in contaminated soils.

Earthworms are widely used as test organisms in single-species laboratory toxicity tests intended to measure the toxicity of pure chemicals, chemical products, or samples of soil contaminated or potentially contaminated with chemicals in the field or (for experimental purposes) in the laboratory. The toxicity of various chemicals or chemical products to earthworms, as determined in the laboratory under standardized conditions using lethal and/or sublethal endpoints and acute (hours or a few days) or prolonged (several weeks) exposures has been reported (Heimbach, 1984; Neuhauser *et al*., 1985, 1986; van Gestel, 1991; Edwards and Bohlen, 1992; Heimbach, 1993; Callahan *et al.*, 1994; Leon and van Gestel, 1994). Most of the studies reviewed in these reports refer to tests for acute lethality (*LC50*s), with exposures of 14 days or less. Edwards and Bohlen (1992) appraised various lethal and sublethal responses to chemicals or chemical products reported for laboratory toxicity tests involving acute or more prolonged ("*chronic*") exposures.[2]

---

[2] The term "*chronic*" is defined herein as occurring during a relatively, long period, usually a significant portion of the life span of the organism (e.g., 10% or more). Since the life span of *Eisenia* spp. can be as long as 4–5 years (Reynolds, 1977; Section 1.2.1), the term "*chronic*" is not applied to the eight-week test for effects of prolonged exposure defined in Section 4.3. Notwithstanding, the intent of this test is to estimate, approximately, what *chronic toxicity* might be.

The use of earthworm toxicity tests as "ecotoxicological assessment tools" for appraising the toxicity of contaminated or potentially contaminated *site soil* is increasing in Canada and elsewhere (Callahan, 1988; Menzie *et al.*, 1992; Römbke *et al.*, 1994; Kula and Larink, 1997; Spurgeon *et al.*, 1994; Spurgeon and Hopkin, 1995, 1996a; Yeardley *et al.*, 1996; Chang *et al.*, 1997; Meier *et al.*, 1997; Stephenson *et al.*, 1997; Aquaterra Environmental, 1998a; Saterbak *et al.*, 1999; Stephenson *et al.*, 2002; Stephenson, 2003a). Studies comparing the results of single-species toxicity tests performed in the laboratory with related field surveys for effects on terrestrial biota have generally found a strong correlation between the laboratory findings and the field results (Edwards and Bohlen, 1992; Kula and Kokta, 1992; Menzie *et al.*, 1992; van Gestel, 1992, 1997; Heimbach, 1993, 1997; Christensen and Mather, 1994; Kula, 1995). Scientists, however, have frequently commented that it is difficult to extrapolate results for single-species laboratory tests with earthworms to the field situation. A number of researchers have discussed how to improve the predictive worth of the laboratory toxicity tests (i.e., their ability to discern adverse environmental conditions or effects). Promising improvements include reliable procedures for estimating the bioavailability of inorganic and organic contaminants in soil, tiered testing approaches, and *risk* assessment schemes for soil toxicity that include earthworm toxicity tests (Bouché, 1988; Callahan, 1988; Lofs-Holmin and Bostrom 1988; Keddy *et al.*, 1995; NERI, 1993; Leon and van Gestel, 1994; Christensen and Mather, 1994; Sauvé *et al.*, 1996, 1998, 2000; Barber *et al.*, 1997; Meier *et al.*, 1997; Saterbak *et al.*, 1999; Conder and Lanno, 2000; and Wells and Lanno, 2001).

A number of investigators have studied the effects of variations in natural characteristics of *chemical-spiked soil* or *site soil*, on the soil's toxicity to earthworms. Variables investigated include soil pH, organic carbon content, particle size, and moisture content (Heimbach and Edwards, 1983; van Gestel and van Dis, 1988; van Gestel, 1991; Christensen and Mather, 1994; Spurgeon and Hopkin, 1996b; Yeardley *et al.*, 1996; Bauer and Römbke, 1997; Puurtinen and Martikainen, 1997; Meharg *et al.*, 1998; Aquaterra Environmental and ESG, 2000). The influence of these soil variables on chemical toxicity depends on interactions between the physicochemical characteristics of the soil and the type(s) and speciation of chemical *contaminant(s)* therein.

A variety of test methods have been used to measure the acute lethal toxicity of chemicals to earthworms in the laboratory (see Section 1.3.1). Laboratory tests which measure the effects of *contaminated soil* on the behaviour of earthworms are increasingly used (see Section 1.3.2), as are those which measure the effects of prolonged exposures on earthworm survival, reproduction, and growth (see Section 1.3.3). Certain researchers have also studied or reviewed other sublethal (e.g., gametogenic, teratogenic, neurotoxic, or immunotoxic) effects of chemical-contaminated soil on earthworms (Drewes *et al.*, 1984; Zoran *et al.*, 1986; Edwards and Bohlen, 1992; Fitzpatrick *et al.*, 1992; Cikutovic *et al.*, 1993; Goven *et al.*, 1993, 1994; Christensen and Mather, 1994; Suzuki *et al.*, 1995; Brousseau *et al.*, 1997; Giggleman *et al.*, 1998; Scott-Fordsmand *et al.*, 2000).

1.3.1 Acute Lethality Tests

A number of diverse laboratory methods have been used to measure the *acute toxicity* of specific chemicals or chemical products to earthworms, including (Edwards and Bohlen, 1992):

- **immersion tests**—groups of earthworms are immersed for a fixed time in solutions containing ranges of chemical concentrations;
- **topical application tests**—chemicals are applied to the surface of earthworms using a paint brush or microapplicator and the animals are observed for post-treatment mortality rates;
- **injection tests**—solutions of chemicals are injected into the coelomic cavity of earthworms that are then held in soil and assessed for post-treatment mortality rates;
- **forced feeding tests**—test chemicals suspended in agar gel are injected into the esophagus of earthworms after which they are placed onto moistened paper for observation of post-treatment mortality rates;
- **voluntary feeding tests**—chemical-contaminated food is offered to earthworms after which they are observed for mortality rates;
- **contact filter paper tests**—standard filter papers are treated with a range of concentrations of the test chemical after which individual earthworms are

exposed to filters held in vials (one/vial) for a fixed period (typically 48 or 72 h; OECD, 1984);

- **"artisol" tests**—chemicals are applied to a matrix of silica that is then moistened and suspended in glass balls, followed by the exposure of groups of earthworms for 14 days for determinations of mortality rates;

- **funnel tests**—individual earthworms are placed into funnels filled with soil, then burrows are established and chemicals are added to the surface of the soil, with observations of mortality and *sublethal effects* on growth;

- **artificial soil tests**—a standard formulation of artificial soil is prepared and mixed ("*spiked*") with a range of concentrations of the test chemical(s), followed by the addition of a group of earthworms to each mixture in a test chamber and their observation for mortality rates after 7 and/or 14 days; and

- **natural soil tests**—a range of concentrations of the test chemical(s) is spiked in a *clean* (uncontaminated) natural soil, followed by the addition of a group of earthworms to each mixture in a test chamber and their observation for mortality rates after 7 and/or 14 days.

The advantages and disadvantages of these varied approaches for measuring the acute toxicity of chemicals to earthworms have been assessed and compared (Heimbach, 1984, 1985, 1988; Neuhauser *et al.*, 1986; van Gestel and van Dis, 1988; Edwards and Bohlen, 1992; ESP, 1992; Römbke *et al.*, 1992; Callahan *et al.*, 1994; Christensen and Mather, 1994; Leon and van Gestel, 1994; Robidoux *et al.*, 1999). Certain investigators have proposed or applied modifications to these methods for measuring the acute toxicity of chemicals or *test soils* (e.g., Karnak and Hamelink, 1982; Fitzpatrick *et al.*, 1992; Sheppard and Evenden, 1992; Giggleman *et al.*, 1998).

Appendix F summarizes the procedures and conditions recommended historically or currently by various agencies, for measuring the acute lethal toxicity to earthworms of samples of soil contaminated with chemical(s) or chemical product(s). Some of these test methods (i.e., OECD, 1984; ISO, 1993) are designed and intended explicitly to measure the acute toxicity of specific chemicals spiked in artificial soil at a range of *concentrations*. Others are designed and intended for measuring the acute lethal toxicity of samples of contaminated or potentially contaminated *site soil*, although these methods (with appropriate modifications) can also be applied when testing chemicals spiked in soil (i.e., USEPA, 1989; EC, 2000b). The ASTM (1999b) offers a standard guide, which can be applied to either *chemical-spiked soil* or *site soil*, for testing the acute lethal toxicity of soil to earthworms. Each of these biological test methods, for which procedures and conditions are summarized in Appendix F, uses *E. andrei/fetida* as the test organism and has a test duration of 14 days. These methodology documents have been used as guidance in developing Environment Canada's standard biological test method for measuring the acute lethality of samples of *chemical-spiked soil* or *site soil* (see Section 4.1).

1.3.2 Acute Avoidance Tests

Lumbricid earthworms including *L. terrestris* and *Eisenia* spp. are known to be highly mobile (Karnak and Hamelink, 1982; Mather and Christensen, 1992). A number of researchers have concluded that a behavioural avoidance response by earthworms to sublethal concentrations of chemicals in soil can have ecological relevance at the population level (Christensen and Mather, 1994; Tomlin, 1995; Yeardley *et al.*, 1996). Some evidence suggests that these and certain other terrestrial invertebrates are able to minimize exposure to harmful chemicals through such behaviour (Yeardley *et al.*, 1996; Haimi and Paavola, 1998). Christensen and Mather (1994) reviewed the use of earthworms as test organisms for evaluating chemical hazards and as part of *ecological risk assessments* for the Danish Environmental Protection Agency. They concluded that, from an ecological viewpoint at the population level, toxicity tests that measured effects on migratory (avoidance) behaviour were amongst those considered as "ideal endpoints" and recommended their application. Advantages of tests for an avoidance response include their short duration (relative to 14-day lethality tests or more prolonged tests for effects on reproduction and/or growth) and their sensitivity (i.e., their ability to detect a behavioural response at sublethal concentrations). Earthworms exposed to contaminated soil typically show an avoidance response to sublethal concentrations within 24–72 h of exposure (Wentsel and Guelta, 1988; Yeardley *et al.*, 1996; Slimak, 1997; Hund, 1998; Stephenson *et al.*, 1998; Hund-Rinke and Wiechering, 2001; ESG and Aquaterra Environmental, 2002; Schaefer, 2003, Stephenson, 2003a).

Laboratory tests that measure acute avoidance are particularly useful, from an ecological perspective, when performed in conjunction with standard toxicity tests such as those that measure lethality or sublethal effects on growth and/or reproduction. The ecological relevance of findings for an acute avoidance test in the absence of comparable data for such standard toxicity tests might produce confusing or questionable results, since the earthworms might avoid concentrations of contaminants that are not damaging to their tissues or might fail to avoid concentrations that are.

Apparatus and procedures used for measuring avoidance responses of earthworms to contaminated soil have been varied, and international efforts to develop a standardized behaviour test method have been lacking. Using a series of rectangular chambers, Wentsel and Guelta (1988) determined that *L. terrestris* avoided soil contaminated with brass powder at sublethal concentrations below those causing weight loss in adults. In tests using circular chambers, Yeardley *et al.* (1996) found that *E. andrei/fetida* avoided sublethal concentrations of chemical-spiked soils or toxic site soils, when exposures were as brief as 1 to 2 days. Slimak (1997) reported that *L. terrestris* avoided sublethal concentrations of several pesticides. Haimi and Paavola (1998) reported an avoidance response of earthworms (*Aporrectodea tuberculata*) when they used pots containing pentachlorophenol-spiked soil in some sectors and *clean* soil in others. Using the test design of Yeardley *et al.* (1996), Hund (1998) reported an avoidance response of earthworms (species unidentified) to organic contaminants at threshold concentrations similar to those causing reproductive effects in 56-day tests. Confounding effects due to differing physicochemical characteristics (e.g., compactness, particle size, organic carbon content) were considered by both Yeardley *et al.* (1996) and Hund (1998) to be minimal.

During the past seven years, a number of studies have been performed to develop and standardize the acute avoidance test for soil toxicity (Section 4.2). The experimental apparatus used is illustrated photographically in Stephenson *et al.* (1998), and schematically in Section 3.1.3 (as Figure 2). Initial studies (Stephenson *et al.*, 1998; Aquaterra Environmental, 1998a), using *L. terrestris*, found that worms given a choice between *clean* natural soil and *clean* artificial soil showed a preference for the natural soil. Subsequent tests with either *L. terrestris* or *E. andrei* given a choice between autoclaved or non-autoclaved *clean* soil (natural or artificial) showed in each instance that each species preferred the non-autoclaved soil. Acute (i.e., 24 or 72 h) avoidance tests with *L. terrestris* or *E. andrei* given a choice between *negative control soil* (natural or artificial) and various concentrations of a condensate-contaminated site soil diluted with the respective negative control soil showed a concentration-dependent avoidance response at sublethal concentrations. Associated prolonged-exposure tests with *E. andrei* and the same sample of condensate-contaminated site soil indicated that the threshold concentration avoided by this species of earthworm was similar to the threshold-effect concentration which reduced reproductive success and subsequent growth of offspring. *L. terrestris* was somewhat more sensitive than *E. andrei* to the condensate-contaminated site soil or to potassium chloride (reference toxicant), although the time to elicit an avoidance response by *L. terrestris* was longer (Stephenson *et al.*, 1998; Aquaterra Environmental, 1998a). Similar results were obtained for a site soil contaminated with amines and glycol products, in that the earthworms avoided sublethal concentrations in soil that resulted in adverse effects on reproduction (Aquaterra Environmental, 1998a; Stephenson, 2003a).

Aquaterra Environmental and ESG (2000) performed acute (72-h) avoidance tests as well as 14-day lethality tests with *L. terrestris* or *E. andrei* exposed to a range of concentrations of copper (as copper sulphate) spiked in each of three *clean* natural soils differing in sand/silt/clay content. In each instance, a concentration-dependent avoidance response to copper was found. Associated 64-day tests for long-term effects (i.e., mortalities, inhibition of reproduction, and/or growth inhibition) of copper-spiked soil on *E. andrei* indicated that, for the same copper-soil mixture, the threshold-effect concentrations affecting adult reproduction or subsequent growth of their progeny were similar to the threshold concentration which caused an acute (72-h) avoidance response. Similar results were observed in tests with the fungicide benomyl (wettable powder formulation). After 72 h, *E. andrei* avoided sublethal concentrations of benomyl in both artificial and field-collected soils similar to the threshold concentrations that resulted in adverse effects on reproduction (Elshayeb *et al.*, 2001; Feisthauer *et al.*, 2001; Stephenson, 2003a).

Standardized procedures and conditions for performing a biological test method that measures the acute avoidance response of earthworms (*L. terrestris*,

E. andrei, or *E. fetida*) to *chemical-spiked soil* or *site soil* are defined in Section 4.2.

1.3.3 Tests for the Effects of Prolonged Exposure on Survival, Reproduction, and Growth

The effects of prolonged exposure to toxic substances or materials on the survival, reproduction, and growth of a single species of test organism, under controlled laboratory conditions, are recognized and accepted by environmental toxicologists as *ecologically relevant* responses. From an ecological viewpoint, these biological effects represent "ideal endpoints" for laboratory toxicity tests with earthworms (Christensen and Mather,1994). Christensen and Mather (1994) recommended their inclusion in an assessment protocol, following their review of the use of earthworms as test organisms for evaluating the ecological risk of toxic chemicals in soil.

In 1988, international efforts were initiated to develop and standardize tests for measuring the effects of long-term exposure to contaminants in the soil on survival, reproduction, and growth of earthworms (van Gestel *et al.*, 1988). A number of standard methods or guidelines were developed using *E. andrei/fetida*; these are now commonly applied and their use is expanding.

E.andrei/fetida is a preferred test organism for studying the effects of prolonged exposure to contaminants on the survival, reproduction, and growth of earthworms, because of the widespread knowledge and experience in culturing this species (see Appendix E), its rapid (relative to other earthworms such as *L. terrestris*) life cycle, its international distribution, and its frequent use in acute lethality tests (see Section 1.3.1 and Appendix F). The development, growth, and reproductive biology of *E. andrei/fetida* under laboratory conditions has been extensively studied and is well documented (e.g., Edwards and Lofty, 1977; Tsukamoto and Watanabe, 1977; Sheppard, 1988; van Gestel *et al.*, 1992a). The toxic effects of prolonged exposure to contaminated soil on the survival, reproduction, and/or growth of *E. andrei/fetida* have been documented in:

- laboratory studies involving samples of soil spiked or contaminated with pesticides (Lofs-Holmin, 1980; Venter and Reinecke, 1988; Neuhauser and Callahan, 1990; van Gestel *et al.*, 1992b; Riepert and Kula, 1996; Bauer and Römbke, 1997; Heimbach, 1997; Kula and Larink, 1997; ESG and Aquaterra Environmental, 2002);
- heavy metals (Neuhauser *et al.*, 1984; van Gestel *et al.*, 1989, 1992b; Spurgeon *et al.*, 1994; Reinecke and Reinecke, 1996; Spurgeon and Hopkin, 1996a; Fischer and Molnar, 1997; Kula and Larink, 1997; Aquaterra Environmental and ESG, 2000; Scott-Fordsmand *et al.*, 2000; ESG, 2002);
- petroleum hydrocarbons (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1998, 1999a, b, 2000a; ESG, 2001); and
- other chemicals including reference toxicants (Hartenstein, 1982; van Gestel *et al.*, 1989, 1992b; Neuhauser and Callahan, 1990; Gibbs *et al.*, 1996; Aquaterra Environmental, 1998a; Robidoux *et al.*, 2000, 2001).

In their initial efforts to develop a standardized test method for determining the effect of chemical substances on the reproduction of *E.andrei/fetida*, van Gestel *et al.* (1988) performed five-week incubation studies that measured cocoon viability and numbers of hatchlings per cocoon, following cocoon recovery from earlier chemical-exposure studies with adult earthworms and their incubation in water or artificial soil. Subsequently, van Gestel *et al.* (1989) described a Dutch test method whereby adult *E. andrei/fetida* were preconditioned for one week in artificial soil and exposed thereafter to a range of concentrations of chemical-spiked artificial soil, after which cocoons produced were incubated for a further five weeks in untreated artificial soil to assess hatchability. This (Dutch) method was submitted to the International Standard Organization (ISO) for consideration as a standard test method for measuring soil toxicity (ISO, 1991; see Appendix G for test specifics).

In 1990, a German working group established by the Biologische Bundesanstalt für Land- und Forstwirtschaft (BBA) and joined by experts from the Netherlands and Switzerland developed a slightly different test method, whereby adult *E. andrei/fetida* were exposed to chemical-spiked artificial soil for four weeks. After this time, the worms were removed and the exposure continued with their progeny for a further four weeks. This (BBA) draft method was introduced in 1990 by Germany to the ISO Working Group "Soil Fauna" (WG 2), and was later published by the BBA (1994) as a guideline for testing the toxicity of pesticides. Following further evaluation and consideration by other scientists, ISO (1998) published

a modified version of the BBA (1994) method. The standard method published by the ISO (1998) consists of a four-week exposure of adult *E. andrei/fetida* to a range of concentrations of chemical-spiked soil with observations thereafter (adult survival rates and their increase or decrease in wet weight), and a subsequent four-week exposure to the same chemical-spiked soils with an endpoint measurement of number of offspring (*juveniles*) produced per *treatment* (see Appendix G).

The Organization for Economic Cooperation and Development (OECD) is presently circulating a draft of a similar test method (OECD, 2000; see Appendix G for specifics), for input to and comments by informed scientists. A shorter (28-day) test, restricted to determinations of the survival and weight change of adult *E. andrei/fetida* during exposure, has been drafted (USEPA, 1996) and applied (Gibbs *et al.*, 1996) as a "cost-effective" method for screening samples of contaminated soil. This shorter test method, however, is not widely used because it does not measure the effects on the reproduction of earthworms and the survival and growth of their progeny.

Standardized procedures and conditions for performing a biological test method that measure the toxic effects of prolonged exposure to *chemical-spiked soil* or *site soil* on the survival, reproduction, and growth of *E. andrei* are defined herein (see Section 4.3). This biological test method is largely in keeping with ISO (1998) and OECD (2000).

Test Organisms

2.1 Species and Life Stage

Both the acute lethality test (Section 4.1) and the acute avoidance test (Section 4.2) may be performed using *Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris* as test organisms. The eight-week test for effects on survival, reproduction, and growth (Section 4.3) must be performed using only laboratory-cultured *E. andrei* (Section 2.3).

The acute lethality test may be performed using either *sub-adult* or *adult* worms, provided that the wet weights of individual worms used to start the test range within 250–600 mg if *E. andrei/fetida*, or 3–10 g if *L. terrestris* (see Section 4.1.1). For an acute avoidance test, the wet weight of each adult worm used to start the test must range within: 250–600 mg if *E. andrei*, 250–800 mg if *E. fetida*, or 3–10 g if *L. terrestris* (see Section 4.2.1). For an eight-week test for effects on the survival, reproduction, and growth of *E. andrei*, the wet weight of each adult worm used to start the test must range within 250–600 mg (see Section 4.3.1).

The identification, distribution, and life history of *E. andrei/fetida* and *L. terrestris* are summarized in Section 1.2. Species identification must be confirmed and documented by qualified personnel experienced with identifying the intended species (see Section 1.2) of earthworm to be used in the toxicity test.[3] Cultures of *Eisenia* sp. held for a prolonged period at a testing laboratory should be identified to species at least once every two years.

2.2 Source

Sources of *E.andrei/fetida* may be government or private laboratories which are culturing this species of earthworm, or a commercial biological supplier (see McCann, 2004 for a list of potential sources). If *E. andrei* or *E. fetida* is used for an acute lethality test (Section 4.1) or an acute avoidance test (Section 4.2), the worms may either be cultured in the laboratory (see Section 2.3) or obtained from outside *cultures* or commercial suppliers and acclimated to laboratory conditions (Section 2.4) before the test is initiated.[4,5] If *E. andrei* is used for an eight-week test for effects on survival, reproduction, and growth (Section 4.3), laboratory-cultured worms of this species (see Sections 1.2.1 and 2.3) must be used as the source of the test organisms. When establishing laboratory cultures of *E. andrei* or *E. fetida*, it is recommended that cocoons (rather than juvenile or adult worms) be obtained to standardize the age and weights of individual worms within the culture.

Sources of *L. terrestris* may be government or private laboratories which are holding populations of this species of earthworm, or a commercial biological

[3] Differing electrophoretic patterns of certain enzymes can be used to distinguish *Eisenia* spp. as either *E. andrei* or *E. fetida* (see Section 1.2.1). Laboratory personnel seeking advice on species identification and confirmation might wish to contact: Dr. W. Diehl (Department of Biological Sciences, Mississippi State University, 130 Harned Hall, Lee Boulevard, Mississippi State, MS 39762; phone: (662) 325-7576; e-mail, wdiehl@biology.msstate.edu); and/or J.I.Princz (Biological Methods Division, Environment Canada, Ottawa, ON, K1A 0H3; phone: (613) 990-9544; fax: (613) 990-0173; e-mail, Juliska.Princz@ec.gc.ca.

[4] Investigators might be concerned with the effects of excessive inbreeding of laboratory cultures, or might wish to use progeny produced from organisms that occupied a particular locale. Accordingly, cultures may also be established using wild populations. If animals are obtained from a wild population, their taxonomy should be confirmed and they or their progeny should be evaluated for sensitivity to *reference toxicant(s)* before being used in toxicity tests.

[5] The practice of obtaining wild populations of adult *E. andrei/fetida* or *L. terrestris* for use in toxicity tests (following their *acclimation* to laboratory conditions; see Section 2.4) should be avoided unless their taxonomy has been confirmed and the collection site is considered to be *clean*. An (atypical) exception to this is when the invetigator(s) wants to use earthworms with a prior history of exposure to contaminated soil, in which case their exposure history should be known and stated. Ideally, any site from which field-collected specimens are taken should be known to be free of any applications or sources of pesticides or fertilizers during the past five years or longer.

supplier. This species of earthworm is extremely difficult to breed in cultures, and its slow rate of reproduction and development precludes mass breeding (Section 1.2.2; Edwards and Bohlen, 1992; C.A. Edwards, personal communication, Department of Entomology, Ohio State University, Columbus, OH, 2001). Accordingly, toxicity tests performed using *L. terrestris* typically rely on field-collected animals as a source of test organisms. Any field-collected earthworms to be used in toxicity tests, whether collected directly by laboratory personnel or others (e.g., a commercial supplier), should be obtained from *clean* (uncontaminated) grassland sites unless the intent is to use worms with a prior history of exposure to contaminants.[5]

Breeding stock of *E. andrei* or E. *fetida* can be obtained by contacting the following Canadian source:

Soil Toxicology Laboratory
Environmental Technology Centre
Environment Canada
335 River Road
Ottawa, ON, K1A 0H3
Phone: (613) 990-9544

Breeding stock of *E. andrei* can also be acquired from the following Canadian sources:

Aquatic Toxicology Section
Pacific Environmental Science Centre
Environment Canada
2645 Dollarton Highway
North Vancouver, BC, V7H 1B1
Phone: (604) 924-2500

Toxicology Laboratory
Atlantic Environmental Science Centre
Environment Canada
P.O. Box 23005
Moncton, NB, E1A 6S8
Phone: (506) 851-3486

All earthworms used in a soil toxicity test must be derived from the same population.

Worms to be used as a source of breeding stock or test organisms should be transported to the laboratory using a portion of the soil or other substrate to which they are adapted. Additional quantities of this substrate might be obtained for culturing and holding purposes, depending on culturing (Section 2.3) and *acclimation* (Section 2.4) conditions and requirements. Shipping and transport containers should be insulated to minimize changes in temperature during transit. If and as necessary, packaged ice or freezer packs should be included in the container(s) to ensure that the temperature in transit remains cool. Live organisms should be transported quickly to ensure their prompt (i.e., within 24 h) delivery. Excessive crowding of animals during shipment or transport should be avoided to minimize stress during transit.

Upon arrival at the laboratory, organisms may be held in the soil (or other substrate) used in transit while temperature adjustments are made, or they may be transferred to other culturing substrate (Section 2.3.5) or that for holding and acclimating test organisms (Section 2.4.5). If the nature (including its *texture* and *moisture content*) of the substrate in which worms were initially held (e.g., by a commercial supplier) or transported differs markedly from that in which they are to be cultured (Section 2.3.5) or acclimated (Section 2.4.5), it might be prudent to adapt the worms to an increasing percentage of the acclimation substrate over several weeks until they are held in 100% of this substrate.[6]

Soil temperature should be adjusted gradually to the exposure temperature to be used during culturing (Section 2.3.4) or when holding and acclimating the worms to test conditions (Section 2.4.4). Guidance for handling worms given in Sections 2.3.7 and 2.4.7 should be followed here when transferring worms from an outside source to culture chambers (Section 2.3.2) or those for holding and acclimating worms (Section 2.4.2). Other conditions during this interim holding period for acclimation of breeding stock or test organisms to laboratory conditions should be as similar as possible to those used for maintaining cultures

[6] Experience at Environment Canada's Soil Toxicology Laboratory (Environmental Technology Centre, Ottawa, ON) indicates that survival rates in cultures can be poor if the nature of the substrate used by a commercial supplier to culture *E. andrei* differs markedly from the laboratory's culturing substrate. Survival rates are markedly improved, in this instance, if the percentage of the laboratory's culturing substrate is increased gradually over several weeks until the earthworms are held in 100% of the culturing substrate (J. Princz, personal communication, Environment Canada, Biological Methods Division, Ottawa, ON, 2004).

(Section 2.3) or for holding and acclimating worms obtained for use in soil toxicity tests (Section 2.4).

2.3 Culturing of E. andrei/fetida

2.3.1 General

General guidance and recommendations for culturing *E. andrei/fetida* in preparation for soil toxicity tests are provided here. In keeping with the premise *"What might work well for one laboratory might not work as well for another laboratory"* (USEPA, 1994a; EC, 1997a, b, 2001), explicit directions regarding many aspects of culturing, including the choice of culture chamber, number of organisms per chamber, soil-renewal conditions, culturing substrate, and food type and ration, are left to the discretion and experience of laboratory personnel, although guidance and recommendations are provided herein. Performance-based indices[7] are used to evaluate the suitability of the cultured organisms for tests, and the acceptability of the test results. Cultures must have low mortalities, to be suitable for use in tests, and the cultured organisms must appear healthy and behave and feed normally. Additionally, those used as controls in the test must have acceptably low mortality rates and meet all criteria for a valid toxicity test (see Sections 4.1.3, 4.2.3, and 4.3.3). The acceptability of the culture should also be demonstrated by concurrent or ongoing tests using a *reference toxicant* (see Section 4.1.8). If a culture of organisms fails to meet these criteria, it should be discarded.

It is the responsibility of the laboratory to demonstrate its ability to obtain consistent, precise results using a *reference toxicant*, when initially setting up to perform soil toxicity tests with cultured *E. andrei* or *E. fetida*. For this purpose, intralaboratory precision, expressed as a *coefficient of variation* for the respective *LC50* data, should be determined by performing five or more tests with different lots (groups) of test organisms from the same source, using the same reference toxicant and identical procedures and conditions for each test (see Section 4.1.8).

When routinely performing soil toxicity tests with *E. andrei/fetida*, *reference toxicity tests* should be conducted monthly with the laboratory's cultures, using the conditions and procedures outlined in Section 4.1.8. If this monthly routine is not followed, the performance of individuals from the culture used to start a soil toxicity test should be evaluated in a reference toxicity test conducted concurrently. Additionally, the performance of any cultures that have been established recently using new breeding stock (Section 2.2) should be checked using a reference toxicity test, and the results determined to be acceptable (see Sections 2.3.9 and 4.1.8) before these cultures are used to provide test organisms.

Cultures of *E. andrei/fetida* should be observed frequently (e.g., once or twice per week). Ideally, records should be maintained documenting:

- the date a culture is started with cocoons and the estimated number of cocoons used to start the culture;
- dates of substrate renewal;
- feeding and watering regime (including type and quantity added on each occasion);
- soil quality measurements (e.g., pH, temperature, moisture content, *water-holding capacity*); and
- observations of culture health (e.g., behaviour and appearance of earthworms in culture, odour of substrate, location of worms in the container, amount of uneaten food in container).

A summary of the various conditions and procedures used by OECD (1984, 2000), USEPA (1989), ISO (1993, 1998), ASTM (1999b), and EC (2000a) for culturing *Eisenia andrei/fetida* is provided in Appendix E. These procedural specifics have presumably worked well in producing adult *E. andrei/fetida* for use in soil toxicity tests and, unless indicated otherwise in this report, provide useful guidance which may also be applied here. A checklist of recommended conditions and procedures for culturing *E. fetida* to generate offspring for use in soil toxicity tests is given in Table 1.

---

[7] Performance-based indices include those related to the survival and condition of cultured *E. andrei/fetida* intended for use in the test (Section 2.3.9); as well as the criteria that must be met by control organisms for a test to be valid (Sections 4.1.3, 4.2.3, and 4.3.3), and those related to the performance of groups of animals in *reference toxicity tests* (Section 4.1.8).

Table 1 Checklist of Recommended Conditions and Procedures for Culturing *Eisenia andrei* or *Eisenia fetida*, to Provide Test Organisms for Use in Soil Toxicity Tests

Source of brood stock for culture	—	cocoons, juveniles, or adults from a government, private, or commercial culture, all from the same source; identification to species (i.e., *E. andrei* or *E. fetida*) confirmed
Acclimation	—	gradually (recommend ≤3 °C/day) for temperature differences upon arrival
Culture chambers	—	breeding boxes of 10–50-litre capacity are suitable (e.g., plastic trays measuring ~30 × 40 × 15 cm or 60 × 40 × 20 cm, covered with perforated lid to allow air exchange and minimize evaporation); sides and/or lid transparent or translucent to enable light to contact surface of culturing substrate; recommended minimum depth, 10 cm
Temperature	—	daily average, 20 ± 2 °C; instantaneous, 20 ± 3 °C
Lighting	—	incandescent or fluorescent; intensity, 400–800 lux at surface of culture chamber; fixed photoperiod (e.g., 16h L:8h D or 12h L:12h D)
Type of substrate	—	optional (e.g., mixture of potting soil, artificial soil, and peat moss; or mixture of shredded un-inked paper, artificial soil, and sphagnum peat moss)
Hydration of substrate	—	hydrated with distilled or de-ionized water; moisture content sufficient to keep surface of bedding moist but with no standing water in the bottom of the culture chamber; soil particles should not adhere to earthworms
pH of substrate	—	adjusted to range within 6.0–7.5 using reagent-grade calcium carbonate
Renewal of substrate	—	as required, and at least once every 2–3 months; sort and transfer worms and cocoons manually; alternatively, prepare new tray of bedding, cover with contents of old tray, leave undisturbed under constant light for two days, then remove and discard old bedding
Monitoring substrate quality	—	temperature, pH, and moisture content measured once per week, in each culture chamber
Feeding	—	either cooked oatmeal, or alfalfa pellets saturated with water; feed once/week by placing in a shallow depression of the substrate and then covering it with a thin layer of substrate, after removing excess (unused) food
Maintenance of culture	—	examine substrate in culture chamber at least once/week; gently turn manually as necessary; remove dead, injured, or atypical (lethargic) worms; record condition of culture; maintain loading density of worms at ≤0.03 g/cm^3
Age/size for test	—	sub-adults or sexually mature adults with clitellum, if to be used in an acute lethality test; clitellated adults only, for an acute avoidance test or an eight-week test; individual wet wt within the size range required for each test method (Section 2.1)
Indices of culture health	—	considered healthy if (1) worms move actively through the substrate, do not try to leave it, and reproduce continuously, and (2) results for reference toxicity tests using worms from the culture fall within historic warning limits; discard culture if >20% of juvenile or adult worms are dead, inactive, or unhealthy at any time

2.3.2 Facilities and Apparatus

Worms must be cultured in a controlled-temperature laboratory facility. Equipment for temperature control (i.e., an incubator or a room with constant temperature) must be adequate to maintain temperature within the required limits (Section 2.3.4). The culturing area must be isolated from any testing, sample storage, or sample-preparation areas, to avoid contamination from these sources. It must be designed and constructed to prevent contamination of cultures (e.g., elimination of copper or galvanized piping or fixtures that could drip metal-contaminated condensates).

All equipment, containers, and accessories that might contact the organisms or substrate within the culturing facility must be clean, rinsed as appropriate, and made of nontoxic materials (e.g., glass, Teflon™, type 316 stainless steel, nylon, Nalgene™, porcelain, polyethylene, polypropylene, fibreglass). Toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must not come in contact with this apparatus and equipment, or the culturing substrate or water.

Various culture chambers, such as plastic trays or breeding boxes of 10–50-L capacity, are suitable for culturing *E. andrei/fetida*. The sides and/or lid should be translucent or transparent, to enable light to contact the surface of the culturing substrate (see Section 2.3.3). Each chamber should have a perforated (e.g., holes covered with fibreglass mesh screening) lid to minimize drying of the surface substrate and the risk of contamination, while allowing air exchange and preventing worms from escaping. Table 2 of Appendix E provides details of the type and size of various chambers recommended by international agencies for culturing this species of earthworm in the laboratory, to generate worms for soil toxicity tests. The use of culture chambers constructed of wood is not recommended, due to the possible presence of toxic contaminants (e.g., plywood glues, antisapstain chemicals, or wood extractives such as resin acids, juvabiones, etc.). The choice of size and numbers of culture chambers required might be influenced by the number of adult earthworms required by the testing facility for one or more series of soil toxicity tests. Each culture container should accommodate a minimum depth of 10 cm of soil or other culturing substrate.

2.3.3 Lighting

Incandescent or fluorescent lights should illuminate the cultures. The *photoperiod* should be regulated (e.g., 16-h light:8-h dark or 12-h light:12-h dark) rather than using continuous (24-h/day) illumination. Light intensity adjacent to the surface of the substrate in culture chambers should range within 400–800 lux. This range is equivalent to a quantal flux of 5.6–11.2 μmol/(m^2 · s) for cool-white fluorescent, 6.4–12.8 μmol/(m^2 · s) for full-spectrum fluorescent, or 7.6–15.2 μmol/(m^2 · s) for incandescent. The lights should be positioned sufficiently far from the culture chambers to prevent evaporation caused by heat buildup.

2.3.4 Temperature

The temperature of the substrate in each culture chamber containing *E. andrei/fetida* should be 20 ± 2 °C as a daily average. Additionally, the instantaneous temperature of this substrate should be 20 ± 3 °C.

2.3.5 Culturing Substrate

Various substrates have been used for culturing *E. andrei/fetida* in preparation for soil toxicity tests (see Table 4 of Appendix E). The choice of substrate for culturing these species is left to the discretion and experience of laboratory personnel; however, the following two culturing substrates are proven and recommended.

A mixture of potting soil (manure/peat/loam blend), *Sphagnum* sp. peat moss, and artificial soil has proven to be a suitable culturing substrate for both *E. andrei* and *E. fetida*. A 10-L batch of this mixture is prepared as follows:

- Mix ~ 3 L of potting soil with ~ 4 L of peat moss (both in their "dry form").
- Then add de-ionized water (~1 L) to the substrate and mix mechanically (handheld mixer) until the moisture content, colour, and *texture* of this mixture appear to be homogeneous.
- Thereafter, add ~1.5 L of artificial soil (see Section 3.2.2).
- Then add de-ionized water (~1 L) to this mixture while stirring mechanically, until a moisture content equivalent to ~70% of the *water-holding capacity* of the mixture is achieved.

- Then measure the soil pH, and, depending on the value, sprinkle ~30 g of calcium carbonate ($CaCO_3$) on the surface of the culturing substrate using a fine sieve, and mix into the soil using a mechanical mixer until no white powder is visible.

This mixture is stored in a covered container at ambient laboratory temperature for three days. The culturing substrate is then stirred, and its pH measured to ensure that it is between 6.0 and 7.5. If the pH is below 6.0, additional calcium carbonate is added (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001).

A substrate comprised of a mixture of shredded un-inked newsprint or shredded brown (unbleached) paper towelling, artificial soil (see Section 3.2.2), *Sphagnum* sp. peat moss, and (depending on availability) worm castings, has been found by Environment Canada laboratory personnel to be suitable for culturing *E. andrei* (EC, 2000a), and is also recommended here. This mixture is prepared by shredding 400–500 g of dry, un-inked paper or brown stock (i.e., un-inked newsprint paper or brown paper towelling) using a single-cut office document shredder, and saturating the shredded paper with dechlorinated municipal drinking water in a ~40-L plastic bin. The bin containing saturated paper is covered and briefly set aside. Approximately 3 kg of artificial soil with a moisture content of ~10% (or ~3 kg of worm castings from another culture) is mixed with 500–600 g of dry *Sphagnum* sp. peat moss in a second bin (~40-L capacity). This mixture is stirred while adding dechlorinated water until it approaches a "mud-like consistency", and its water content is ~80–90% of the mixture's water-holding capacity (D. Moul, personal communication, Environment Canada, Pacific Environmental Science Centre, North Vancouver, BC, 2001). The pH of this soil/peat/water mixture is then measured, and powdered calcium carbonate is added until the pH ranges within 6.0–7.5. The contents of the two bins are then mixed together and placed into a culturing chamber (plastic bin measuring 53 × 38 × 30 cm). Any standing water in the bin is removed and discarded. Thereafter, individuals (~1000 worms) of differing ages and sizes are added. Substrate pH and moisture content are measured periodically (e.g., weekly), and adjusted as necessary. On these occasions, the culturing substrate is gently stirred, and its pH measured to ensure that it is between 6.0 and 7.5. If the pH is below 6.0, additional calcium carbonate is added. Any standing water in the bin is discarded, and test water is mixed in if the culturing substrate appears too dry (D. Moul, personal communication, Environment Canada, Pacific Environmental Science Centre, North Vancouver, BC, 2001).

2.3.6 Food and Feeding

Various types of food and feeding regimes have been used for culturing *E.andrei/fetida* in preparation for soil toxicity tests (see Appendix E; Table 5). Success in culturing these species has been achieved using cooked oatmeal (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1999a, b; Aquaterra Environmental and ESG, 2000) or hydrated alfalfa pellets (USEPA, 1989; ASTM, 1999b; EC, 2000a). Details for preparing these two recommended food types, together with acceptable feeding regimes, follow.

Oats to be used as food for cultures can be purchased from grocery stores. Quaker Oats™ "quick" 3–5 min oatmeal is recommended for this purpose as well as for use in an eight-week test (see Section 4.3.4), since experience with generic or other brands of oatmeal has sometimes indicated problems with respect to excessive mould production in the cultures or during the test. The oatmeal should be hydrated with de-ionized water (one-third volume of dried oatmeal flakes to two-thirds volume of water), cooked in the microwave using the "high" temperature setting, and cooled before feeding to cultures. Using a spoon, one or two teaspoonsful (~5 mL each) of cooked oatmeal should be added to each culture container by making a small depression in the culturing substrate, inserting the food, and then covering it with a thin layer of substrate to minimize mould growth or the proliferation of mite populations (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). During the weekly (or more frequent) checks of the cultures, any old oatmeal and accompanying mould or mites appearing on the surface of the culturing substrate should be removed with a spoon and discarded, taking care not to remove any worms (hatchlings tend to burrow into the oatmeal bolus). Feeding should be once per week.

Alfalfa pellets may also be used as a food source for cultured *E. andrei/fetida* (USEPA, 1989; ASTM,

1999b; EC, 2000a).[8] Dried pellets can be obtained from agricultural feed and supply stores. Before using, the pellets should be saturated with de-ionized or distilled water (at a ratio of ~1 g of dry pellet per 2 mL water). Although ASTM (1999b) recommends that the hydrated alfalfa should be aged for a minimum of two weeks at 4 °C in a covered container, experience at one of Environment Canada's testing laboratories indicates that aging is not necessary and that hydrated alfalfa may be used within hours of hydrating (D. Moul, personal communication, Environment Canada, Pacific Environmental Science Centre, North Vancouver, BC, 2001). The worms in each culture chamber should be fed once per week. At the time of each feeding, any uneaten food observed on the surface of the bedding should be removed with a spoon or forceps, and discarded. Fresh, hydrated alfalfa food is then transferred to the surface of the bedding and covered with a thin layer of bedding substrate to minimize growth of parasites (mites and springtails) (D. Moul, personal communication, Environment Canada, Pacific Environmental Science Centre, North Vancouver, BC, 2001).

The feeding of earthworm cultures with oatmeal or alfalfa should be supplemented with regular additions of small quantities of composted vegetable matter, to improve and sustain the health of the earthworms. Weekly feedings of dehydrated compost (e.g., at a rate of 15–30 mL per culture bin containing ~6–8 L of substrate) should supplement the weekly addition of cooked and hydrated oatmeal or alfalfa. Description of a procedure for preparing this compost, which has been shown to improve the health of cultures of *Eisenia* spp., follows (Stephenson, 2003b).

Vegetable materials (fruits - no banana peels or large pits/stones, bread, vegetables) are placed into a stainless steel pail that is used to collect compostable material from office lunch rooms or households. Once a week, this material is taken to the laboratory and pulverized with a food processor. The addition of water (de-ionized or reverse osmosis) to the food processor containing the compostable material might be required if the material is too dry to pulverize; however, this occurs infrequently. The pulverized material is then placed onto aluminum trays, and distributed to form a thin layer of organic matter (OM). The trays with the OM are placed into a drying oven (90–105 °C) to dry overnight. The next day, the dried OM (with the consistency of dried pabulum) is placed into a food container, and stored in the refrigerator until used (within seven days after preparation). The dehydrated OM is sprinkled weekly on the surface of the substrate in each culture chamber. It is not necessary to rehydrate this material if the culturing substrate is sufficiently moist.

The quantity of food added depends on worm density and developmental stage. The amount of food (cooked oatmeal or hydrated alfalfa) added to each culture chamber should be based on observations and records of food consumed or not consumed, during preceding weekly feedings.

2.3.7 Handling Organisms and Maintaining Cultures

The embryonic (in cocoons), juvenile, and adult life stages of *E. andrei/fetida* should be handled as little as possible, to avoid damage and undue stress. When handling is necessary, it should be done gently, carefully, and quickly to minimize stress to the animals. The use of a gloved hand and/or the arm(s) of rounded forceps are suitable for moving worms to and from culture or test chambers. When handled, any animals that are dropped or injured or appear stressed should be discarded, and must not be used for testing.

Table 6 of Appendix E summarizes useful guidance for maintaining cultures of *E. andrei/fetida.* This information is found in international guides and methodology documents on soil toxicity tests using these earthworms.

In keeping with ASTM (1999b) and EC (2000a), it is recommended that the contents of each culture chamber be inspected just before each weekly feeding, to determine the apparent condition of the worms and the bedding substrate. If, during this inspection, any excess water is observed to have accumulated at the bottom of the substrate, the bedding within the culture chamber should be turned carefully at this time to redistribute the excess water throughout the culturing

[8] This food source has been demonstrated to enable cultures of *E. andrei/fetida* to thrive (USEPA, 1989; ASTM, 1999b; EC, 2000a), and, from a hygienic perspective, is preferable to using animal waste in the laboratory as a food source (as per OECD, 1984, 2000; ISO, 1993, 1998) (see Appendix E; Table 5).

substrate.[9] Care must be taken while turning the chamber's contents, to prevent injuries to the earthworms. Any dead worms observed at these times must be removed and discarded. Any injured or apparently atypical (e.g., lethargic) worms observed should also be removed and discarded. Records should be kept of the apparent condition of the culture (worms and substrate) noted during each observation period (Section 2.3.1).

The loading density of worms in each culture chamber should be restricted to prevent overcrowding and the resulting adverse effects on worm growth, reproduction, and culture health. The maximum loading density of 0.03 g wet wt/cm^3 recommended by ASTM (1999b) (see Appendix E; Table 6, including footnotes 4 and 5 therein) provides useful guidance in this respect. To reduce the population of worms in a crowded culture chamber, either of the following procedures (or some suitable modification thereof) is recommended and should be applied. The first option provides the added advantage of sorting worms into two size classes (i.e., juveniles and adults).

Option 1 (as per ESG International Inc., Guelph, ON) (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001):

- Prepare a fresh mixture of culturing substrate (see Section 2.3.5).
- Thereafter, transfer an aliquot of ~1 L of fresh substrate to each of two temporary holding containers.
- Place the contents of an old (crowded, too wet, or foul smelling) culture chamber onto a plastic sheet or in a shallow plastic container with sufficient surface area to allow the contents to be sorted.
- Remove live and apparently healthy juvenile and adult worms, and transfer them to the two temporary holding containers as two size classes (i.e., juveniles in one container, adults in the other).
- Prepare two new culture chambers, by mixing in a portion of the old substrate and new substrate in a ratio of 1 part old:3 parts new. After mixing, adjust the moisture content and pH of the substrate in each of these culture chambers as required (Section 2.3.5, paragraph 2).
- Transfer the juvenile worms in one of the two temporary holding containers to the surface of the substrate in one of these two culture chambers, and the adult worms to the other.
- Gently distribute individual worms evenly over the surface area, so that they enter the substrate throughout the chamber.
- Label each culture chamber and record the species, life stage, source of worms, approximate number of individuals per chamber, and the date that the substrate was renewed.

Option 2 (as per ASTM 1999b and EC 2000a; see Appendix E, Table 6):

- Set up a culture chamber with new (freshly prepared) substrate but no worms, and place half of its contents onto a plastic sheet.
- Transfer the contents of the crowded culture chamber to a separate plastic sheet.
- Then, carefully remove the worms (including cocoons, juveniles, and adults) from the substrate, and transfer equal numbers and age classes (approximately) temporarily to each of two suitable transfer containers.
- Thereafter, transfer half of the old substrate to the new culture chamber, and mix the contents gently using gloved hands or a spatula or plastic spoon.

---

[9] An alternate approach for redistributing excess water throughout the culturing substrate is to invert culture chambers (with lids in place) each week for a minimum of 1 h. This approach is less labour intensive than turning the substrate in each culture chamber once per week, and may be applied for this purpose. A disadvantage of this procedure is that it does not enable concurrent observations of earthworms that are evident when turning the contents of a culture chamber by hand. If this procedure is followed without turning and the contents of the culturing substrate is not turned manually, gentle stirring of the surface of the substrate on a weekly basis (or more frequently) is recommended, to minimize the proliferation of populations of mites. The use of culture chambers with small mesh-covered holes in the bottom will prevent the buildup of any excess water in the bottom of the chambers.

- Then, mix half of the new substrate on the plastic sheet with the remaining half of the old substrate, and transfer the mixture to the previously crowded culture chamber.

- Thereafter, transfer one of the two groups of worms held briefly in each transfer container to each of the two freshly prepared culture chambers.

Renew the bedding in each culture chamber as required and at least once every 2–3 months, regardless of worm loading densities.[10] An efficient procedure to achieve this (ASTM, 1999b; EC, 2000a) is to prepare a new tray of bedding, and place the contents of the old bedding (including the worms therein) on top of the new bedding. Hold the stacked (old on new) bedding in an uncovered tray at 20 ± 2 °C under continuous illumination for two days, to encourage the worms to burrow into the new bedding. At the end of the two-day period, remove the old bedding from the new bedding and discard it.[11]

Monitor the pH, temperature, and moisture content of the bedding in each culture chamber weekly, and make adjustments as and if necessary (see Sections 2.3.4 and 2.3.5).

2.3.8 Worms for Toxicity Tests

To be successful, the culturing procedures used must produce the required number of healthy test organisms of a known developmental stage, similar age, and similar size. The wet weight of individual *E. andrei/fetida* used to initiate each of the soil toxicity tests described in Sections 4.1, 4.2, or 4.3 must be within the range identified in Section 2.1. Additionally, the cultured organisms must meet specific health and performance-related indices (Section 2.3.9).

Any laboratory-cultured *E. andrei* or *E. fetida* used to start an acute lethality test (including that with a reference toxicant) or an acute avoidance test should be acclimated in the laboratory to the temperature conditions representing those in the toxicity test for a minimum of seven days (see Sections 4.1.2 and 4.2.2). Cultured *Eisenia* spp. to be used in an acute lethality test should also be acclimated during this period to the lighting conditions to be used in the test (Section 4.1.2). Cultured *Eisenia* spp. to be used in an acute avoidance test need not be acclimated beforehand to the conditions of complete darkness that occur throughout this test (see Section 2.4.3 and 4.2.2).

Any laboratory-cultured *E. andrei* used to start an eight-week toxicity test for effects on survival, reproduction, and growth (Section 4.3) must be acclimated in the laboratory to conditions for this toxicity test, for a minimum of seven days. If the culturing substrate used is essentially soil (or a mixture of soil and peat moss; see Section 2.3.5), and the food provided to cultures is the same as that used in the eight-week test (i.e., cooked oatmeal; see Sections 2.3.6 and 4.3.4), then acclimation to these test conditions has been achieved and any additional transfer and handling of worms for this purpose is not advised. However, if culturing conditions of substrate and/or food differ appreciably from those/that to which worms in the negative control soil will be exposed during an eight-week test, all worms to be used in the toxicity test must be held for a minimum of seven days in *negative control soil*. During this acclimation period, lighting and temperature conditions must be the same as those to be used in the eight-week test, and worms must be fed cooked oatmeal (see Sections 2.3.3, 2.3.4, 2.3.6, and 4.3.2).

2.3.9 Health and Performance Indices

Each culture chamber should be checked at least once per week, during which time culture performance should be monitored and recorded (see Sections 2.3.1, 2.3.6, and 2.3.7). Procedures and conditions used to maintain each culture should be evaluated routinely, and adjusted as necessary to maintain or restore the health of the culture. Any juvenile or adult worms that appear to be dead, inactive, not burrowing in the bedding substrate, or otherwise unhealthy or atypical should be discarded. If the culture appears unhealthy or atypical during any weekly (or more frequent) check, it should then be checked daily to make sure that "cascade mortality" (i.e., rate of death increasing exponentially over time) is not occurring. Any worms

---

[10] Signs of deteriorating substrate include differences in colour between the bottom layer and upper few centimetres, and a strong odour indicative of anaerobic conditions (EC, 2000a).

[11] This procedure does not recover the cocoons, and some of the juvenile and adult worms will likely remain in the old bedding (ASTM, 1999b).

on the surface which appear to be dead, inactive, or otherwise unhealthy or atypical should be discarded during these daily observations. If more than 20% of the juvenile or adult *E. andrei/fetida* in a culture chamber appear to be dead, inactive, or unhealthy during any period of observation, the entire group in the container should be discarded. Also, if the combined number of daily mortalities and apparently unhealthy worms observed on the surface of the culturing substrate persists or increases over several days, the contents of the culture chamber should be discarded.

One or more seven-day reference toxicity tests must be conducted using a portion of the population of adult earthworms taken from a particular culture to start a *definitive* soil toxicity test (see Sections 4.1.8 and 4.3.8). Ideally, a reference toxicity test should be performed together with each soil toxicity test. Laboratories routinely undertaking soil toxicity tests using cultured *E. andrei/fetida* may choose instead to conduct one or more routine reference toxicity tests (i.e., at least once each month), using a portion of the adult worms held in the particular culture(s) from which the test organisms for the soil toxicity test(s) are obtained. All tests with the reference toxicant(s) should be performed using the conditions and procedures outlined in Section 4.1.8. Test-related criteria used to judge the validity of a particular soil toxicity test (and, indirectly, the health of the culture), based on the performance of test organisms in the *negative control soil*, are given in Sections 4.1.3, 4.2.3, and 4.3.3.

A laboratory that routinely (e.g., once per month or more) performs eight-week toxicity tests (Section 4.3) might find it useful to monitor the data on number of progeny produced in negative control soil, as a measure of culture health and performance. A plot of such data over time can show problems with respect to reproductive success that are attributable to diet or other conditions to which cultures are exposed (G. Stephenson, personal communication, Stantec Consulting Ltd., Guelph, Ontario, 2004).

2.4 Holding and Acclimating Worms

2.4.1 General

Any group of earthworms (*E. andrei/fetida* or *L. terrestris*) used in one or more of the soil toxicity tests described herein must first be acclimated to the laboratory conditions (temperature and lighting) to which they will be exposed during the test(s). Procedures and conditions for the holding and acclimation (*"rearing"*) of any group of *E. andrei/fetida* or *L. terrestris* transported to the laboratory for their use in an acute lethality test (Section 4.1) or an acute avoidance test (Section 4.2) are described here, and summarized in Table 2. Guidance on sources of earthworms to be delivered to a testing laboratory for use in an acute lethality test or an acute avoidance test is provided in Section 2.2. Refer to Section 2.3 for guidance on conditions and procedures for culturing earthworms (*E. andrei/fetida*) to be used in either of these acute toxicity tests (rather than worms imported for direct testing purposes) or in an eight-week test (restricted to cultured *E. andrei*) for the toxic effects from prolonged exposure to sublethal concentrations of contaminants (Section 4.3).

As with initial tests using cultured *E. andrei/fetida*, it is the responsibility of each laboratory not experienced with the biological test method(s) described in this document to demonstrate its ability to obtain consistent, precise results using a reference toxicant, when initially setting up to perform soil toxicity tests with groups of earthworms (*E. andrei, E. fetida*, or *L. terrestris*) obtained from an outside source. For this purpose, intralaboratory precision, expressed as a coefficient of variation for the respective LC50 data, should be determined by performing five or more tests with different lots (groups) of test organisms from the same source, using the same reference toxicant and identical procedures and conditions for each test (see Section 4.1.8). The laboratory should also confirm its test precision at this time by conducting five or more toxicity tests using *negative control soil* and different lots of test organisms (USEPA, 1994a; EC, 1997a, 1997b, 2001). The conditions and procedures used to perform these initial tests with *negative control soil* should be identical and according to Section 4.1 (if acute lethality tests are intended), Section 4.2 (if acute avoidance tests are intended), or Section 4.3 (if tests for effects of more prolonged exposure on survival, reproduction, and growth are intended).

2.4.2 Facilities and Apparatus

Worms must be held and acclimated in a controlled-temperature laboratory facility isolated from any testing, sample storage, or sample-preparation areas.

Table 2 Checklist of Recommended Conditions and Procedures for Holding and Acclimating *Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris*, to Provide Test Organisms for Use in Soil Toxicity Tests

Source of worms	—	government or private laboratory or a commercial supplier, or collected from grassland known to have not been treated with pesticides or fertilizer for ≥5 years
Life stage and size on receipt	—	depending on timing of toxicity test, may be obtained as juveniles or as sexually mature worms with clitellum; individual wet weight within the indicated range (Section 2.1)
Chamber(s) for holding and acclimation	—	10–50-L "breeding" boxes (e.g., plastic trays measuring ~30 × 40 × 15 cm), covered with perforated lid to allow air exchange and minimize evaporation; sides and/or lid transparent or translucent to enable light to contact surface of substrate; recommended minimum depth, 10 cm
Temperature	—	recommend adjusting gradually (e.g., ≤3 °C/day) for temperature differences upon arrival; thereafter, maintain *Eisenia* sp. at a daily average temperature of 20 ± 2 °C and an instantaneous temperature of 20 ± 3 °C; adjust field-collected *L. terrestris* to a daily average temperature of 20 ± 2 °C for ≥7 days before testing; alternatively, adjust *L. terrestris* to a cooler temperature (e.g., ≤15 ± 2 °C) and hold for several weeks or months at this temperature followed by adjustment to the test temperature over a minimum 6-h period immediately preceding the test
Lighting	—	incandescent or fluorescent; intensity, 400–800 lux at surface of holding/acclimation chamber; fixed photoperiod (e.g., 16 L:8 D or 12 L:12 D); acclimate to these conditions for a minimum seve-day period immediately preceding the test
Type of substrate	—	options include: negative control soil (natural or artificial); a mixture of potting soil, artificial soil, and peat moss; or a mixture of shredded paper (i.e., un-inked newsprint paper or brown paper towelling), artificial soil, and sphagnum peat moss
pH of substrate	—	≥6.0; no adjustment if natural (field-collected) negative control soil; adjusted to range within 6.0–7.5 using reagent-grade calcium carbonate if artificial substrate
Hydration of substrate	—	hydrated with distilled or de-ionized water; moisture content sufficient to keep surface of bedding moist but with no standing water in the bottom of the holding/acclimation chamber; soil particles should not adhere to earthworms
Renewal of substrate	—	as required and at least once every 2–3 months, if worms held for an extended period before use in soil toxicity test; sort and transfer worms manually; alternatively, prepare new tray of bedding, cover with contents of old tray, leave undisturbed under constant light for two days, then remove and discard old bedding
Duration of acclimation	—	for a minimum of seven days during the period immediately preceding the test, to laboratory conditions; earthworms obtained from a commercial supplier should be acclimated to laboratory conditions for a minimum period of 14 days immediately preceding the test
Monitoring substrate quality	—	temperature and moisture content measured at least once per week, in each holding/acclimation chamber
Feeding	—	either cooked oatmeal, or alfalfa pellets saturated with water; feed only cooked oatmeal for ≥7-day period immediately preceding test if acclimating *E. andrei* for use in eight-week test; feed once/week by placing in a shallow depression of the substrate and then covering it with a thin layer of substrate, after removing excess (unused) food and scraping off any visible mould or mites nearby

Weekly maintenance	—	rehydrate as necessary and manually turn substrate in holding/acclimation tray(s), gently, once/week; remove dead, injured, or atypical (lethargic) worms; record apparent condition of substrate and worms; maintain loading density of worms at ≤0.03 g/cm^3
Age/size for test	—	sub-adults or sexually mature adults with clitellum, if to be used in an acute lethality test; clitellated adults only, for an acute avoidance test or an eight-week test; individual wet wt in each instance, within the size range indicated in Section 2.1
Health indices	—	worms in holding/acclimation chamber(s) considered healthy if (1) they appear to be active when observed, and do not try to leave the substrate, and (2) results for reference toxicity tests using worms from the holding/acclimation chamber(s) fall within historic warning limits; discard entire group if >20% of juvenile or adult worms are dead, inactive, or unhealthy at any time

See Section 2.3.2 for further guidance on holding/acclimation facilities and suitable containers (i.e., culture chambers) and lids for holding and acclimating worms to be used in soil toxicity tests.

2.4.3 Lighting

Incandescent or fluorescent lights should illuminate the chamber(s) used to hold and acclimate worms to be used in soil toxicity tests. Photoperiod should be regulated (e.g., 16-h light and 8-h dark; or 12-h light and 12-h dark) rather than continuous (i.e., 24-h illumination). Light intensity adjacent to the surface of the substrate in the holding/acclimation chamber(s) should range within 400–800 lux. This range is equivalent to a quantal flux of 5.6–11.2 μmol/(m^2 · s) for cool-white fluorescent, 6.4–12.8 μmol/(m^2 · s) for full-spectrum fluorescent, or 7.6–15.2 μmol/(m^2 · s) for incandescent. Worms should be acclimated to these lighting conditions for a minimum of seven days before being used in a test.[12]

2.4.4 Temperature

The temperature of the substrate in each holding/acclimation chamber should be 20 ± 2 °C as a daily average, throughout the acclimation period. Additionally, the instantaneous temperature of the substrate should be 20 ± 3 °C throughout this period. An incubator or temperature-controlled room isolated from the testing facility should be used to achieve this.

Upon the receipt of worms at the testing laboratory, the temperature of the substrate within the transport container should be measured and recorded. Worms should be adjusted gradually to the acclimation temperature. For groups of cultured *Eisenia* sp. transferred to a testing laboratory from another laboratory or culturing facility, the temperature of the substrate and worms therein should be adjusted gradually (e.g., ≤3 °C/d) to that at which the cultures are to be held (see Section 2.3.4); these worms should be acclimated to the mean test temperature (i.e., 20 ± 2 °C) for a minimum period of seven days immediately preceding their use in any toxicity test. For field-collected specimens of *L. terrestris* (Section 2.2) to be used in acute lethality tests or acute avoidance tests, worms should be adjusted gradually to the mean test temperature (i.e., 20 ± 2 °C) and held within this temperature range for a minimum period of seven days before they are used in a toxicity test. Alternatively, the field-collected *L. terrestris* may be adjusted gradually to a lower temperature (e.g., 15 ± 2 °C or lower), held at that temperature for an extended period of time (e.g., several weeks or months), and then adjusted to the mean test temperature (i.e., 20 ± 2 °C) before they are used in a toxicity test.[13]

[12] This acclimation period is recommended to enable the adjustment of the worms to test conditions before the toxicity test is initiated. Additionally, it will provide a minimum number of days (≥7)for recovery from any stress due to transfer to the testing laboratory, before they are used in a toxicity test. An exception to acclimating worms to lighting conditions during the test applies to the avoidance test (see Section 4.2), in which instance worms may be acclimated to a fixed photoperiod (e.g., 16-h light:8-h dark or 12-h light:12-h dark) although the test conditions result in complete darkness during the exposure period.

[13] The optimum temperature for holding *L. terrestris* in the laboratory and performing acute toxicity tests with this species is thought to be ≤15 °C (Sheppard, 1994). However, experience indicates that the criterion for test validity using *L. terrestris* in a 14-day lethality test (see Section 4.1.3) or a sublethal test for acute avoidance (see Section 4.2.3) can be readily achieved when these worms are acclimated to and tested at a temperature of 20 ± 2 °C, even if they are held at lower temperatures and then adjusted to the test temperature over a minimum 6-h period (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). Accordingly, and to minimize acclimation requirements while facilitating a single test temperature for each of the earthworm species and tests defined herein, *L. terrestris* may be held at a temperature range below that to be used in a toxicity test for an extended period, provided that

2.4.5 Substrate

Bedding material for holding and acclimating earthworms, in preparation for one or more soil toxicity tests, may be the same substrate as that intended to be used as *negative control soil* in the test(s). This may be either natural, field-collected soil from an uncontaminated site (Section 3.2.1) or artificial soil (Section 3.2.2). Alternatively, the bedding material recommended for culturing *E. andrei/fetida* (Section 2.3.5) may be used for holding and acclimating *E. andrei/fetida* or *L. terrestris*.

The moisture content of the substrate should be sufficient to keep the bedding moist, while not causing water to pool in the bottom of the holding/acclimation chamber. Adjustments for moisture content might be necessary[14] (see Section 2.3.5).

The pH of substrate used to hold and acclimate test organisms should be near neutral, so that it is not stressful to them. Adjustments may be made (see Section 2.3.5), as necessary, to bring the pH of the bedding material into a suitable range (i.e., 6.0–7.5).

2.4.6 Food and Feeding

Worms placed into one or more chambers for holding and acclimating should be fed at that time, using either cooked oatmeal or hydrated alfalfa (see Section 2.3.6). Those held for periods of longer than one week should be fed once per week. Guidance in Section 2.3.6 for food preparation and feeding should be followed. Guidance in Section 2.3.8 should be followed when feeding *E. andrei* during the acclimation period preceding an eight-week test (Section 4.3).

2.4.7 Handling and Maintaining Organisms

Guidance in Section 2.3.7 applies when handling and maintaining worms held in the laboratory before their use in toxicity tests. If the period for holding and acclimation exceeds seven days, the contents of each holding/acclimation chamber should be manually turned just before each weekly feeding. At this time, the apparent condition of the bedding substrate and the worms should be observed and recorded. Any dead, injured, or apparently atypical (e.g., lethargic) worms observed should be removed and discarded.

The loading density of worms in each holding/acclimation chamber should be restricted to prevent overcrowding and the resulting adverse effects on worm condition, performance, and health. The maximum loading density of 0.03 g/cm^3 recommended by ASTM (1999b) for cultures of *E. andrei/fetida* (see Section 2.3.7) should be used as a guide in this respect. If the period for holding earthworms in the laboratory is extended (e.g., several months), and worm density increases during this time, overcrowding should be prevented by splitting the population in a holding/acclimation chamber and adding new bedding material (see Section 2.3.7).

If worms in a holding/acclimation chamber are to be maintained in the laboratory for 2–3 months or longer, the bedding material should be renewed at each 2–3-month interval regardless of worm loading density. Guidance in Section 2.3.7 should be followed to achieve this.

At the start of the holding and acclimation period(s), the temperature, pH, and moisture content of the substrate in each holding/acclimation chamber should be measured and recorded. Weekly measurements of each of these soil quality variables should be made if the acclimation period extends beyond seven days, and adjustments made as and if necessary.

2.4.8 Worms for Toxicity Tests

All earthworms used in a soil toxicity test must appear healthy, and be of similar size. Additionally, they must have been held and acclimated according to the procedures and conditions described herein (Sections 2.4.1 to 2.4.7, inclusive). For *E. andrei/fetida*, each worm to be used in a toxicity test must have a wet weight ranging within that identified in Section 2.1. Each *L. terrestris* used to start a toxicity test must have a wet weight ranging within 3–10 g. Animals used in a toxicity test must satisfy specific health and performance-related indices (Section 2.4.9).

Test-specific acclimation conditions and procedures described in Section 2.3.8 apply when acclimating cultured *Eisenia* spp. for use in an acute lethality or avoidance test, or an eight-week test for effects on survival, reproduction, and growth. Conditions for lighting and temperature applicable to the acclimation of field-collected *L. terrestris* to be used in an acute lethality or avoidance test are as described in Sections 2.4.3 and 2.4.4.

---

they are adjusted within ≥6 h to the test temperature (20 ± 2 °C) before their transfer to test chambers. Alternatively, *L. terrestris* may be acclimated to the test temperature for a more prolonged period following receipt in the laboratory, before their transfer to test chambers. In each instance, the worms should be held at the testing laboratory for a minimum of seven days before testing, to ensure their recovery from the stress of collection and transfer.

[14] If soil particles are observed to be adhering to the worms, the soil is too dry and its moisture content should be increased.

Any *E. andrei* obtained from an outside culture, for initiating an eight-week toxicity test for effects of prolonged exposure on survival, reproduction, and growth (Section 4.3), must be acclimated to the toxicity test conditions in the laboratory for a minimum of seven days. Applicable guidance provided in Sections 2.3.8 and 2.4, must also be followed. If earthworms to be used in an acute or eight-week toxicity test are obtained from a commercial source or supplier, it is strongly recommended that these earthworms be acclimated in the laboratory for a minimum of 14 days preceding their use in the test.

2.4.9 Health and Performance Indices

Each holding/acclimation chamber should be checked at least once per week, during which time the condition of the worms and substrate therein should be monitored and recorded (see Section 2.4.7). Procedures and conditions used to maintain the worms in each holding/acclimation chamber should be evaluated routinely, and adjusted as necessary to optimal levels. Any juvenile or adult worms that appear to be dead, inactive, not burrowing in the bedding substrate, or otherwise unhealthy or atypical, should be discarded. If more than 20% of the juvenile or adult earthworms in a holding/acclimation chamber appear to be dead, inactive, or unhealthy during any period of observation, the entire contents of the container must be discarded.

The sensitivity of each group of acclimated worms used in a definitive soil toxicity test must be measured, using a seven-day reference toxicity test (see Section 4.1.8). Ideally, a seven-day reference toxicity test should be performed together with each definitive soil toxicity test. However, laboratories routinely undertaking soil toxicity tests using healthy *E. andrei/fetida* or *L. terrestris* acclimated to laboratory conditions (according to Sections 2.4.1 to 2.4.8 inclusive) and held in the laboratory for an extended period (months) may choose instead to conduct routine reference toxicity tests (i.e., at least once per month), using a portion of the population(s) of worms held for the definitive soil toxicity tests. All tests with reference toxicant(s) should be performed using the conditions and procedures outlined in Section 4.1.8.

Test-related criteria used to judge the validity of a particular soil toxicity test (and, indirectly, the health of the population of acclimated worms), based on the performance of test organisms in the *negative control soil*, are given in Sections 4.1.3, 4.2.3, and 4.3.3.

Test System

3.1 Facilities and Apparatus

Tests may be performed in an environmental chamber or equivalent facility having acceptable temperature and lighting control (see Sections 4.1.2, 4.2.2, and 4.3.2). The test facility should be well ventilated to prevent personnel from being exposed to harmful fumes, and isolated from physical disturbances or any contaminants that might affect the test organisms. The area used to prepare test soils should also be properly ventilated.

The test facility should be isolated from the area where the worms are cultured (Section 2.3) or held and acclimated (Section 2.4), to avoid potential contamination. Additionally, the test facility should be removed from places where samples are stored or prepared, to prevent the possibility of contamination of test chambers and contents from these sources. The ventilation system should be designed, inspected, and operated to prevent air within the testing facility from contaminating the culturing or holding/acclimation facilities. Return air from sample handling and storage facilities or those where chemicals are processed or tested should not be circulated to the area of the laboratory where tests are conducted.

Any construction materials that might contact the organisms, water, or test chambers within this facility must be nontoxic (see Section 2.3.2) and should minimize sorption of chemicals. Borosilicate glass, nylon, high-density polyethylene, high-density polystyrene, polycarbonate, fluorocarbon plastics, and type 316 stainless steel should be used whenever possible to minimize chemical sorption and leaching.

The test facility must have the basic instruments required to monitor the quality (e.g., temperature and pH) of the *test soil* and associated *test (hydration) water*. Additionally, the laboratory should be equipped to facilitate prompt and accurate analysis of the moisture content of test soils. Equipment requirements include a drying oven which can be set at 90 °C for drying test organisms and 105 °C for drying soils, a weighing balance accurate to the nearest 0.1 mg, and a pH meter. Safety apparatus, including a respirator with dust protection, gloves, laboratory clothing, and glasses for eye protection, are required when preparing mixtures and aliquots of test soil.

All test chambers, equipment, and supplies that might contact site soils, test soils, test (hydration) water, *stock solutions*, or test solutions must be clean and rinsed with de-ionized or distilled water (i.e., *test water*), before being used. All nondisposable materials should be washed after use. The following cleaning procedure is recommended (EC, 1997a, b, 2001).

1. soak in tap water (with or without detergent added) for 15 minutes, then scrub with detergent or clean in an automatic dishwater;

2. rinse twice with tap water;

3. rinse carefully with fresh, dilute (10%, v:v[15]) nitric (HNO_3) or hydrochloric acid (HCl) (metal-free grade) to remove scale, metals, and bases;

4. rinse twice with de-ionized water (or other test water);

5. rinse once with full-strength, pesticide-grade acetone to remove organic compounds and with reagent-grade (e.g., HPLC grade, ≥98.5% purity) hexane for oily residues (use a fume hood or canopy);[16]

6. allow organic solvent to volatilize from dishware in fume hood and rewash with detergent (scrub if necessary); and

7. rinse three times with de-ionized water (or other test water).

Test chambers and apparatus that might contact soil or test (hydration) water should be thoroughly rinsed with test water, immediately before being used in the test.

3.1.1 Initial Tests

Before definitive soil toxicity tests using any of the test methods defined in Sections 4.1, 4.2, or 4.3 are performed for the first time by a testing laboratory, it is recommended that a minimum of five control performance tests with one or more samples of uncontaminated natural or artificial soil intended (or under consideration) for use in one or more definitive soil toxicity tests as *negative control soil* (see Section 3.2) be undertaken by laboratory personnel. Additionally, a minimum of five reference toxicity

[15] To prepare a 10% solution of acid, carefully add 10 mL of concentrated acid to 90 mL of de-ionized water.

[16] Rinsing Plexiglas™ with acetone or hexane is not recommended, since the Plexiglas™ can become pitted and etched by these solvents and can turn from transparent to opaque.

tests using one or more samples of a candidate artificial or natural negative control soil intended for routine use in conjunction with definitive soil toxicity tests (see Sections 4.1.8, 4.2.7, and 4.3.8), should be undertaken by laboratory personnel. These initial tests are recommended to confirm that acceptable performance of the test species (*E. andrei, E. fetida*, or *L. terrestris*) can be achieved in a candidate natural or artificial negative control soil using that laboratory and the culturing or holding/acclimation conditions and procedures specified in this report (see Sections 2.3 and 2.4).

The conditions and procedures used to perform these initial tests with negative control soil should be identical and according to Section 4.1 (if acute lethality tests are intended), Section 4.2 (if acute avoidance tests are intended), or Section 4.3 (if tests for effects of more prolonged exposure on survival, reproduction, and growth are intended). The conditions and procedures used to perform these initial tests with one or more reference toxicants should be identical and according to Section 4.1.8. Each test with *negative control soil* or *reference toxicant(s)* should be performed using a different lot of test organisms of the same species from the same source.

Data from the control performance tests ($\geq$5) must show that the criterion or criteria for test validity (see Sections 4.1.3, 4.2.3, and 4.3.3) can be met for the intended test species using a natural or artificial soil intended for use as negative control soil in a definitive soil toxicity test. Data from the initial reference toxicity tests ($\geq$5) should be compared by calculating and appraising the magnitude of the coefficient of variation (CV) for the respective series of tests and endpoint values.

3.1.2 Acute Lethality Test

Wide-mouthed glass jars (e.g., mason canning jars), with a 500-mL capacity, are to be used as test chambers. Each glass jar must be cleaned thoroughly before and after use, and rinsed well with de-ionized or other test water immediately before use. Each test chamber should be covered with a new piece of thin transparent or translucent material (e.g., Saran wrap™ or Parafilm™) , that is perforated with $\geq$5 small (e.g., ~1–2 mm) holes (to minimize evaporation and allow air exchange) and secured to the lip of each jar using a rubber band.

3.1.3 Acute Avoidance Test

The recommended test apparatus for performing an acute avoidance test with earthworms is illustrated in Figure 2.[17] The design of each test unit consists of a circular container with an outer diameter of ~230 mm. Each test unit is partitioned into a central cylinder with an inner diameter of ~54 mm, and six pie-shaped interconnecting compartments each with a capacity for ~350 mL of soil. A series of 1-cm holes drilled in the bottom of the central chamber (two per compartment) and on the sides of each pie-shaped compartment (three per side) enable the free movement of earthworms from the central cylinder (devoid of substrate) to the test compartments containing test soils, and free movement of test organisms between the compartments. A set of six removable side partitions, made of rigid steel sheeting (see Figure 2 for illustration and dimensions) is required for insertion alongside each of the six walls separating compartments at the end of the test (Section 4.2.6).

The apparatus is constructed of high quality stainless steel sheeting (1–4 mm thick) or Plexiglas™ sheeting (5–6 mm thick), and includes a removable lid (also made of stainless steel or Plexiglas) which does not seal and enables an exchange of air within the test compartments. Avoidance apparatus constructed of stainless steel is recommended when testing soils contaminated or spiked with organic compounds (particularly petroleum products), since this material sorbs fewer organics than Plexiglas and can be rinsed with acetone and/or hexane without damaging it. Apparatus constructed of Plexiglas is recommended when testing soil contaminated with heavy metals.

A minimum of five test units are required for each single-concentration or multi-concentration soil toxicity test, and more units (e.g., 6–10) are recommended for each multi-concentration test (see Section 4.2.1). Each test unit must be cleaned thoroughly before and after use, and rinsed well with de-ionized or other test water immediately before use.

---

[17] The experimental apparatus used in this test is called the "Kaushik chamber", after the professor responsible for its design (N. Kaushik, personal communication, Department of Environmental Biology, University of Guelph, Guelph, ON, 1995). The prototype was used in the early 1960s to investigate the preference of aquatic oligochaetes to sediment with different grain size characteristics. The design was modified to accommodate the larger terrestrial earthworm species.

The dimensions for the apparatus depicted in Figure 2 represent a test unit constructed of stainless steel; similar dimensions apply when constructing the avoidance apparatus using Plexiglas™.

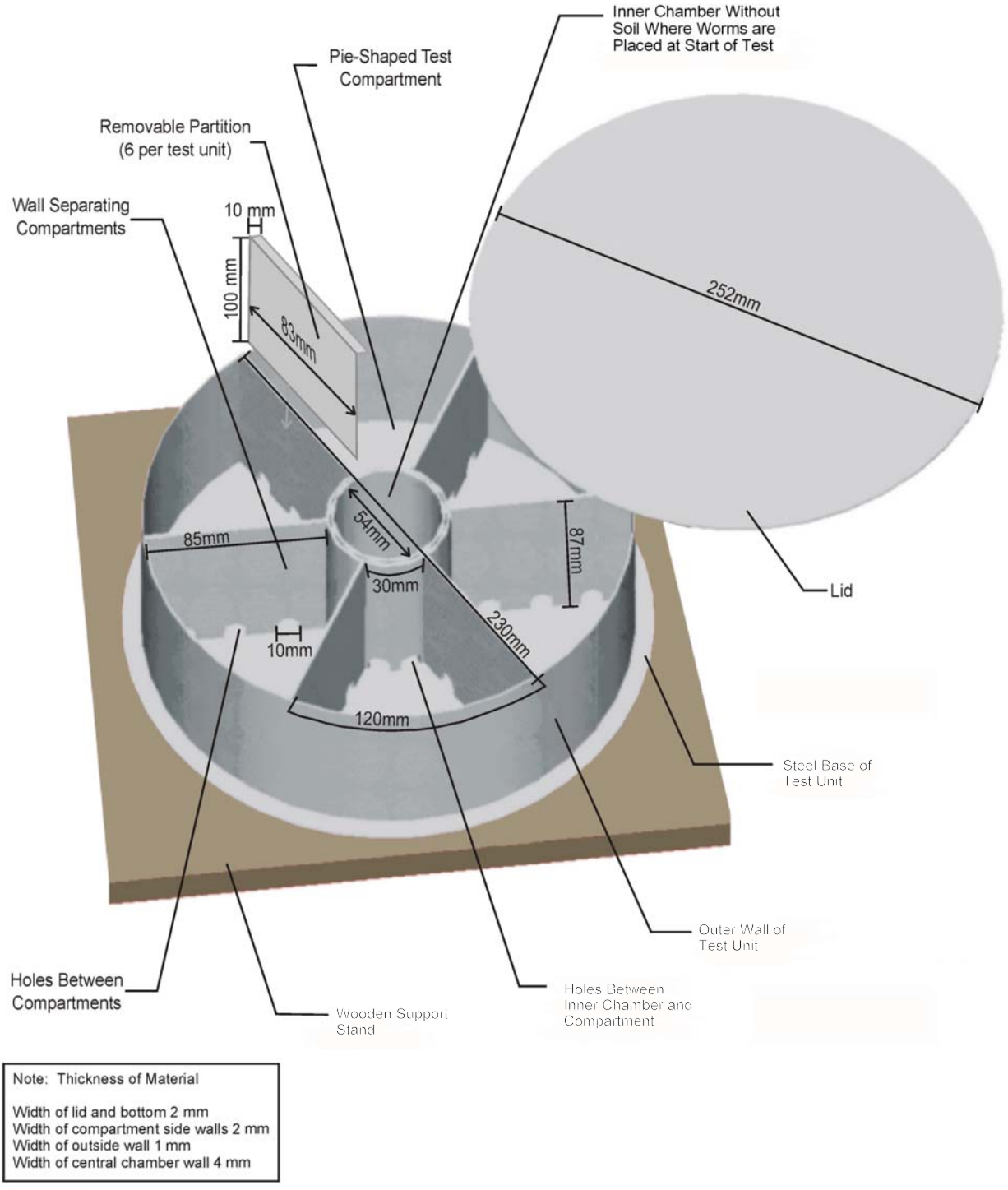


Figure 2 Recommended Design of Test Unit for Performing an Acute Avoidance Test Using Earthworms (*E. andrei, E. fetida,* or *L. terrestris*) and *Clean* or Contaminated Soil

3.1.4 Test for Effects of Prolonged Exposure on Survival, Reproduction, and Growth

Wide-mouthed glass jars (e.g., mason canning jars), with a 500-mL capacity, are to be used as test chambers. Guidance on cleaning each jar before and after use, and on their covers, is described in Section 3.1.2.

3.2 Negative Control Soil

Each soil toxicity test must include *negative control soil* as one of the experimental *treatments*. *Negative control soil* is essentially free of any contaminants that could adversely affect the performance of earthworms during the test. The use of *negative control soil* provides a measure of test acceptability, evidence of the health and performance of the test organisms, assurance as to the suitability of the test conditions and procedures, and a basis for interpreting data derived from the test soils.

A soil toxicity test may use *clean* (uncontaminated) natural soil and/or artificial soil as the negative control soil. The selection of an appropriate negative control soil depends on considerations such as the study design, physicochemical characteristics of the test soil(s), and the availability of suitable *clean* natural soil with acceptable properties.[18] There should also be prior experimental evidence that the soil chosen for use as negative control soil will consistently and reliably meet the criteria for test validity defined herein for each test method (Sections 4.1.3, 4.2.3, and 4.3.3).

The biological test methods described herein have been developed and tested using five negative control soils with diverse physicochemical characteristics (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1999a, b, 2000a; Aquaterra Environmental and ESG, 2000; ESG 2000, 2001, 2002; and ESG and Aquaterra, 2002). These *clean* soils included one artificial soil and four natural soils (i.e., samples of sandy loam and silt loam agricultural soils from southern Ontario, a clay loam prairie soil from Alberta, and a forest loam soil from the Canadian Shield in northern Ontario). These soils differed in composition with respect to the physicochemical characteristics that could potentially influence the fate and effects of contaminants. All of the field-collected soils originated from uncontaminated areas that had not been subjected to any direct application of pesticides in recent previous years and were therefore considered to be "clean". The origin and physicochemical characteristics of these natural soils are further described in Appendix H. The test validity criteria for *Eisenia* spp. or *L. terrestris* described in Sections 4.1.3, 4.2.3, and 4.3.3 are based on the performance data for these earthworms in negative control soil, that were generated for each of these five diverse soils (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1999a, b, 2000a; Aquaterra Environmental and ESG, 2000; ESG 2000, 2001, 2002; and ESG and Aquaterra, 2002).

3.2.1 Natural Soil

Negative control soil may be natural soil collected from a *clean* (uncontaminated) site known to have been free of pesticide or fertilizer applications for at least five years. The source of this *negative control soil* might be the same as that where earthworms were collected to establish a culture or to obtain test organisms (Section 2.2). Before using a sample of *clean* field-collected soil as *negative control soil* in a definitive toxicity test, the test laboratory must have previous experimental evidence showing that natural soil from this source can meet the criteria that must be achieved for the results of a toxicity test to be considered valid (see Sections 4.1.3, 4.2.3, and 4.3.3).

Accordingly, preliminary tests involving a sample of this soil must be performed using earthworms of the intended test species, to confirm that test organisms of this species are able to meet the criterion or criteria for test validity that apply to the particular test method(s) to be undertaken. Thereafter, and assuming that the preceding results for these preliminary bioassays are satisfactory, it is recommended that samples of natural soil selected for possible use as *negative control soil* in soil toxicity tests (as well as samples of candidate *reference soil*) be analyzed for the following

---

[18] The Canadian Council of Ministers of the Environment (CCME) provides a comprehensive Web site on Canadian Environmental Quality Guidelines including those for soil www.ccme.ca. This information is useful when reviewing analytical data (e.g., values for metals or PAHs) for samples of field-collected soil from a location under consideration as a source of natural soil suitable for use as negative control soil in toxicity tests. The summary table of CCME's Environmental Quality Guidelines can be accessed directly at www.ccme.ca/assets/pdf/e1_06.pdf. These Web sites and associated links will assist the investigator(s) reviewing the physicochemical characteristics of presumably *clean* natural soils under consideration for use as negative control soil in soil toxicity tests. The CCME can also be contacted by toll-free phone (1-800-805-3025) or e-mail (info@ccme.ca).

physicochemical characteristics:

– pH,
– particle size distribution,
– *conductivity*,
– *texture*,
– *fertility*,
– *total organic carbon* content (%),
– *organic matter* content (%),
– cation exchange capacity,
– major cations,
– total nitrogen,
– total phosphorus,
– bulk density,
– WHC,
– metals,
– petroleum hydrocarbons (including PAHs),
– organophosphorus insecticides,
– organochlorine insecticides, and
– a suite of herbicides.

Pesticide and metal concentrations should not exceed the Canadian Soil Quality Guidelines, if available (see footnote 18). If the results of both the preliminary biological tests and the physicochemical analyses are satisfactory, a larger sample of this natural soil can be collected, air dried to a moisture content of between 10 and 20%, coarse-screened (4–6 mm), transferred to clean plastic pails, and stored in darkness at 4 ± 2 °C until required.

3.2.2 Artificial Soil

Negative control soil may be artificial soil formulated in the laboratory. The use of artificial soil offers a consistent, standardized approach and is advantageous when testing the toxicity of chemicals or chemical products spiked in *negative control soil* (Section 6).

The formulation of artificial soil used internationally in various soil toxicity tests with earthworms is very similar. Appendix F (Tables 3, 4, and 6) provides a summary of the ingredients and preparation of artificial soil recommended by various agencies (OECD, 1984; USEPA, 1989; ISO, 1993, ASTM, 1999b; EC, 2000b) for use as *negative control soil* in 14-day lethality tests with earthworms.

Appendix G (Tables 4, 5, and 6) summarizes the formulation and preparation of artificial soil recommended for use as *negative control soil* by ISO (1991, 1998) and OECD (2000) in laboratory tests of the effects of contaminated soil on the reproduction of earthworms.

In keeping with the formulation of artificial soil used by OECD (1984, 2000), USEPA (1989), ISO (1991, 1993, 1998) and ASTM (1999b), the following ingredients should be used to prepare artificial soil to be used in the biological test methods described herein:

- 10% *Sphagnum* sp. peat, air dried and sieved through a 2-mm mesh screen (ASTM, 1999b)

- 20% kaolin clay with particles <40 µm (ASTM, 1999b)

- 70% "grade 70" silica sand (USEPA, 1989; ASTM, 1999b)

The ingredients should be mixed thoroughly in their dry form using a mechanical stirrer and/or gloved hands.[19] Reagent-grade calcium carbonate should be added to the dry mixture in a quantity sufficient to attain a pH for the artificial soil ranging within 6.0–7.5 once it is hydrated.[20] Thereafter, the mixture should be

---

[19] It is recommended that the dry ingredients initially be mixed (to incorporate the calcium carbonate) using a mechanical stirrer. Mixing should be completed using a gloved hand, to ensure that all of the soil from the corners of the container have been well mixed. Personnel must take the appropriate precautions for protection to prevent the inhalation of and contact with these ingredients.

[20] The amount of calcium carbonate ($CaCO_3$) required to adjust the pH of artificial soil to within this range depends on the nature (i.e., acidity) of the ingredients (and, in particular, that of the *Sphagnum* sp. peat). A quantity of 10–30 g of $CaCO_3$ for each kg of peat might prove adequate. A pH as low as 4.5 can occur when the soil is first formulated without the addition of $CaCO_3$. The initial pH adjustment should attempt to raise pH to range within 7.0–7.5, since the pH of artificial soil typically drops slightly (to 6.5–7.0) during the three-day equilibration period, before it stabilizes. The pH of stored samples of artificial soil should be checked regularly (e.g., once every two weeks) to ensure that it has not changed dramatically; adjustments should be made as necessary by adding additional quantities of $CaCO_3$ (Aquaterra Environmental, 1998a; G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON , 2001).

A mixture of formulated artificial soil can also be stored dry, followed by partial hydration to ~20% moisture content, storage at 20 ± 2 °C for a minimum 3-day period, and subsequent hydration to ~70% WHC when required

hydrated gradually using *test water* (i.e., de-ionized or distilled water) until its moisture content is ~20% (which is ~28% of the soil's *water-holding capacity*), while mixing further until the soil is visibly uniform in colour and texture. As necessary, reagent-grade calcium carbonate should be added to the hydrated mixture in a quantity sufficient to maintain a pH ranging within 6.0–7.5. Samples of pH-adjusted artificial soil should be stored in darkness at 20 ± 2 °C for a minimum of three days before being used in a toxicity test, to enable adequate time for pH equilibration.[20] Thereafter, artificial soil can be stored at 4 ± 2 °C. As and when required for a soil toxicity test, a suitable quantity of stored artificial soil should be hydrated further using test water until its moisture content is ~70% of the water-holding capacity.

3.3 Positive Control Soil

The use of one or more samples of *positive control soil* is recommended for inclusion in each series of soil toxicity tests with earthworms, to assist in interpreting the test results. In choosing a positive control soil, the intent is to select a toxic soil that will elicit a response in the test organisms which is predictable based on earlier toxicity tests with this material. The positive control soil might be a sample of negative control soil that is spiked with a reference toxicant for which historic data are available on its toxicity to earthworms using specified test conditions and procedures. For each of the three biological test methods described herein, one or more reference toxicants must be used as a *positive control* soil when appraising the sensitivity of the test organisms and the precision and reliability of results obtained by the laboratory for that material (see Sections 4.1.8, 4.2.8, and 4.3.8). A test might also include a sample of negative control soil (natural or artificial; see Section 3.2) that has been spiked experimentally (Section 6) with one or more toxic chemicals or chemical products of particular concern when evaluating the sample(s) of test soil, at a concentration toxic to earthworms according to the biological test method to be used. In some instances, a test might include a positive control soil that is comprised of a highly contaminated sample of field-collected soil or sludge shown previously to be consistently toxic to earthworms according to the biological test method to be used.

3.4 Reference Soil

One or more samples of *reference soil* might be included in a soil toxicity test using earthworms.[21] The type and nature of the sample(s) of soil used as reference soil in a particular study depend on the experimental design and the study's objectives. If the toxicity of samples of field-collected soil from a contaminated or potentially contaminated site is under investigation, the reference soil included in the study might be one or more samples of field-collected soil taken from a *clean* (uncontaminated) site where the physicochemical properties (e.g., organic carbon content, organic matter content, particle size distribution, texture, pH) represent the sample(s) of test (contaminated) soil as much as possible. Ideally, the reference soil is collected near the site(s) where samples of test soil are collected, but is removed from the source(s) of contamination. One or more samples of field-collected *clean* reference soil from sites removed from the test site(s) might also be chosen due to their known lack of toxicity in previous tests with earthworms, and their possession of physicochemical characteristics similar to the samples of test soil. The sample(s) of field-collected reference soil used in a study could be tested for toxic effects at full strength only, or this soil could be mixed with the sample(s) of test soil to prepare a range of concentrations to be included in a multi-concentration

---

for use in a toxicity test. If storing formulated artificial soil dry, it is necessary to partially hydrate (to ~20% moisture) and equilibrate thereafter (for ≥3 days) to provide conditions for pH equilibrium similar to that recommended herein using artificial soil stored partially hydrated. Using this optional approach, the interim storage as partially hydrated artificial soil is necessary to enable the addition of more water (and, in certain instances, the addition of a chemical solution) as required when finalizing the pH and moisture content (i.e., adjusted to ~70% WHC) of artificial test soil. Storage of artificial soil that is partially hydrated, rather than dry, is considered a preferred approach since it enables laboratory personnel to more quickly hydrate to the desired moisture content (i.e., ~70% WHC) while ensuring pH equilibrium, and reduces any further delay in time associated with the dry storage of artificial soil.

[21] The use of field-collected *reference soil* might not be appropriate for certain toxicity tests such as those using samples of sludge (Section 5) or chemical-spiked soil (Section 6). Table 4 of Appendix F summarizes the applications of reference soil described in certain other test-method documents.

test[22] (see Sections 3.5 and 5.5, as well as the introductory comments in Section 4). Samples of reference soil should not be collected from sites known to have received applications of pesticides or fertilizers within the past five years or more.

An investigator might choose to include one or more samples of artificial soil as *reference soil* in a particular test. For instance, these samples could be used in multi-concentration tests with site soils or chemical-spiked soils to investigate the influence of certain physicochemical characteristics (e.g., a number of artificial reference soils prepared to provide a range of differing values for texture and/or percent organic matter content; Sheppard and Evenden, 1998; Stephenson *et al.*, 2002) on the toxicity of a contaminated site soil or a chemical-spiked soil. Multiple samples of *clean* field-collected soil collected from various sites, which differ markedly with respect to one or more physicochemical characteristics, might also be used for this purpose. For such a study, a portion of each reference soil used to prepare a series of concentrations of the the test soil should be included in the test without dilution (i.e., 100% reference soil).

Each test involving one or more samples of reference soil must include a sample of *negative control soil* (see Section 3.2). Conversely, certain tests (e.g., one involving a series of concentrations of chemical-spiked soil prepared using artificial or natural negative control soil) need not involve a sample of reference soil. For tests with field-collected site soil, the inclusion of one or more samples of reference soil from a neighbouring site is a preferred approach for comparative purposes (see Section 5.5); a decision to dilute site soil with reference soil (rather than negative control soil) when preparing multiple concentrations for testing depends on the study objectives.

3.5 Test Soil

These biological test methods are intended to measure the toxicity of one or more samples or mixtures of contaminated or potentially contaminated soil (*test soil*), using earthworms as test organisms. The sample(s) of test soil might be either field-collected soil from an industrial or other site of concern, or industrial or municipal biosolids (e.g., dredged material, municipal sludge from a sewage treatment plant, composted material, or manure) under consideration for possible land disposal. A sample of field-collected test soil might be tested at a single concentration (typically, 100%) or evaluated for toxicity in a multi-concentration test whereby a series of concentrations are prepared by mixing measured quantities with either *negative control soil* or *reference soil* (see Section 5). The test soil might also be one or more concentrations of a chemical-spiked soil, prepared in the laboratory by mixing one or more chemicals or chemical products with *negative control soil*, *reference soil*, or *site soil* (see Section 6).

---

[22] Alternatively, the series of test concentrations used in a multi-concentration test could be prepared using negative control soil. The choice might be influenced by whether or not the candidate reference soils are known to likely be non-toxic in the test to which they are to be applied, or a desire to prepare a range of concentrations of test soil using a *clean* soil with characteristics (e.g., texture, organic matter content) that closely match that of the test soil.

Universal Test Procedures

General procedures and conditions described herein for each of the three biological test methods with earthworms apply when testing the toxicity of samples of soil, particulate waste (e.g., sludge), or chemical, and also apply to their associated reference toxicity tests. More specific procedures for conducting tests with field-collected samples of soil or other similar particulate material (e.g., sludge, de-watered mine tailings, drilling mud residue, compost, biosolids) are provided in Section 5. Guidance and specific procedures for conducting tests with *negative control soil* or other soils spiked (amended) experimentally with chemical(s) or chemical product(s) are given in Section 6.

All aspects of the test system described in Section 3 must be incorporated into these universal test procedures. Those conditions and procedures described in Section 2 for culturing *E. andrei/fetida*, or for holding and acclimating *E. andrei/fetida* or *L. terrestris* in preparation for soil toxicity tests, also apply.

For any multi-concentration test to be performed according to the test procedures described herein (see Sections 4.1, 4.2, and 4.3), concentrations should be chosen to span a wide range, including a low concentration that obtains effects like the negative control, and a high concentration that results in "complete" or severe effects. It is a common mistake to anticipate the endpoint and bracket it with a closely spaced series of concentrations, all of which turn out to be either too low or too high. To keep the wide range of concentrations, and also obtain the important mid-range effects, it might be necessary to use additional treatments to split the selected range more finely. In any case, a consistent geometric series should be used. Additional guidance on selecting test concentrations that applies here is found in EC (2004b).

4.1 Acute Lethality Test

Table 3 provides a summary checklist of recommended conditions and procedures to be universally applied to each acute lethality test with samples of contaminated or potentially contaminated soil, as well as those for testing specific types of test materials or substances.

These could include samples of site soil, biosolids (e.g., dredged material, sludge from a sewage treatment plant, composted material, or manure), or *negative control soil* (or other soil, contaminated or clean) spiked in the laboratory with one or more test chemicals or chemical products.

This biological test method uses either *sub-adult* or *adul*t earthworms within a specified weight range (see Section 4.1.1) as test organisms, and measures survival (mortality) as the biological endpoint. Test organisms are either laboratory-cultured *E. andrei/fetida* (see Section 2.3) or worms (*E. andrei/fetida* or *L. terrestris*) obtained from an outside source (Section 2.2) and acclimated to laboratory conditions (Section 2.4) before their use in the test. Test duration is 14 days. The test organisms are not fed during the test, and the test soils are not renewed. This 14-day test method was applied and validated by seven participating laboratories, in a series of concurrent multi-concentration tests using artificial soil spiked with boric acid (EC, 2004a).[23]

4.1.1 Beginning the Test

Each test chamber (see Section 3.1.2) placed within the test facility must be clearly coded or labelled to enable identification of the sample and (if diluted) its concentration. The date and time when the test is started must be recorded, either directly on the labels or on separate data sheets dedicated to the test. The test chambers should be positioned for ease while making observations and measurements. Treatments

[23] In this series of tests, each of the participating laboratories was able to achieve valid test results (i.e., 14-day survival rates in negative control soil were ≥90%). The mean 14-day LC50 for boric acid spiked in artificial soil was 3524 mg H_3BO_3/kg soil dry wt, with values for individual laboratories ranging from 2228–4677 mg/kg. The interlaboratory coefficient of variation for these LC50s was 25%, which is considered to be an acceptable level of precision between laboratories (EC, 2004a).

Table 3 Checklist of Recommended Conditions and Procedures for Conducting 14-Day Lethality Tests for Soil Toxicity Using Earthworms (*Eisenia andrei, Eisenia fetida*, or *Lumbricus terrestris*)

Universal		
Test type	—	whole soil toxicity test; no renewal (static test)
Test duration	—	14 days
Test organisms	—	*E. andrei, E. fetida,* or *L. terrestris*; juveniles (i.e., sub-adults) or sexually mature adults with clitellum; if *E. andrei/fetida*, individual wet wt 250–600 mg; if *L. terrestris*, individual wet wt 3–10 g ; 5 worms/test chamber if *E. andrei/fetida* and 3 worms/test chamber if *L. terrestris*
Negative control soil	—	depends on study design and objectives; *clean* field-collected soil or artificial soil if testing site soils; recommend artificial soil for tests with chemicals or chemical products spiked in soil
Test chamber	—	500-mL glass jar; perforated translucent or transparent cover (e.g., Saran wrap™ or Parafilm™), secured with a rubber band recommended as cover
Amount of soil/ test chamber	—	identical wet wt, equivalent to a volume of ~350 mL; ~200 g dry wt if artificial soil
Moisture content, test soils	—	hydrate to the optimal percentage of its water-holding capacity (WHC) if field-collected soil (see Section 5.3), or to ~70% of WHC if artificial soil
Number of replicates	—	≥5 replicates/treatment if single-concentration test (e.g., site soil tested at 100% concentration only); ≥3 replicates/treatment if multi-concentration test using *E. andrei/fetida*; ≥5 replicates/treatment if multi-concentration test using *L. terrestris*
Temperature	—	daily average, 20 ± 2 °C; instantaneous, 20 ± 3 °C
Lighting	—	incandescent or fluorescent; intensity, 400–800 lux adjacent to surface of soil in test chamber; fixed photoperiod (e.g.,16L:8D or 12L:12D)
Feeding	—	do not feed
Measurements during test	—	soil moisture content and pH in each treatment/concentration, at start and end; temperature in test facility, daily or continuously
Observations during test	—	number of live worms in each test chamber on Days 0, 7 (optional), and 14; number of worms seen on surface of soil in each test chamber on Days 0 (i.e., at 1 h), 7 (optional), and 14; obvious pathological symptoms (e.g., open wounds, discolourations, segmental swellings) or distinct behavioural abnormalities (e.g., lethargy, coiling, non-burrowing) for worms in each test chamber on Days 0, 7 (optional), and 14
Biological endpoint	—	number of live worms in each replicate (i.e., in each test chamber) on Day 7 (optional) and 14
Statistical endpoints	—	percent survival, each test chamber and each treatment/concentration, on Days 7 (optional) and 14; 7-day LC50 (optional) and 14-day LC50 if multi-concentration test

Test validity	—	invalid if mean 14-day survival in negative control soil <90%
Test with reference toxicant	—	perform once/month, or in conjunction with definitive test(s) with soil samples; use boric acid; prepare and test ≥5 concentrations plus a negative control, using artificial soil as substrate; ≥3 replicates/concentration and 5 worms/replicate if *E. andrei/fetida*; ≥5 replicates/concentration and 3 worms/replicate if *L. terrestris*; follow procedures and conditions for a multi-concentration acute lethality test for soil toxicity described herein; determine 7-day LC50 (including 95% confidence limits); express as mg boric acid/kg, dry wt

Field-collected Soil

Transport and storage	—	seal in plastic and minimize air space; transport in darkness (e.g., using an opaque cooler, plastic pail, or other light-tight container); do not freeze or overheat during transport; store in dark at 4 ± 2 °C; test should start within two weeks, and must start within six weeks unless soil contaminants are known to be stable
Negative control soil	—	either natural, uncontaminated field-collected soil or artificial soil, for which previous 14-day tests with the test species showed a mean survival rate of ≥90%
Reference soil	—	one or more samples for tests with field-collected soil; ideally, taken from site(s) presumed to be clean but near sites of test soil collection; characteristics including percent organic matter, particle size distribution, and pH similar to test soil(s)
Characterization of test soils	—	at least percent moisture content, WHC, pH, conductivity, percent total organic carbon (TOC), percent organic matter, and particle sizes (% sand, % silt, % clay); optionally, contaminants of concern [e.g., metals, polycyclic aromatic hydrocarbons (PAHs), pesticides]
Preparation of test soils	—	if necessary, remove debris and indigenous macro-organisms using forceps; if necessary, press through a sieve of suitable mesh size (e.g., 4–6 mm); mix; determine soil moisture content; hydrate with de-ionized or distilled water (or, if and as necessary, dehydrate) to the optimal percentage of its WHC (see Section 5.3); mix; dilute with negative control soil or reference soil if multi-concentration test; ensure homogeneity

Soil Spiked with Chemical(s) or Chemical Product(s)

Negative control soil	—	recommend artificial soil or a *clean* field-collected soil
Characterization of chemical(s) or chemical product(s)	—	information on stability, water solubility, vapour pressure, purity, and biodegradability of chemical(s) or chemical product(s) spiked into negative control soil should be known beforehand
Solvent	—	de-ionized water is the preferred solvent; if an organic solvent is used, the test must include a solvent control
Preparation of mixtures	—	procedure depends on the nature of the test substance(s) and the test design and objectives; chemical/soil mixtures may be prepared manually or by mechanical agitation; test substance(s) may be added as measured quantities in solution (i.e., in water or an organic solvent) or as a solid material comprised partly or completely of the test substance(s); ensure homogeneity
Concentration within soil mixture of chemical(s) or chemical product(s) added	—	normally measure at beginning and end of test, in high, medium, and low strengths as a minimum

should be positioned randomly within the test facility (EC, 1997a, b, 2001).

The day that animals are initially exposed to samples of test materials or substances is designated Day 0. On the day preceding the start of the test (i.e., Day -1), each sample or subsample of test soil or similar particulate material, including *negative control soil* and, if used, *reference soil*, should be mixed thoroughly[24] (see Sections 5.3 and 6.2) to provide a homogeneous mixture consistent in colour, texture, and moisture. If field-collected samples of *site soil* are being prepared for testing, large particles (stones, thatch, sticks, debris) should be removed before mixing, along with any vegetation or macroinvertebrates observed (see Section 5.3).

The quantity of each test soil mixed as a *batch* should be enough to set up the *replicates* of that treatment (see Table 3) plus an additional amount for the physicochemical analyses to be performed (Section 4.1.5) and a surplus to account for that unused portion of soil that adheres to the sides of the mixing chamber after removing aliquots for the testing and analyses. The moisture content (%) of each test soil should be known or determined, and adjustments made as necessary by mixing in *test water* (or, if and as necessary, by dehydrating the sample) until the desired moisture level is achieved (see Sections 5.3 and 6.2). Quantitative measures of the homogeneity of a *batch* might be made by taking aliquots of the mixture for measurements such as particle size analysis, total organic carbon content (%), organic matter content (%), moisture content (%), and concentration of one or more specific chemicals.

Immediately following the mixing of a *batch*, an identical wet weight of test soil equivalent to a volume of ~350 mL should be transferred to each replicate test chamber.[25] The soil added to each test chamber should be smoothed (but not compressed) using a spoon, or by gently tapping the glass jar ≥3 times on the benchtop or with a hand.

For a single-concentration test (e.g., *site soil* tested at full strength only), a minimum of five replicate test chambers must be set up by adding an identical wet weight (equivalent to a volume of ~350 mL) of the same *batch* to each replicate chamber. For a multi-concentration test using *E. andrei/fetida*, a minimum of three replicate test chambers per *batch* must be prepared. For a multi-concentration test using *L. terrestris*, a minimum of five replicate test chambers per *batch* must be set up. A preliminary or *range-finding* test might prove worthwhile for selecting concentrations for the definitive test, in which instance the number of replicates per concentration could be reduced.

Following the addition of test soil to each test chamber, unperforated covers (Section 3.1.2) should be placed on them to minimize loss of moisture. The test chambers should be held overnight under test conditions (Section 4.1.2), for equilibration of the test soils (e.g., for chemical equilibration of chemical-spiked soil or site soil diluted with control soil). On Day 0 (i.e., when starting the test), each cover should be perforated.[26]

Test organisms are transferred to each test chamber the next day (i.e., Day 0 of the toxicity test). A number in excess of those required for the test should be removed

---

[24] Any liquid that has separated from a sample or subsample of test soil during transport and/or storage must be remixed into the sample.

[25] The wet weight of soil required to achieve a volume of ~350 mL depends on the moisture content, bulk density, and other characteristics of the soil, and will vary from sample to sample. Accordingly, the wet weight of each sample required to achieve this volume should be determined by transferring the amount of sample required to fill a preweighed (or tared) 500-mL glass beaker or jar to a 350-mL mark scribed on its side, followed by smoothing the surface and gently tapping the container on the bench top, three times. Thereafter, the wet weight of that quantity should be determined and recorded, and an identical wet weight added to each replicate test chamber.

[26] For a test involving a sample of contaminated soil with volatile compounds, it is recommended that opaque aluminum foil (rather than a transparent or translucent material such as Saran wrap™ or Parafilm™ that is shrunk or distorted by volatiles) be used as covers for the test chambers. These covers should not be perforated during the first week of the test to minimize gaseous emissions and to increase the exposure of worms to these volatile compounds. In this instance, the covers should be perforated on Day 7 (Stephenson *et al.*, 2001). If opaque (e.g., aluminum foil) covers are used in a test, the use of side lighting as well as overhead lighting is recommended to ensure that the minimal light intensity required at the surface of the soil in each test chamber is achieved (see Section 4.1.2). All test chambers, including those containing negative control soil, must be treated identically.

from a culture (Section 2.3) or acclimation (Section 2.4) chamber. Either *adult* (fully clitellated) or *sub-adult* worms within the acceptable size range (i.e., wet wt of individual worms, 250–600 mg if *E. andrei/fetida*; 3–10 g if *L. terrestris*) should be selected from this chamber, removed by gloved hand or using the blunt arm(s) of rounded forceps, and transferred briefly to a clean, shallow dish or tray where they are quickly rinsed in clean test water (i.e., de-ionized or distilled water). Thereafter, these worms are placed into a transfer container (e.g., a glass or aluminum tray measuring ~10 × 10 cm) lined with paper towel dampened with test water. The worms in this container should be given a final observation to confirm that their appearance is normal; any atypical worms should be discarded. Individual worms of as similar size as possible should be carefully selected, while confirming that they are within the acceptable size range, and then transferred individually (using one arm of rounded forceps) to the surface of the soil in each test chamber. The group of earthworms transferred to each test chamber should be randomly allocated with respect to treatment. If using *E. andrei/fetida*, five worms should be placed into each randomly chosen test chamber. For *L. terrestris*, three worms should be placed into each randomly chosen test chamber.[27]

Worms chosen should be similar in size (i.e., similar wet weights) and only those appearing healthy, similar in colouration, and active when removed from the bedding substrate should be selected. A minimum of 10 worms, taken randomly as surplus worms from the group selected for use in the test, must be weighed individually to determine the variability in size from worm to worm, for this sample. These individual weights must be recorded, and the mean (± SD) weight calculated and recorded (Section 7).

4.1.2 Test Conditions

- This is a 14-day soil toxicity test, during which the soil in each test chamber is not renewed.
- The test chamber is a 500-mL wide-mouthed glass jar. Its contents (i.e., a 350-mL volume of test soil) should normally be covered with a piece of perforated transparent or translucent material (e.g., Saran wrap™ or Parafilm™) (Section 3.1.2) secured to the lip of the jar using a rubber band. If the test material is known to contain volatile compounds (e.g., PAHs), the use of opaque aluminum foil as covers is recommended, together with side lighting sufficient to achieve the minimal acceptable light intensity at the surface of the soil (see Section 4.1.1 and the last bullet in 4.1.2).
- For a single-concentration test, at least five replicate test chambers must be set up for each test soil (i.e., each treatment). For a multi-concentration test, a minimum of three replicate test chambers per concentration must be set up for tests using *E. andrei/fetida*, and a minimum of five replicate chambers per concentration must be set up for tests with *L. terrestris;* at least five concentrations plus the appropriate control treatment(s) must be used.
- The test must be conducted at a daily mean temperature of 20 ± 2 °C. Additionally, the instantaneous temperature must always be 20 ±3 °C.
- Test chambers must be illuminated with a fixed daily photoperiod (e.g., 16-h light and 8-h dark, or 12-h light and 12-h dark), using incandescent or fluorescent lights. The photoperiod chosen should be the same as that to which the worms are acclimated before the test (see Sections 2.3.3 and 2.4.3). Light intensity adjacent to the surface of the soil in each test chamber should be 400–800 lux, and must be at least 400 lux as a minimum. This range is equivalent to a quantal flux of 5.6–11.2 μmol/(m^2 · s) for cool-white fluorescent, 6.4–12.8 μmol/(m^2 · s) for full-spectrum fluorescent, or 7.6–15.2 μmol/(m^2 · s) for incandescent.

4.1.3 Criterion for a Valid Test

For a valid test, the mean survival rate for earthworms held in *negative control soil* for 14 days must be ≥90% at the end of the test.

4.1.4 Food and Feeding

No supplementary feeding is to be provided during the 14-day test.

---

[27] Fewer *L. terrestris* are placed into each test chamber because of their larger size and associated loading density (i.e., biomass of worms per volume of test soil) in test chambers. For a multi-concentration test, the number of replicate test chambers per concentration (including the *negative control soil*) is increased from three (if *E. andrei/fetida*) to five (for *L. terrestris*) to enable the same number of test organisms per concentration (i.e., 15 worms/concentration).

4.1.5 Observations and Measurements During the Test

The primary biological endpoint for the test is the number of live worms in each test chamber on Day 14. Determining the number of live worms in each test chamber on Day 7 is also useful and frequently done, although such observations are optional. Depending on study objectives (e.g., if a determination of the median lethal time to 50% mortality of earthworms is desired), additional observations (e.g., at 1, 3, 6, and/or 24 h) of the number of live worms in each test chamber might also be included.

The condition, appearance, and number of live worms transferred to each test chamber on Day 0 must be observed and recorded during their transfer. At 1 h following their transfer, the number of worms on the surface of the soil in each jar, or against the glass on its inner sides or bottom, should be noted and recorded.

Observations of worms on Day 7 are optional but recommended (e.g., to determine the 7-day mortality rate for a single-concentration test, or the 7-day LC50 for a multi-concentration test). If the experimental design includes such measurements and calculations, the number of live and dead worms observed on Day 7 on the surface of the soil in each test chamber, or against the glass on its inner sides or bottom, should be determined and recorded. Thereafter, the contents of each test chamber should be transferred to a sorting tray or plastic sheeting, and the number of live and dead worms counted and recorded. Worms appearing to be dead should be touched gently on their anterior end with a glass rod or spatula; failure of any response is defined as death. All dead worms must be removed and discarded. Any worm fragments observed should also be discarded. Missing worms must be counted as dead, since dead worms can decompose rapidly and not be found. The appearance (e.g., normal or signs of discolouration or lesions) and behaviour (e.g., normally active, coiling, or lethargic) of each surviving worm should be noted and recorded. Immediately after this evaluation, the test soil must be returned to the jar, and all surviving worms transferred gently (by hand or using one arm of rounded forceps) to the surface of the soil. Further observations of survival, appearance, and behaviour are made at the end of the test (Section 4.1.6).

The application of a qualitative "squeeze test" (see Section 5.3) to each aliquot of test soil temporarily removed from a test chamber on Day 7 is useful to determine if its moisture content is acceptable. Another good indicator of acceptable soil moisture is the absence of soil particles adhering to the outer surface of the test organisms. If the outer surface of the surviving earthworms temporarily removed from the soil within a test chamber on Day 7 is relatively free of soil particles, then the moisture level for this soil is generally adequate. If, however, there are particles of soil adhering to the bodies of the earthworms, water should be added to the soil at this time. The squeeze test is used to confirm that the water content in the soil is adequate or insufficient. If deemed to be insufficient, a quantity of de-ionized water sufficient to achieve the homogeneous, crumbly consistency of the soil at the start of the test (Section 5.3) should be sprayed or gently misted onto the surface of the soil particles while they are distributed on the sorting tray or plastic sheeting.

Air temperature in the test facility (Section 3.1) must be measured daily (e.g., using a maximum/minimum thermometer) or continuously (e.g., using a continuous chart recorder).

The pH and moisture content of at least one replicate of each treatment (including the *negative control soil* and, if used, *reference soil*) must be measured and recorded at the beginning and end of the test. The initial (Day 0) measurements should be made using subsamples of each *batch* of test soil used to set up replicates of a particular treatment (see Section 4.1.1).[28] The final (i.e., Day 14) measurements should be made using subsamples of the replicates of each treatment to which worms were exposed, following the end-of-test observations of worm survival, condition, and behaviour.

Soil pH should be measured using a $CaCl_2$ slurry method (modified from Hendershot *et al*., 1993).[29] For

---

[28] On Day -1, one or more additional replicates of each test soil should be placed into a test chamber within the test facility. These replicates (with no worms added) should be reserved for physicochemical analyses of Day-0 conditions to which the worms are exposed. If desired, a separate set of replicates (with no worms added) should also be set up on Day -1, for physicochemical analyses of Day-14 conditions.

[29] The method by Hendershot *et al.* (1993) includes a step that involves air drying the sample for 48 h before its analysis for pH. Experience by Environment Canada investigators is that this step is needlessly time consuming (K. Doe, personal communication, Environment Canada,

these analyses, 4 g of hydrated soil [30] is placed into a 30-mL glass beaker (~ 3 cm in diameter and ~7 cm high) with 20 mL of 0.01 *M* $CaCl_2$.[31] The suspension should be stirred intermittently for 30 min (e.g., once every 6 min). The slurry should then be left undisturbed for ~1 h. Thereafter, a pH probe is immersed into the supernatant and the pH recorded once the meter reading is constant.

The moisture content of each test soil is measured by placing a 3–5 g subsample of each test soil into a pre-weighed aluminum weighing pan, and measuring and recording the wet weight of the subsample. Each subsample should then be placed into a drying oven at 105 °C until a constant weight is achieved; this usually requires a minimum of 24 hours. The dry weight of each subsample should then be measured and recorded. Soil moisture content must be calculated (on a dry-weight basis) by expressing the moisture content as a percentage of the soil dry weight:

$$\text{Moisture content (\%)} = \frac{\text{wet weight (g)} - \text{dry weight (g)}}{\text{dry weight (g)}} \times 100$$

It is important that moisture content (%) be calculated on a dry-weight basis (not on a wet-weight basis), since the results of these calculations are used with calculations of water-holding capacity (also calculated on a dry-weight basis) to express the optimal moisture content in test soils (see Section 5.3).

Depending on the nature of the test and the study design, concentrations of chemical(s) or chemical product(s) of concern might be measured for test soils or selected concentrations thereof, at the beginning and end of the test. For a test using a sample of field-collected *site soil*, the chemical(s) or chemical product(s) measured will depend on the contaminant(s) of concern (see Section 5.4). For a multi-concentration test with chemical-spiked soil, such measurements should be made for the high, medium, and low strengths tested, as a minimum (see Section 6.3). Aliquots for these analyses should be taken as described previously for pH, conductivity, and moisture content; analyses should be according to proven and recognized (e.g., SAH, 1992 and Carter, 1993) analytical techniques.

4.1.6 Ending the Test

The test is terminated after 14 days of exposure. At that time, the number of live and dead worms on the surface of the soil in each test chamber, or against the glass on its inner sides or bottom, should be determined and recorded. Thereafter, the contents of each test chamber must be transferred to a sorting tray or plastic sheeting, and the number of live and dead worms counted and recorded. Worms appearing to be dead should be touched gently with a glass rod or spatula; absence of any response is defined as death. Dead worms are discarded. Missing worms must be counted as dead. The appearance (e.g., normal or signs of discolouration or lesions) and behaviour (e.g., normally

---

Atlantic Environmental Science Centre, Moncton, NB; J. Princz, personal communication, Environment Canada, Biological Methods Division, Ottawa, ON, 2004), and does not appreciably modify the pH relative to that for hydrated (i.e., as per the toxicity test) soil (Courchesne *et al.*, 1995; J. Princz, personal communication, Environment Canada, Biological Methods Division, Ottawa, ON, 2004).

Becker-van Slooten *et al*. (2004) assessed three different soil slurry methods for measuring pH. The need for this testing was identified during EC's soil toxicity workshop in Vancouver, BC (February, 2003) where certain participants recommended that a commonly used and "universally standardized" method for measuring soil pH be incorporated into each of EC's soil toxicity test methods (EC, 2004c). The following three methods for measuring soil pH were compared: 1) 1 *M* KCl in water; 2) 0.01 *M* $CaCl_2$ in water; and 3) water only. Results of this investigation showed that there were advantages and disadvantages with each of these methods for measuring pH. However, based on practical considerations and the recommendations of the workshop participants (i.e., that a widely used method for characterizing soil pH be applied), the 0.01 *M* $CaCl_2$ method was recommended as the most appropriate for EC's soil toxicity test methods (Becker-van Slooten *et al*., 2004).

[30] It might be necessary to use a lower soil:$CaCl_2$ solution ratio (e.g., 2 g of soil with 20 mL of $CaCl_2$) for soils with a high organic matter content (i.e., for soils where the slurry does not yield a supernatant).

[31] To prepare 0.01 *M* $CaCl_2$, dissolve 2.940 g of calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$) with distilled water, in a 2000-mL volumetric flask. The conductivity of the $CaCl_2$ solution should be between 224 and 240 mS/m at 25 °C, and the pH should range within 5.5–6.5 at 25 °C (Hendershot *et al.*, 1993). If the pH is outside this range, it should be adjusted to the ranges using an HCl or $Ca(OH)_2$ solution. If the conductivity is not within the acceptable range, a new solution must be prepared.

active or lethargic) of each surviving worm should be noted and recorded. Immediately after this evaluation, subsamples of each test soil (including the *negative control soil* and, if included in the test, *reference soil*) should be taken for determinations of pH and moisture content (Section 4.1.5). Analyses for other chemical constituents (i.e., concentrations of contaminants) should also be made at this time using representative subsamples of each test soil (Section 4.1.5).

4.1.7 Test Endpoints and Calculations

The percent survival (percent mortality) in each test chamber on Day 14 of the test must be calculated and reported for each test. The mean (± SD) percent survival for all worms exposed to each treatment must also be calculated and reported for Day 14, using the survival data determined for all replicates.[32] Any optional observations of survival rates taken on Day 7 (see Section 4.1.5) should also be calculated and reported as percent survival in each test chamber, as well as mean (± SD) percent survival for each treatment.

For a single-concentration test (Section 4.1.1), the mean (± SD) value for percent survival determined for each treatment is compared with that for the sample(s) of *reference soil* or, as necessary and appropriate, compared with that for the negative control soil (see Section 5.5). For a multi-concentration test (see Sections 4.1.1, 5.3, and 6.2), the 14-day LC50 must be calculated and reported (data permitting). If 7-day observations of percent survival in each concentration were made during a multi-concentration test, it is recommended that the 7-day LC50 also be calculated and reported (data permitting). Environment Canada's guidance document on statistical methods for estimating endpoints of toxicity tests (EC, 2004b) provides definitive direction and advice for calculating LC50s, which should be followed (see Section 6.4.1, herein).

[32] This is based on a sample size (n) of at least 15 worms per treatment, if a multi-concentration test (e.g., three replicates of 5 worms/test chamber per treatment, if *E. andrei/fetida*, or five replicates of 3 worms/test chamber per treatment, if *L. terrestris*). For a single-concentration test involving five replicates per treatment, n = 25 if *E. andrei/fetida*, and n = 15 if *L. terrestris* (see Section 4.1.1).

4.1.8 Tests with a Reference Toxicant

Table 12 of Appendix F summarizes the guidance for performing reference toxicity tests given in other documents describing procedures and conditions for conducting 14-day lethality tests of soil toxicity using earthworms. Described herein are the procedures and conditions to be followed when performing reference toxicity tests in conjunction with a 14-day lethality test of soil toxicity using earthworms, as well as those reference toxicity tests undertaken in conjunction with an acute avoidance test using earthworms (see Section 4.2.8) or a test for effect on the prolonged survival, reproduction, and growth of earthworms (see Section 4.3.8). The procedures herein also apply to tests for assessing the acceptability and suitability of cultures of *E. andrei/fetida* or that of groups of *E. andrei/fetida* or *L. terrestris* being held for use in soil toxicity tests; and should be applied to assess intralaboratory precision when a laboratory is inexperienced with the biological test methods defined in this document and is initially setting up to perform them (see Sections 2.3.1, 2.3.9, and 2.4.9).

The routine use of a *reference toxicant* is necessary to assess, under standardized test conditions, the relative sensitivity of a portion of the population of adult or sub-adult earthworms within a particular culture (Section 2.3.9) or a particular acclimation chamber (Section 2.4.9) from which test organisms are selected for use in one or more definitive soil toxicity tests. Tests with a reference toxicant also serve to demonstrate the precision and reliability of data produced by the laboratory personnel for that reference toxicant, under standardized test conditions. A reference toxicity test, conducted according to the procedures and conditions described in Section 4.1, must be performed according to one of the following regimes:

(1) at least once per month using sub-adult or adult earthworms taken from the population of earthworms that is being cultured for use in the definitive test(s) (Section 2.3).

(2) at the same time as the definitive soil toxicity test(s), using sub-adult or adult worms taken from the same population as those used for the definitive test(s) (see Sections 2.3.9 and 2.4.9).

A laboratory that cultures earthworms (*Eisenia* sp.) and frequently performs soil toxicity tests using these organisms might choose to routinely monitor (e.g.,

monthly) the sensitivity of their culture(s) to one or more reference toxicants, while including a reference toxicity test using a portion of the worms used to start a definitive soil toxicity test. Alternatively, a laboratory culturing *Eisenia* sp. might choose to monitor the sensitivity of their culture(s) to a reference toxicant less frequently (e.g., once every two months), and to perform a reference toxicity test at the time that each definitive soil toxicity test is performed. A laboratory using earthworms (*L. terrestris* or *Eisenia* sp.) transported to the laboratory and acclimated to test conditions (see Section 2.4) in preparation for a specific set of definitive soil toxicity tests might choose to only perform one or more reference toxicity tests using a portion of the worms used to start the definitive tests. If *L. terrestris* are being held at a laboratory for an extended period (several months), that laboratory might also choose to routinely monitor (e.g., monthly) the sensitivity of the population being held.

Each reference toxicity test performed in conjunction with this acute lethality test for soil toxicity must be conducted as a *static* seven-day multi-concentration test. The test conditions and procedures described herein (Section 4.1) for performing a multi-concentration acute (14-day) lethality test with samples of test soil apply equally to each reference toxicity test, except that the test duration is restricted to seven days. Procedures given in Section 6 for the preparation and testing of chemicals spiked in negative control soil also apply here, and should be referred to for further information. Environment Canada's guidance document on using negative control sediment spiked with a reference toxicant (EC, 1995) provides useful information that is also applicable when performing reference toxicity tests with negative control soil spiked with a reference toxicant.

The reference toxicity test should be performed using 500-mL glass jars as test chambers (Section 3.1.2) and a 350-mL aliquot of test soil representing each treatment (concentration) in each test chamber. The number of replicate test chambers per concentration is species dependent, as is the number of worms per test chamber (Section 4.1.1). Wet weights of individual worms must be within the acceptable size range for each species that is specified in Section 4.1.1. Worms should not be fed during the seven-day test period.

Procedures for starting and ending a reference toxicity test should be consistent with those described in Sections 4.1.1 and 4.1.6, with the exception of the shorter (seven-day) test duration. Test conditions described in Section 4.1.2 apply. Test observations and measurements given in Section 4.1.5 should be followed.

To be valid, the mean seven-day survival rate for earthworms held in the aliquots of negative control soil used in a particular reference toxicity test must be at least 90%. Test endpoints to be calculated and reported include the mean percent survival in each treatment on Day 7, and the seven-day LC50 (including its 95% confidence limits). Results for a reference toxicity test should be expressed as mg reference chemical/kg soil, on a dry-weight basis.

Appropriate criteria for selecting the reference toxicant to be used in conjunction with an acute lethality test for soil toxicity include the following (EC, 1995):

- chemical readily available in pure form;
- stable (long) shelf life of chemical;
- can be interspersed evenly throughout *clean* substrate;
- good concentration-response curve for test organism;
- stable in aqueous solution and in soil;
- minimal hazard posed to user; and
- concentration easily analyzed with precision.

Stantec and Aquaterra Environmental (2004) considered the known properties of reagent-grade boric acid in these respects, and concluded that it was a suitable chemical for routine use as a reference toxicant alongside soil toxicity tests with earthworms.

The seven-day reference toxicity test requires a minimum of six treatments (i.e., *negative control soil* and five concentrations of reference toxicant). Reagent-grade boric acid is recommended for use as the reference toxicant(s) when performing soil toxicity tests with earthworms, although other chemicals may be used if they prove suitable.[33] Each test concentration should

[33] Aquaterra Environmental (1998a) initially evaluated the performance of various chemicals as candidate reference toxicants for use in conjunction with acute lethality tests for measuring soil toxicity to *E. andrei* or *L. terrestris*.

be made up according to the guidance in Sections 4.1.1 and 6.2, using artificial soil (Section 3.2.2) as substrate.

Routine reference toxicity tests (e.g., those performed monthly or in conjunction with each definitive test for soil toxicity) using boric acid (or another suitable reference chemical, such as potassium chloride) spiked in *negative control soil* should consistently apply the same test conditions and procedures described herein. A series of test concentrations should be chosen[34], based on preliminary tests, to provide partial mortalities in two or more concentrations and enable calculation of a 7-day LC50 (see Section 6.4).

Once sufficient data are available (EC, 1995), all comparable LC50s for a particular reference toxicant derived from these toxicity tests must be plotted successively on a *warning chart*. Each new LC50 for the same reference toxicant should be examined to determine whether it falls within ± 2 SD of values obtained in previous comparable tests using the same reference toxicant and test procedure (EC, 1997a, b, 2001). A separate warning chart must be prepared and updated for each dissimilar procedure (e.g., differing species of test organism, or differing reference toxicants). The warning chart should plot logarithm of concentration on the vertical axis against date of the test or test number on the horizontal axis. Each new LC50 for the reference toxicant should be compared with established limits of the chart; the LC50 is acceptable if it falls within the *warning limits*.

The logarithm of concentration (including LC50) should be used in all calculations of mean and standard deviation, and in all plotting procedures. This simply represents continued adherence to the assumption by which each LC50 was estimated based on logarithms of concentrations. The warning chart may be constructed by plotting the logarithmic values of the mean and ±2 SD on arithmetic paper, or by converting them to arithmetic values and plotting those on the logarithmic scale of semi-log paper. If it were demonstrated that the LC50s failed to fit a log-normal distribution, an arithmetic mean and SD might prove more suitable.

The mean of the available values of log(LC50), together with the upper and lower warning limits (±2 SD), should be recalculated with each successive LC50 for the reference toxicant until the statistics stabilize (EC, 1995, 1997a, b, 2001). If a particular LC50 fell outside the warning limits, the sensitivity of the test organisms and the performance and precision of the test would be suspect. Since this might occur 5% of the time due to chance alone, an outlying LC50 would not necessarily indicate abnormal sensitivity of the culture or population of earthworms from an outside source being held in the laboratory, nor unsatisfactory precision of toxicity data. Rather, it would provide a warning that there might be a problem. A thorough check of all culturing, holding/acclimation, and test conditions and procedures should be carried out. Depending on the findings, it might be necessary to

---

Subsequent studies by Aquaterra Environmental (2001) using boric acid spiked in negative control soil demonstrated that this chemical offered considerable promise as a reference toxicant in soil toxicity tests with earthworms. Further investigations by Stantec and Aquaterra Environmental (2004) confirmed the usefulness of boric acid as a suitable reference toxicant when performing monthly 7-day LC50s and maintaining a warning chart for this chemical using *E. andrei*. Testing by Stantec and Aquaterra Environmental (2004) also demonstrated the sensitivity of *E. andrei* to boric acid in eight-week tests for effects on their survival, reproduction, and growth, and showed similar findings for a number of eight-week tests with this chemical performed according to Section 4.3 herein.

[34] Stephenson (2003b) and Stantec and Aquaterra Environmental (2004) demonstrated endpoint values ranging from 3295–4915 mg boric acid/kg soil (dry wt) in their results for seven-day LC50s with boric acid mixed in artificial soil or a *clean* field-collected clay-loam soil using *E. andrei, E. fetida*, or *L. terrestris* and the test method for a reference toxicity test described herein. Values differed little due to earthworm species or the type of soil in which this reference toxicant was mixed. For these tests, the lowest test concentration (2000 mg/kg) caused no demonstrable effect, 4000–6000 mg/kg caused partial mortalities, 8000 mg/kg killed some or all worms, and the highest test concentration (10 000 mg/kg) killed all worms within seven days. See Appendix I for guidance in selecting an appropriate series of test concentrations (assuming a log-concentration response) for use in toxicity tests with this or other chemicals to be used in a reference toxicity test.

As part of a series of interlaboratory studies performed to validate Environment Canada's acute lethality test described in Section 4.1, 11 laboratories undertook concurrent seven-day LC50s with *E. andrei* and multiple concentrations of boric acid spiked in artificial soil. Each of the participating laboratories achieved valid test results (i.e., the seven-day survival rates in negative control soil were consistently ≥90%). For these tests, the mean seven-day LC50 for boric acid in artificial soil derived for the data provided by the laboratories was 3826 mg H_3BO_3/kg dry wt, with values for individual laboratories ranging from 3236–4198 mg/kg. The CV of 9% for these LC50s was low, showing good interlaboratory precision (EC, 2004a).

repeat the reference toxicity test, establish a new culture, select worms from an alternate culture, or obtain a new population of test organisms from an outside source, before undertaking further soil toxicity tests.

Results that remained within the warning limits might not necessarily indicate that a laboratory was generating consistent results. Extremely variable historic data for a reference toxicant would produce wide warning limits; a new data point could be within the warning limits but still represent undesirable variation in test results. A CV of no more than 30%, and preferably 20% or less, has been suggested as a reasonable limit by Environment Canada (EC, 1995, 2004b) for the mean of the available values of log(LC50) (see preceding paragraph). For this biological test method, the CV for mean historic data derived for reference toxicity tests performed using boric acid should not exceed 30%.[35]
Concentrations of reference toxicant in all stock solutions should be measured chemically using appropriate methods (e.g., analytical methods involving AES with ICAP scan, for concentration of boron). Test concentrations of reference toxicant in soil are prepared by adding a measured quantity of the stock solution to negative control soil[36], and mixing thoroughly.[37] Upon preparation of the mixtures of the reference toxicant in soil, aliquots should be taken from at least the negative control soil as well as the low, middle, and high concentrations.[38] Each aliquot should either be analyzed directly, or stored for future analysis (i.e., at the end of the test) if the seven-day LC50 based on nominal concentrations was found to be outside the warning limits. If stored, sample aliquots must be held in the dark at 4 ± 2 °C. Stored aliquots requiring chemical measurement should be analyzed promptly upon completion of the reference toxicity test. The seven-day LC50 should be calculated based on the measured concentrations if they are appreciably (i.e., ≥20%) different from nominal ones and if the accuracy of the chemical analyses is satisfactory.

If boric acid is used as a reference toxicant, the following analytical method applies (OMEE, 1996). A 1–5 g subsample of soil spiked with boric acid is dried at 105 °C to constant weight. A 1-g aliquot is then extracted using an 0.01 *M* solution of $CaCl_2$, by boiling a slurry of soil in 50 mL of this extraction solution and then re-adjusting the final volume to 50 mL using more extraction solution. The 50-mL extract is then filtered through a #4 Whatman™ filter, and diluted to a final volume of 100 mL. A blank sample is prepared in a similar manner. The filtrate is analyzed for elemental

---

[35] The technical report by Stantec and Aquaterra Environmental (2004) includes a seven-day warning chart for boric acid and *E. andrei*, based on nine reference toxicity tests performed with mixtures of this reference chemical in artificial soil over a 21-month period. All but one of the data points fell within the upper and lower warning limits. The CV derived for these mean historic data was 14.7%.

As part of a series of tests to validate Environment Canada's acute lethality test described in Section 4.1, 11 laboratories undertook a seven-day LC50 with *E. andrei* exposed to a range of concentrations of boric acid in artificial soil. The interlaboratory CV for the mean seven-day LC50s derived from these 11 tests was low, at 9% (EC, 2004a).

[36] Section 6.2 "*Preparing Test Mixtures*" includes an example showing the amounts of de-ionized or distilled water and boric acid to be added to artificial soil, to prepare a given treatment for a reference toxicity test with a specific concentration of boric acid in artificial soil. The calculations in this example show the amount of water necessary to adjust the moisture content of the artificial soil to a fixed percentage (i.e., 70%) of the soil's water-holding capacity, while taking into account the volume of the stock solution of boric acid as part of the overall adjustment for soil moisture content.

[37] An accepted procedure is to add a precalculated volume of stock solution (using volumetric and/or graduated pipets) to a glass Erlenmeyer™ flask, diluting to a graduated mark using de-ionized water, and then adding a measured volume to the soil. The flask is then rinsed three times with de-ionized water, and the rinsate is added to the soil. The mixture of soil and stock solution is then mixed thoroughly (for approximately three minutes) with a mechanical mixer (e.g., a hand-held mixer with revolving stainless steel beaters) until the soil appears homogeneous in colour, texture, and moisture content. During the mixing process, the soil in the mixing bowl should also be stirred intermittently using a large stainless steel spoon to facilitate homogenization.

[38] If the LC50 for each reference toxicity test is to be based on measured concentrations, it is recommended that one or more aliquots of the chemical-in-soil mixture representing each test concentration be collected and analyzed. If, however, the LC50 for each test is based on nominal concentrations, sampling and analysis of aliquots from at least the low, middle, and high test concentrations is recommended.

boron using ICAP/AES. The boric acid concentration in the soil is then calculated using the following equation:

$$\text{boric acid (mg/kg, dry wt)} = \frac{\mu\text{g B/mL (measured)} \times \text{final volume (mL)} \times MW_{\text{boric acid}}/MW_{\text{boron}}}{1000\ (\mu\text{g}) \times \text{weight of sample (mg dry wt)}} \times 10^6$$

The analytical limit of detection for boric acid in soil is reportedly 1 mg boric acid/kg soil dry wt in most instances (Stephenson, 2003b).

4.2 Acute Avoidance Test

Table 4 provides a summary checklist of recommended conditions and procedures to be universally applied to each acute avoidance test with samples of contaminated or potentially contaminated soil, as well as those for testing specific types of test materials or substances (e.g., samples of biosolids, or *negative control soil* spiked in the laboratory with one or more test chemicals or chemical products).

Universal procedures for performing an acute avoidance test are described herein. This biological test method uses adult earthworms as test organisms, and measures avoidance of test soils[39] as the biological endpoint. Test organisms are either laboratory-cultured *E. andrei/fetida* (Section 2.3) or worms (*E. andrei, E. fetida,* or *L. terrestris*) obtained from an outside source (Section 2.2) and acclimated to laboratory conditions (Section 2.4) before their use in the test. Test duration is 48 hours if *E. andrei* or *E. fetida*, and 72 hours if *L. terrestris*.[40, 41] The test organisms are not fed during the test, and the test soils are not renewed.

4.2.1 Beginning the Test

A minimum of five test units is required for a single-concentration test, with each unit containing the same two treatments (i.e., a single site soil or a single concentration of test soil, plus a *negative control soil* or *clean reference soil*) in alternating compartments. For a multi-concentration test, a minimum of five test units is also required, with each test concentration restricted to a single test unit (together with aliquots of a *clean* soil in alternating compartments). The use of more test units (e.g., 6–10) in a multi-concentration test is recommended to provide for testing of a broader range of test concentrations and to increase the likelihood of achieving the statistical endpoint sought for this test (see Section 4.2.7). Another option for a multi-concentration test is to conduct a preliminary range-finding test for acute avoidance responses to a wide range of concentrations using this test apparatus (n = 10 worms per test concentration), followed by the replication of each test concentration by using two test units per concentration (n = 20 worms per test concentration). In certain instances (e.g., highly contaminated soil, or soil for which the LC50 is unknown), a preliminary acute lethality test (see Section 4.1) might be performed to assist in selecting test concentrations to be used in a (sublethal) avoidance test.

---

[39] Using the recommended test apparatus defined and illustrated in Section 3.1.3 (Figure 2), groups of earthworms are given a choice between *negative control soil* or *reference soil* and a test soil (e.g., a field-collected soil from a contaminated or potentially contaminated site; or a particular concentration of a chemical or chemical product spiked into *negative control soil*). Each worm (n = 10) within a test unit is free to move between the *clean* soil (i.e., *negative control soil* or *reference soil*) held in three compartments, and the test soil held in three alternating compartments (six compartments/test unit) therein. At the end of a defined exposure period (i.e., 48 h if *E. andrei* or *E. fetida*; 72 h if *L. terrestris*) , the total numbers of worms in the *clean* soil and the test soil are determined and compared statistically (see Section 4.2.7).

[40] A test duration of 48 h has been chosen for *Eisenia* spp. to harmonize with the avoidance test using *E. fetida* or *E. andrei* to be published by ISO (2003). A test duration of 72 h has been chosen for *L. terrestris* based on past experience with this species in avoidance tests and, in certain instances, its somewhat slower avoidance response than *Eisenia* spp. (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1998).

[41] A 48-h test for avoidance by *E. andrei* of multiple concentrations of boric acid spiked in a field-collected reference soil (Alberta Black Chernozem soil) was performed by four laboratories to validate this test method. Each of these laboratories was able to achieve valid test results (see Section 4.2.3). The mean 48-h EC50 for boric acid spiked in this reference soil was 874 mg H_3BO_3/kg soil dry wt, with values for individual laboratories ranging from 757–979 mg/kg. The interlaboratory CV for these EC50s was 11%, which is considered to be well within an acceptable level of precision between laboratories (EC, 2004a).

Table 4 **Checklist of Recommended Conditions and Procedures for Conducting an Acute (48 or 72 h) Sublethal Test for the Effect of Contaminated Soil on the Avoidance Behaviour of Earthworms (*Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris*)**

Universal		
Test type	—	whole soil toxicity test; no renewal (static test)
Test duration	—	48 hours if *E. andrei* or *E. fetida*; 72 hours if *L. terrestris*
Test organisms	—	*E. andrei, E. fetida,* or *L. terrestris*; sexually mature adults with clitellum; individual wet wt ranging within 250–600 mg if *E. andrei*, 250–800 mg if *E. fetida*, or 3–10 mg if *L. terrestris*; 10 worms per test unit
Negative control soil	—	depends on study design and objectives; *clean* field-collected soil or artificial soil if testing site soils; recommend artificial soil for tests with chemicals or chemical products spiked in soil
Test unit	—	circular container with central chamber (ID, ~5.4 cm) and six pie-shaped interconnecting compartments, with fitted lid; constructed of high-quality stainless steel or Plexiglas™; OD, ~23 cm, height, ~9 cm, 1.0-cm holes in bottom of central chamber (two/compartment) and sides of compartments (three/side) for worm movement
Amount of soil per central chamber	—	none
Amount of soil per test compartment	—	identical wet wt, equivalent to a volume of ~350 mL
Moisture content, test soils	—	hydrate to the optimal percentage of its WHC if field-collected soil (see Section 5.3), or to ~70% of WHC if artificial soil
Number of compartments per test unit with same treatment	—	three (negative control soil or reference soil in each of three compartments, with a single test soil or concentration thereof in each of three alternate compartments)
Number of treatments per test unit	—	two (negative control soil or reference soil, plus a single sample or concentration of a test soil (e.g., a field-collected test soil at 100% or lower concentration, or a single concentration of a chemical-spiked soil); alternate treatment in each neighbouring compartment
Number of test units per test	—	≥5 (normally one per concentration, if a multi-concentration test; at least five test units, if a single-concentration test)
Temperature	—	daily average, 20 ± 2 °C; instantaneous, 20 ± 3 °C
Lighting	—	continuous darkness (test units wrapped in aluminum foil if made of transparent or translucent Plexiglas™)
Feeding	—	do not feed

Measurements during test — soil moisture content and pH in each treatment/concentration, at start and end; temperature in test facility, daily or continuously

Observations during test — compartment (treatment) entered by each worm at start of test; number of live worms in each test compartment, at test end; number of dead worms in each test compartment, at test end; number of live or dead worms seen on surface of soil in each test compartment at test end; obvious pathological symptoms (e.g., open wounds) or distinct behavioural abnormalities (e.g., lethargy) for surviving worms in each test compartment at test end

Biological endpoint — number of live worms per treatment in each test unit (i.e., total number of live worms in the three compartments containing the same test soil, for each of the two treatments), at test end

Statistical endpoints — percent of live worms per treatment in each test unit at test end; EC50 (or other ECp) for avoidance if multi-concentration test

Test validity — invalid if percent survival of worms in any test unit <90% at test end

Test with reference — perform once/month, or in conjunction with definitive test(s) with soil samples; use toxicant boric acid; prepare and test ≥5 concentrations plus a negative control, using artificial soil as substrate; determine seven-day LC50 (including 95% confidence limits) according to procedures and conditions defined in Section 4.1.8; express as mg boric acid/kg, dry wt

Field-Collected Soil

Transport and storage — seal in plastic and minimize air space; transport in darkness (e.g., using an opaque cooler, plastic pail, or other light-tight container); do not freeze or overheat during transport; store in dark at 4 ± 2 °C; test should start within two weeks, and must start within six weeks unless soil contaminants are known to be stable

Negative control soil — either natural, uncontaminated field-collected soil or artificial soil, for which previous acute avoidance tests with the test species showed that the criterion for test validity could be regularly met

Reference soil — one or more samples for tests with field-collected soil; ideally taken from site(s) presumed to be clean but near sites where test soils are collected; characteristics including percent organic matter, particle size distribution, and pH similar to test soil(s)

Characterization of test soils — at least percent moisture, WHC, pH, conductivity, percent TOC, percent organic matter, and particle sizes (% sand, % silt, % clay); optionally, contaminants of concern (e.g., metals, PAHs, pesticides)

Preparation of test soils — if necessary, remove debris and indigenous macro-organisms using forceps; if necessary, press through a sieve of suitable mesh size (e.g., 4–6 mm); mix; determine soil moisture content; hydrate with de-ionized or distilled water (or, if and as necessary, dehydrate) to the optimal percentage of its WHC (see Section 5.3); mix; dilute with negative control or reference soil if multi-concentration test; ensure homogeneity

Soil Spiked with Chemical(s) or Chemical Product(s)

Negative control soil	—	recommend artificial soil or a *clean* field-collected soil
Characterization of chemical(s) or control chemical product(s)	—	information on stability, water solubility, vapour pressure, purity, and biodegradability of chemical(s) or chemical products(s) spiked into negative soil should be known beforehand
Solvent	—	de-ionized water is the preferred solvent; if an organic solvent is used, the test must include a solvent control
Preparation of mixtures	—	procedure dependent on the nature of the test substance(s) and the test design and objectives; chemical/soil mixtures may be prepared manually or by mechanical agitation; test substance(s) may be added as measured quantities in solution (i.e., in water or an organic solvent) or as a solid material comprised partly or completely of the test substance(s); ensure homogeneity
Concentration within soil mixture of chemical(s) or chemical product(s) added	—	normally measure at beginning and end of test, in high, medium, and low strengths as a minimum

Each test unit (see Section 3.1.3) placed within the test facility must be clearly coded or labelled to enable identification of the test soil or its concentration therein. Each of the six compartments within each test unit (Section 3.1.3) must also be coded (e.g., identified by numbers or letters) or otherwise marked to distinguish the test soil therein. The date and time when the test is started must be recorded, either directly on the labels or on separate data sheets dedicated to the test. The test units should be positioned randomly within the test facility.

On the day preceding the start of the test (i.e., on Day -1), each sample or subsample of test soil or similar particulate material, including *negative control soil* and, if used, *reference soil*, should be mixed thoroughly[42] (see Sections 5.3 and 6.2) to provide a homogeneous mixture consistent in colour, texture, and moisture. If field-collected samples of *site soil* are being prepared for testing, large particles should be removed before mixing, along with any vegetation or macroinvertebrates observed (see Section 5.3).

The quantity of each test soil mixed as a *batch* should be enough to set up the replicates of that treatment (see Table 4) plus an additional amount for the physicochemical analyses to be performed (Section 4.2.5). The moisture content of each test soil should be known or determined, and adjustments made as necessary by mixing in *test water* (or, if and as necessary, by dehydrating the sample) until the desired moisture level is achieved (see Sections 5.3 and 6.2). Quantitative measures of the homogeneity of a *batch* might be made by taking aliquots of the mixture for measurements such as particle size analysis, total organic carbon content (%), organic matter content (%), moisture content (%), and concentration of specific chemicals.

Immediately following the mixing of a *batch*, an identical wet weight of *negative control soil* (natural or artificial; see Section 3.2) or *reference soil* (see Section 3.4), equivalent to a volume of ~350 mL[43], should be

[42] Any liquid that has separated from a sample or subsample of test soil during transport and/or storage must be remixed into the sample.

[43] The wet weight of soil required to achieve a volume of ~350 mL depends on the moisture content, bulk density, and other characteristics of the soil; and will vary from sample to sample. Accordingly, the wet weight of each sample required to achieve this volume should be determined by transferring that amount of sample required

transferred to every second test compartment (three/test unit) within each test unit included in an acute avoidance test. Thereafter, a weighed quantity (sample dependent; equivalent to a volume of ~350 mL)[43] of *test soil* from the same *batch* of mixed test material should be transferred to the other three test compartments within a test unit. Depending on the nature of the test (e.g., a single-concentration test with five replicate test units per treatment; or a multi-concentration test with each concentration restricted to a single test unit), the same test material (i.e., test soil from the same *batch*) is placed into each of three alternating compartments within one or more test units. The soil added to each test compartment should be smoothed (but not compressed) using a spoon, to evenly distribute the soil therein.

Following the addition of *clean* soil (i.e., *negative control soil* or *reference soil*) and a *test* soil (i.e., contaminated or potentially contaminated soil from the same *batch*) to each of three alternating compartments within each test unit, each compartment should be covered with a lid (Section 3.1.3) to minimize moisture loss. The test units should be held overnight at the test temperature (Section 4.2.2) to enable chemical equilibration of the test soils therein.

Ten test organisms (Section 2) are transferred individually (i.e., one worm at a time) to each test unit the next day (i.e., on Day 0). At that time, a number of worms in excess of those required for the test should be removed from a culture (Section 2.3) or acclimation (Section 2.4) chamber. Adult (fully clitellated) worms within the acceptable size range (i.e., wet wt of individual worms, 250–600 mg if *E. andrei*, 250–800 mg if *E. fetida*, or 3–10 g if *L. terrestris*) should be selected from this chamber, removed by gloved hand or by using the blunt arm(s) of rounded forceps, and transferred briefly to a clean, shallow dish or tray where they are quickly rinsed in clean test water (i.e., de-ionized or distilled water). Thereafter, these worms are placed into a transfer container (e.g., a glass or aluminum tray measuring ~10 × 10 cm) lined with paper towel dampened with test water. A final observation should be made of the worms in this container to confirm that their appearance is normal. Any atypical worms should be discarded. Thereafter, individual worms of as similar size as possible should be carefully selected while confirming that they are within the acceptable size range, and then transferred one at a time to the central chamber of each test unit (devoid of soil substrate). After the first worm is placed into this chamber, it is observed until it has moved into a compartment containing soil. The second worm is then added, and observed until it has disappeared from the central chamber into a neighbouring compartment. This procedure is repeated sequentially until the full complement of 10 worms per test unit has been added.

The test compartment (and its contents) entered by each worm should be noted and recorded.[44] Any worm that does not enter a test compartment within 30 minutes should be removed and discarded, and replaced with another worm from the transfer container.[45] Once a group of 10 worms has been added to a test unit, and all worms have moved into *clean* or *test* soil within the compartments, the time must be recorded (t = 0 h) and the lid placed on the test unit. Any test units constructed of transparent or translucent Plexiglas™ (see Section 3.1.3) must either be wrapped with opaque sheeting (e.g., aluminum foil) or held in a darkened facility throughout the test period.[46]

Worms chosen should be similar in size, and only those appearing healthy, similar in colouration, and active when removed from the bedding substrate

---

to fill a preweighed (or tared) 500-mL glass beaker or jar to a 350-mL mark scribed on its side, after gently smoothing (not compressing) the surface of the soil at this mark. Thereafter, the wet weight of that quantity should be determined and recorded, and an identical wet weight transferred to each of three (alternate) compartments within a test unit.

[44] Records of entry to test compartments provide useful information as to whether the worms enter randomly and are initially randomly distributed within the test compartments (see Section 4.2.7), or whether they show a preference for *clean* soil (i.e., *negative control soil* or *reference soil*) versus the *test* soil in alternate compartments within a test unit.

[45] Individual worms typically move from the central chamber of a test unit to a test compartment within three to five minutes (Stephenson *et al.*, 1998).

[46] Worms must be held under conditions of continuous darkness during this test, to prevent light from affecting their behavioural response. Use of test units constructed of stainless steel or opaque Plexiglas™ effectively provides darkened conditions; otherwise, testing in darkness or shrouding the test units in aluminum foil or other opaque wrapping material is required.

should be selected. A minimum of 10 worms, taken randomly as surplus worms from the group selected for use in the test, must be weighed individually to determine the variability in size from worm to worm, for this sample. These individual weights must be recorded, and the mean (± SD) weight calculated and reported (Section 7).

4.2.2 Test Conditions

- This is a sublethal test for acute (48 or 72 h) avoidance of test soils by adult earthworms, during which the test soils are not renewed and the worms in each test unit are left undisturbed to choose between *clean* soil (i.e., *negative control soil* or *reference soil*) and a single *test* soil (e.g., a field-collected test soil at 100% or lower concentration, or a single concentration of a chemical-spiked soil)

- Each test unit is comprised of a central chamber devoid of soil and six identical pie-shaped test compartments which are interconnected and enable movement of worms from compartment to compartment. Three of the test compartments in each test unit contain *clean* soil from the same *batch*, and three contain a single *test* soil from the same batch. The location of *clean* and *test* soil in the six compartments within a test unit alternates between compartments (i.e., each neighbouring compartment contains an alternate treatment).

- For a single-concentration test, at least five replicate test units must be used. Each test unit has three compartments containing *clean* soil from the same *batch*, and three compartments containing a single *test* soil from the same batch. Identical aliquots of *clean* and *test* soil from the same two *batches* are placed into alternating compartments within all five test units.

- For a multi-concentration test, at least five test units must be used and more (i.e., six to ten) are recommended. Each test unit has three compartments containing *clean* soil from the same *batch*, and a single concentration of *test* soil from the same batch. Identical aliquots of *clean* soil from the same batch are placed into alternating compartments within each of these test units. The concentration of test soil in the three alternating compartments of a single test unit is identical (from the same batch); however, concentrations of test soil differ from unit to unit.

- The test must be conducted at a daily mean temperature of 20 ± 2 °C. Additionally, the instantaneous temperature must always be 20 ± 3 °C.

- Test organisms are held in continuous darkness throughout the test period.

4.2.3 Criterion for a Valid Test

For a valid test, the percent survival of all earthworms in each test unit containing negative control soil or reference soil, plus test soil, must be ≥90% at test end.

4.2.4 Food and Feeding

No supplementary feeding is to be provided during the test.

4.2.5 Observations and Measurements During the Test

When adding test organisms to the central chamber of each test unit at the start of the test, the compartment entered by each worm should be observed and recorded (see Section 4.2.1). Observations are not possible once the test begins (t = 0). Care should be taken to not move or otherwise disturb the test units throughout the test period (or until after side partitions have been inserted at the end of the test to segregate worms within each compartment; see Section 4.2.6).[47] The primary observation for the test is the number of live worms residing in each test compartment at the end of the test (see Section 4.2.6).

Air temperature in the test facility (Section 3.1) must be measured daily (e.g., using a maximum/minimum thermometer) or continuously (e.g., using a continuous chart recorder).

The pH and moisture content of at least one replicate of each test soil (including the *negative control soil* and, if used, *reference soil*) must be measured and recorded at the beginning and end of the test. Additionally, it is recommended that conductivity be measured at the beginning and end of the test in

---

[47] Banging, jarring, or other related activities (such as moving the test units during the test or upon its completion but before the insertion of side partitions) which disturb the worms can cause them to start moving from compartment to compartment, and can result in spurious findings (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001).

instances where the test soil is anticipated to have a high salt content. The initial measurements should be made using subsamples of each *batch* of test soil used to set up replicates of a particular treatment (see Section 4.2.1). The final (i.e., t = 48 h if *Eisenia* spp., t = 72 h if *L. terrestris*) measurements should be made using subsamples of the replicates of each treatment to which worms were exposed, following the end-of-test observations of worm distribution, survival, appearance, and behaviour (see Section 4.2.6). Measurements of soil pH and moisture content should be made according to the guidance in Section 4.1.5.

The test soils might be analyzed for concentrations of chemical(s) or chemical product(s) of concern. Guidance in Section 4.1.5 applies here. Sections 5.4 and 6.3 should be consulted for further advice.

4.2.6 Ending the Test

The test is terminated after 48 hours of exposure if using *E. andrei* or *E. fetida*, and after 72 hours of exposure if using *L. terrestris*. At that time, the lid of each test unit is removed without any other movement of or disturbance to the test apparatus. Then, a side partition (see Section 3.1.3, including Figure 2) is quickly inserted adjacent to the side of each test compartment to confine the test organisms to each compartment. Immediately thereafter, the number of live and dead worms on the surface of the soil in each test compartment must be observed and recorded. The contents of each test compartment should then be gently removed with a spoon or spatula, placed into a sorting tray or onto a plastic sheet, and the number of live and dead worms counted and recorded.[48] Worms appearing to be dead should be touched gently on their anterior end with a glass rod or spatula; absence of any response is defined as death. Dead worms are discarded. Missing worms must be counted as dead. The appearance (e.g., normal or signs of discolouration or lesions) and behaviour (e.g., normally active or lethargic) of each surviving worm should be noted and recorded.

Immediately after this evaluation, subsamples of each test soil (including the *negative control soil* and, if included in the test, *reference soil*) should be taken for determinations of pH and moisture content (Section 4.2.5). Analyses for other chemical constituents (i.e., concentrations of contaminants) should also be made at this time using representative subsamples of each test soil (Section 4.2.5).

4.2.7 Test Endpoints and Calculations

For each test, the total number of surviving worms in the test soil (i.e., contaminated or potentially contaminated soil) and the *clean* soil (i.e., *negative control soil* or *reference soil*) within each test unit at the end of the test must be determined and recorded.

For a single-concentration test, the mean (± SD) number of surviving worms recovered from the test soil and the *clean* soil in each of the five (or more) replicate test units must be calculated. These values should be compared statistically using a one-tailed *Student's t-test* or another appropriate statistic for pairwise comparisons (see Section 5.5). Results showing a significantly lower mean number of surviving worms in the test soil, relative to those in the *clean* soil, indicate an avoidance response to the test soil (or a preference response to the *clean* soil).

For a multi-concentration test, the total number of surviving worms in each concentration of test soil within each test unit must be tabulated. The presence and extent of an apparent avoidance response to each test concentration is determined based on the (lesser) number of worms in the test soil relative to a *neutral* (no avoidance, no preference) response. A neutral response is defined as the presence of an equal number of worms in the test soil and the *clean* soil (i.e., *negative control soil* or *reference soil*) at the end of the exposure period. Using this definition, the total number of worms determined to be in a particular test soil within a test unit is converted to a value indicative of percent avoidance as follows:[49]

---

[48] Rarely, a worm is severed by the insertion of a side partition at the end of the test. If any severed segment of a worm is found in a test compartment, it should be counted and recorded only if that segment is from the anterior end of the worm.

[49] Using this equation, and assuming that numbers are based on the distribution of each group of 10 worms within a single test unit, the avoidance response to each test soil is calculated as follows: (a) if 5 or more worms in test soil, (5 – 5) ÷ 10 × 100 = 0% avoidance; (b) if 4 worms in test soil, (6 – 4) ÷ 10 × 100 = 20% avoidance; (c) if 3 worms in test soil, (7 – 3) ÷ 10 × 100 = 40% avoidance; (d) if 2 worms in test soil, (8 – 2) ÷ 10 × 100 = 60% avoidance; (e) if 1 worm in test soil, (9 – 1) ÷ 10 × 100 = 80% avoidance; and (f) if 0 worms in test soil, (10 – 0) ÷ 10 × 100 = 100% avoidance. If the experimental

$$\text{avoidance (\%)} = \frac{\text{no. in clean soil} - \text{no. in test soil}}{\text{total number of worms}} \times 100$$

where:

"no. in clean soil"	is the number of live worms found in all compartments containing *clean* soil, at the end of the test;
"no. in test soil"	is the number of live worms found in all compartments containing test soil, at the end of the test; and
"total number of worms"	is the total number of live worms found in all compartments, at the end of the test.

Data permitting, the *median effective concentration* (*EC50*) and, if desired, any other ECp (e.g., EC20 or EC25) causing avoidance should then be estimated (see Section 6.4) based on the percent avoidance responses determined for each test concentration.

An investigator might wish to analyze the data showing numbers of worms entering each test compartment at the start of the test (see Sections 4.2.1 and 4.2.5), to test for randomness of this response. Chi-square analysis (EC, 2004b) is suitable for this purpose. A significant difference due to treatment (i.e., *clean* soil versus *test* soil) indicates an initial detection and response (i.e., avoidance or preference) of earthworms to the test soil. A significant difference among compartments suggests a lack of random movement of earthworms into the test compartments at the start of the test.

4.2.8 Tests with a Reference Toxicant

The routine use of a reference toxicant is necessary to assess, under standardized test conditions, the relative sensitivity of a portion of the population of adult earthworms within a particular culture (Section 2.3.9) or a particular acclimation chamber (Section 2.4.9) from which test organisms are selected for use in one or more definitive soil toxicity tests including those used in acute avoidance tests. Tests with a reference toxicant also serve to demonstrate the precision and reliability of data produced by the laboratory personnel for that reference toxicant, under standardized test conditions.

A seven-day lethality test with a reference toxicant, conducted according to the procedures and conditions described in Section 4.1.8, must be performed using a portion of any population of adult earthworms being cultured (Section 2.3) or acclimated to test conditions (Section 2.4) and used to provide worms for an acute avoidance test. Either of the two regimes described in Section 4.1.8 for performing this reference toxicity test may be applied in conjunction with an acute avoidance test.

4.3 Test for the Effects of Prolonged Exposure on Survival, Reproduction, and Growth

The intent of this biological test method is to measure the effects of prolonged exposure of earthworms (laboratory-cultured *E. andrei*) to contaminated soil on their survival and reproductive success, and on the subsequent growth of their progeny.

Table 5 provides a summary checklist of recommended conditions and procedures to be universally applied when performing this biological test method, as well as those for testing specific types of test materials or substances (e.g., samples of biosolids, or *negative control soil* spiked in the laboratory with one or more test chemicals or chemical products). This test method has been developed using guidance provided by ISO (1991, 1998) and OECD (2000) for the performance of tests for the effects of chemical-spiked soil on the reproduction of *E. andrei/fetida* (see Appendix G).

---

design includes two test units per concentration (see Section 4.2.1), with 10 worms per unit (i.e., n = 20, the same equation applies when calculating percent avoidance for each concentration. For instance, if 10 or more worms were found in test soil within these two units, (10 - 10) ÷ 20 × 100 = 0% avoidance; if 9 worms were found in the test soil, (11 - 9) ÷ 20 × 100 = 10% avoidance; if 8 worms were found in the test soil, (12 - 8) ÷ 20 × 100 = 20% avoidance, etc.

Table 5 Checklist of Recommended Conditions and Procedures for Conducting Tests for Effects of Prolonged Exposure to Contaminated Soil on the Survival, Reproduction, and Growth of Earthworms (*Eisenia andrei*)

Universal		
Test type	—	whole soil toxicity test; no renewal (static test)
Test duration	—	≥56 days = ≥8 weeks
Test organisms	—	cultured *E. andrei*; sexually mature adults with clitellum; individual wet wt, 250–600 mg; choose worms as similar in wet wt as possible; two worms per test chamber and 20 worms per test treatment; hold for seven days in negative control soil containing added food identical to that used in the test
Number of replicates	—	10 replicates/treatment; each replicate consisting of two worms in a test chamber
Negative control soil	—	depends on study design and objectives; *clean* field-collected soil or artificial soil if testing site soils; recommend artificial soil for tests with chemicals or chemical products spiked in soil
Test chamber	—	500-mL glass jar; perforated translucent or transparent cover (e.g., Saran wrap™ or Parafilm™), secured with a rubber band, is recommended
Amount of soil/ test chamber	—	identical wet wt, equivalent to a volume of ~350 mL; ~200 g dry wt, if artificial soil
Moisture content, test soils	—	hydrate to the optimal percentage of its WHC if field-collected soil (see Section 5.3), or to ~70% of WHC if artificial soil
Temperature	—	daily average, 20 ± 2 °C; instantaneous, 20 ± 3 °C
Lighting	—	incandescent or fluorescent; intensity, 400–800 lux adjacent to surface of soil in test chamber; fixed photoperiod (e.g., 16 L:8 D or 12L:12D)
Feeding	—	cooked oatmeal (recommend Quaker Oats™ "quick" 3–5 min oatmeal); 5 mL (= 1 teaspoonful) per test chamber each feeding; placed in a shallow depression in the centre of the soil surface in each test chamber on Days 0, 14, 28, and 42 only
Measurements during test	—	temperature in test facility, daily or continuously; soil moisture content and pH in each treatment/concentration, at start and end
Observations during test	—	total number of live adult worms in each test chamber on Days 0 and 28; optionally, number of live and dead adults worms on surface of the soil in each test chamber at the start of the test (t = 1 h) and on Day 28; presence of uneaten food in surficial layer of soil in each test chamber on each feeding occasion; number of live juvenile worms in each test chamber on Day 56; obvious pathological symptoms (e.g., open wounds) or distinct behavioural abnormalities (e.g., lethargy) for worms in each test chamber
Biological endpoints	—	total number of live adult worms in each replicate (i.e., in each test chamber) on Day Day 28; total dry wt and number of live juvenile worms in each replicate on Day 56

Statistical endpoints	—	mean (± SD) percent survival of adults in each treatment, on Day 28; mean (± SD) number of live juveniles in each treatment, on Day 56; mean (± SD) dry wt of live juveniles in each treatment, on Day 56; if multi-concentration test: 28-day LC50 for adults worms, ICp for numbers of live juveniles produced in each concentration during 56-day test, and ICp for mean dry wt of individual worms surviving in each concentration at test end
Test validity	—	invalid if mean 28-day survival of adults in negative control soil <90%; invalid if mean reproduction rate for adults in negative control soil <3 live juveniles/adult; invalid if mean dry wt of individual live juveniles in negative control soil at test end <2.0 mg
Test with reference toxicant	—	must perform once/month, or in conjunction with definitive test(s) with soil samples; use boric acid; prepare and test ≥5 concentrations plus a negative control, using artificial soil as substrate; ≥3 replicates/concentration and 5 worms/replicate; follow procedures and conditions for seven-day soil toxicity tests described in Section 4.1.8; determine seven-day LC50 (including 95% confidence limits); express as mg boric acid/kg, dry wt; also recommend eight-week tests with boric acid performed according to Section 4.3 at least twice a year or in conjunction with definitive eight-week test

Field-Collected Soil

Transport and storage	—	seal in plastic and minimize air space; transport in darkness (e.g., using an opaque cooler, plastic pail, or other light-tight container); do not freeze or overheat during transport; store in dark at 4 ± 2 °C; test should start within two weeks, and must start within six weeks unless soil contaminants are known to be stable
Negative control soil	—	either natural, uncontaminated field-collected soil or artificial soil, for which previous 56-day tests with *E. fetida* showed that all criteria for test validity could be regularly met
Reference soil	—	one or more samples for tests with field-collected soil; taken from site(s) presumed to be clean but in the general vicinity of sites where test soils are collected; characteristics including percent organic matter, particle size distribution, and pH similar to test soil(s)
Characterization of test soils	—	at least percent moisture, WHC, pH, conductivity, percent TOC, percent organic matter, and particle sizes (% sand, % silt, % clay); optionally, contaminants of concern (e.g., metals, PAHs, pesticides)
Preparation of test soils	—	if necessary, remove debris and indigenous macro-organisms using forceps; if necessary, press through a sieve of suitable mesh size (e.g., 4–6 mm); mix; determine percent moisture content; hydrate with de-ionized or distilled water (or, if and as necessary, dehydrate) to the optimal percentage of its WHC (see Section 5.3); mix; dilute with control or reference soil if multi-concentration test; ensure homogeneity

Soil Spiked with Chemical(s) or Chemical Product(s)

Negative control soil	—	recommend artificial soil or a *clean* field-collected soil
Characterization of chemical(s) or chemical product(s)	—	information on stability, water solubility, vapour pressure, purity, and biodegradability of chemical(s) or chemical product(s) spiked into negative control soil should be known beforehand

Solvent	—	de-ionized water is the preferred solvent; if an organic solvent is used, the test must include a solvent control
Preparation of mixtures	—	procedure dependent on the nature of the test substance(s) and the test design and objectives; chemical/soil mixtures may be prepared manually or by mechanical agitation; test substance(s) may be added as measured quantities in solution (i.e., in water or an organic solvent) or as a solid material comprised partly or completely of the test substance(s); ensure homogeneity
Concentration within soil mixture of chemical(s) or chemical product(s) added	—	normally measure at beginning and end of test, in high, medium, and low strengths as a minimum

Universal procedures for performing this test for effects on the survival, reproduction, and *growth* of earthworms are described herein. This is a whole soil toxicity test, with no renewal of test soils during the 56-day (eight-week) test duration. The test begins with adult earthworms taken from laboratory cultures of *E. andrei*.[50,51] The experimental design involves 10 replicate test chambers per treatment, with two adult worms added to each test chamber. Following a 28-day (four-week) exposure of adult worms of this species, numbers surviving in each test chamber and treatment are determined and recorded and they are discarded[52], while continuing the test for a further 28 days to measure effects on the growth of their progeny (i.e., juvenile worms). Food for the adult worms and their progeny is provided throughout the eight weeks of the test.

4.3.1 Beginning the Test

The test is performed using 500-mL glass jars as test chambers (Section 3.1.4). Each test chamber (see Section 3.1.4) placed within the test facility must be clearly coded or labelled to enable identification of the sample or its concentration. The date and time when the test is started must be recorded, either directly on the labels or on separate data sheets dedicated to the test. The test chambers should be positioned for ease while making observations and measurements. Treatments should be positioned randomly within the test facility (EC, 1997a, b, 2001).

The day that animals are initially exposed to samples of test materials or substances is designated Day 0. On the day preceding the start of the test (i.e., Day -1), each sample or subsample of test soil or similar particulate material, including *negative control soil* and, if used, *reference soil*, should be mixed thoroughly[53] (see Sections 5.3 and 6.2) to provide a

[50] This test method is presently restricted to *E. andrei*, since a criterion for test validity based on a minimum acceptable dry weight of juvenile worms at test end has not been established for *E. fetida* using the feeding regime required herein (see Sections 4.3.3 and 4.3.4).

[51] These worms should be obtained from cultures maintained in the laboratory using the guidance provided in Section 2.3. In situations where a testing laboratory is unable to provide test organisms using in-house cultures, the worms may be obtained as juveniles or adults from another source which uses culturing conditions, procedures, and *quality assurance* consistent with the guidance in Section 2.3. In this situation, however, the worms must be held and acclimated to laboratory conditions within a facility provided by the testing laboratory, according to the guidance in Section 2.4, before their use in this biological test method. The use of earthworms collected from the field or purchased from a commercial supplier with unknown or no *quality assurance* or *quality control* is unacceptable for purposes of this test method.

[52] If body or tissue residues of contaminants are of interest, the worms might be frozen (-20 °C) for analyses at a later date instead of being discarded.

[53] Any liquid that has separated from a sample or subsample of test soil during transport and/or storage must be remixed into the sample.

homogeneous mixture consistent in colour, texture, and moisture. If field-collected samples of *site soil* are being prepared for testing, large particles should be removed before mixing, along with any vegetation or macroinvertebrates observed (see Section 5.3).

The quantity of each test soil mixed as a *batch* should be enough to set up the replicates of that treatment (see Table 5) plus an additional amount for the physicochemical analyses to be performed (Section 4.3.5) and a surplus to compensate for the portion that adheres to the sides of the mixing chamber after removing aliquots for the testing and analyses. The moisture content of each test soil should be known or determined, and adjustments made as necessary by mixing in *test water* (or, if and as necessary, by air drying the sample) until the desired moisture level is achieved (see Sections 5.3 and 6.2). Quantitative measures of the homogeneity of a *batch* might be made by taking aliquots of the mixture for measurements such as particle size analysis, total organic carbon content (%), organic matter content (%), moisture content (%), and concentration of specific chemicals.

Immediately following the mixing of a *batch*, an identical wet weight of soil equivalent to a volume of ~350 mL should be transferred to each replicate test chamber[54] (Section 3.1.4). The soil added to each test chamber should be smoothed (but not compressed) using a spatula or by gently tapping the glass jar against the bench top, three times.

For each treatment (e.g., a particular *site soil* tested at 100% concentration only; or a particular concentration of test soil or chemical-spiked soil mixed with *clean* soil), 10 replicate test chambers should be set up using aliquots of the same *batch* with identical wet weights (equivalent to ~350 mL/aliquot). Ten replicate test chambers containing aliquots of *negative control soil* (Section 3.2) from the same *batch* must be included in each test. For a multi-concentration test, at least seven concentrations plus the control treatment(s) must be set up, and more (i.e., ≥10 plus controls) are recommended.[55] In some instances, preliminary or range-findings tests might be warranted in which instance the number of replicates per concentration might be reduced (see Section 6.2).

It is recommended that a minimum of one additional test chamber containing *negative control soil* and one additional test chamber containing *reference soil* and/or the lowest concentration of test soil (if a multi-concentration test) be included in the test. These extra replicates, for which data are not included in the analyses and no reporting requirements pertain, are useful in providing a preliminary assessment as to whether or not acceptable production of young in these treatments has occurred by Day 28 (see Sections 4.3.3 and 4.3.5). If acceptable production of young in these treatments has not occurred by Day 28, the investigator may choose to extend the duration of exposure of adult earthworms in the definitive test chambers from 28 days to 35 days (see Section 4.3.5), in which instance the test duration would be 63 days rather than 56 days.

Following the addition of a measured (identical wet wt; volume ~350 mL) aliquot of test soil to each test chamber, an unperforated cover (see Section 3.1.4) should be placed over each test chamber to minimize moisture loss. The test chambers should be held overnight under test temperature and lighting conditions (Section 4.3.2) for chemical equilibration of the test soils. On Day 0 (i.e., when starting the test), each cover should be perforated.[56]

---

[54] The wet weight of soil required to achieve a volume of ~350 mL depends on the moisture content, bulk density, and other characteristics of the soil; and will vary from sample to sample. Accordingly, the wet weight of each sample required to achieve this volume should be determined by transferring that amount of sample required to fill a preweighed (or tared) 500-mL glass beaker or jar to a 350-mL mark scribed on its side, after gently smoothing (not compressing) the surface of the soil at this mark. Thereafter, the wet weight of that quantity should be determined and recorded, and an identical wet weight transferred to each replicate test chamber.

[55] The use of 10 or more concentrations (plus the controls) is advised. A large number of test treatments is necessary to show the shape of the concentration-response relationship and to choose the appropriate linear or nonlinear regression model (see Section 6.4.2.1). Use of 10 or more concentrations is particularly prudent if the investigators wish to determine a 28-day LC50 for the adult worms, as well as ICps for reproductive and growth inhibition (see Section 4.3.7). In certain tests for effects of prolonged exposure, the investigators might wish to focus on the sublethal endpoints and not derive a 28-day LC50, in which instance 7–9 test concentrations (plus the controls) might prove adequate for this purpose.

[56] For a test involving a sample of contaminated soil with volatile compounds, it is recommended that the covers not be perforated during the first week of the test to minimize gaseous emissions and to increase the exposure of worms to these volatile compounds. In this instance, the covers should be perforated on Day 7 (Stephenson *et al.*, 2001).

Test organisms (see Section 2.3.8) are transferred to each test chamber the next day (i.e., Day 0 of the toxicity test). Adult (fully clitellated) worms which are within the acceptable size range (i.e., individual wet wt, 250–600 mg) must be used for this test. The adult worms used in the test must previously have been held in a facility within the testing laboratory for a minimum of seven days, during which time they are acclimated to *negative control soil* under feeding, temperature, and lighting conditions identical to those used in the test (see Sections 2.3.8 and 2.4 including 2.4.1 and 2.4.8). A number in excess of those required for the test should be removed from a culture chamber (Section 2.3) or holding/acclimation chamber (Section 2.4). Worms chosen for use in the test should be as similar in size (i.e., initial wet wt) as possible, based on the range of individual wet weights within the culture from which they are selected. Only those worms appearing healthy, similar in colouration, and active when removed from the bedding substrate should be selected. Those worms should be selected from this chamber, removed by gloved hand or using the blunt arm(s) of rounded forceps, and transferred briefly to a clean, shallow dish or tray where they are quickly rinsed in clean test water (i.e., de-ionized or distilled water). Thereafter, these worms are placed into a transfer container (e.g., a glass or aluminum tray measuring ~10 × 10 cm) lined with paper towel dampened with test water. A final observation should be made of the worms in this container to confirm that their appearance is normal. Any atypical worms should be discarded. Thereafter, individual worms of as similar size as possible should be carefully selected while confirming that they are within the acceptable size range, and then transferred individually (by hand or using the blunt arms of rounded forceps) to the test chambers. The order of adding earthworms to test chambers should be randomly allocated with respect to treatment.

Worms are placed onto the surface of the test soil in each test chamber; two per chamber. The number of worms not burrowed into the soil in each jar after 1 h following their introduction should be noted and recorded, for each test chamber.[57] Individual wet weights for a minimum of twenty worms must be measured and recorded when the worms are introduced to the test chambers, to determine the variability in initial size of worms used in the test. These weights may either be based on surplus worms that are from the group selected for use in the test, or on the weights of individual worms representing the various treatments as they are weighed and transferred to the test chambers. The mean (± SD) weight for these worms must be calculated and reported (Section 7).

4.3.2 Test Conditions

- This is a 56-day (eight-week) whole soil toxicity test[58], during which the soil in each test chamber is not renewed.

- The test chamber is a 500-mL wide-mouthed glass jar. Its contents (i.e., a 350-mL volume of test soil) should normally be covered with a piece of perforated transparent or translucent material (e.g., Saran wrap™ or Parafilm™) (Section 3.1.2) secured to the lip of the jar using a rubber band. If the test material is known to contain volatile compounds (e.g., PAHs), the use of opaque aluminum foil as covers is recommended, together with side lighting sufficient to achieve the minimal light intensity required at the surface of the soil (see last bullet in 4.3.2).

- Ten test chambers (with two adult worms per chamber) are required for each treatment, when setting up this test. For a multi-concentration test, at least seven concentrations plus the appropriate control treatment(s) must be set up using 10 replicates per concentration, and more concentrations (i.e., ≥10 plus controls) are recommended.

- The test must be conducted at a daily mean temperature of 20 ± 2 °C. Additionally, the instantaneous temperature must always be 20 ± 3 °C.

All test chambers, including those containing negative control soil, must be treated identically.

[57] A lack of burrowing might reflect an avoidance response by the worms. It could also indicate their poor condition at the start of the test. A comparison of the mean (± SD) percentage of worms burrowing in *negative control soil* (and, if used, *reference soil*) during the first hour of the test, versus percent worms burrowing in each test soil at that time (or thereafter; see Section 4.3.5), would provide insight into the possibility that the worms are showing an avoidance response to one or more of the test treatments.

[58] The investigator may choose to extend the duration of the test to 63 days (9 weeks) (see Sections 4.3.1 and 4.3.5).

- Test chambers must be illuminated with a fixed daily photoperiod (e.g., 16-h light and 8-h dark, or 12-h light and 12-h dark), using incandescent or fluorescent lights. The photoperiod chosen should be the same as that to which the worms are acclimated before the test (see Sections 2.3.3 and 2.4.3). Light intensity adjacent to the surface of the soil in each test chamber should be 400–800 lux, and must be at least 400 lux as a minimum. This range is equivalent to a quantal flux of 5.6–11.2 μmol/(m^2 · s) for cool-white fluorescent, 6.4–12.8 μmol/(m^2 · s) for full-spectrum fluorescent, or 7.6–15.2 μmol/(m^2 · s) for incandescent.

- Worms in each test chamber must be fed an identical quantity (i.e., 5 mL = 1 teaspoonful) of hydrated, cooked oatmeal (see Section 4.3.4), on Days 0, 14, 28, and 42 only.

4.3.3 Criteria for a Valid Test

For the results of this biological test method to be considered valid, each of the following must be achieved:

- the mean survival rate for the adult worms held in *negative control soil* for 28 days must be ≥90%.[59, 60]

- the reproduction rate for the adult worms in *negative control soil* must average at least three live juveniles per adult.[60]

- the mean dry weight of individual live juveniles in *negative control soil* must be ≥2.0 mg at test end.

4.3.4 Food and Feeding

Worms in each test chamber are normally fed a measured quantity of hydrated, cooked oatmeal (see Section 2.3.6) on each of the following days of the test, only: Day 0, Day 14, Day 28, and Day 42.[61] The quantity of food added to each test chamber on each of these occasions must be an identical volume (i.e., 5 mL = 1 teaspoonful) of hydrated, cooked oatmeal with a dry weight equivalent to 230–250 mg dry oatmeal (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). This is prepared fresh on each feeding occasion, by adding one-third volume of dried oatmeal (recommend Quaker Oats™ "quick" 3–5 min oatmeal; see Section 2.3.6) to two-thirds volume of de-ionized water, and cooking in a microwave for 3–5 minutes only using the "high" temperature setting, followed by cooling (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). Alternatively, the dried oatmeal may be boiled in de-ionized water for 3–5 minutes using a hot plate or stove element, during which time it is stirred periodically. The mixture should then be cooled and stirred while removing each 5-mL aliquot to be added to each test chamber. The recommended consistency of cooked oatmeal to be used is a "viscous slurry"; additional water can be added to the cooked oatmeal (or the cooking period extended as necessary) to achieve this viscosity.

On Day 0, before adding the earthworms to the test chambers, a small hollow should be made in the centre of the soil surface within each chamber. A 5-mL (= 1 teaspoonful) volume of cooked oatmeal should be placed into this depression, and covered with a thin layer of surrounding soil to reduce fungal growth. The adult earthworms should then be added, two/chamber (see Section 4.3.1). On Day 14, the cover of each test chamber is removed and an additional 5-mL aliquot of cooked oatmeal added (as per the procedure for Day 0). On Day 28, and following the removal of adult worms and the return of the remainder of the contents of the jar to each test chamber (Section 4.3.5), another 5-mL aliquot of cooked oatmeal should be added to each test chamber (for development and growth of their progeny) in the same manner as before. A final

---

[59] An investigator may choose to extend the duration of exposure of adult worms to 35 days (see Sections 4.3.1 and 4.3.5), in which instance this survival rate also applies.

[60] ISO (1998) and OECD (2000) used this (or an equivalent value) as a criterion for a valid test for effects of chemicals on the reproduction of *E. andrei/fetida*.

[61] If, on Days 14 and/or 28 only, uneaten food (i.e., residual oatmeal bolus) is evident within the surficial layer of the soil in certain or all test jars representing any treatment, no additional food should be added to those test chambers whereas all other test chambers with no food evident should receive the normal supplement (i.e., 5 mL of cooked oatmeal). The discretionary practice of skipping feeding(s) at these times, on an individual (jar-by-jar) basis, is advisable to avoid the risk of overfeeding and the risk of excessive mould or sorption of toxic contaminants caused by uneaten food, which can be "chamber specific". For any treatment containing or comprised of artificial soil, feeding should without exception be provided on each of the scheduled days (i.e., on Days 0, 14, 28, and 42), since the artificial soil is otherwise deficient in natural organic food.

5-mL aliquot of cooked oatmeal should be added to each test chamber on Day 42. When adding food to the test chambers on Days 14, 28, and 42, any old food evident in the surficial layer of the soil within each test chamber should be left undisturbed (since hatchling worms frequently burrow within it).

4.3.5 Observations and Measurements During the Test

The condition, appearance, and number of live worms transferred to each test chamber on Day 0 must be observed and recorded. At 1 h following their transfer, the number of worms on the surface of the soil in each jar, or against the glass on the inner sides or bottom of each test chamber, should be noted and recorded (see Section 4.3.1).

At the time of each feeding (see Section 4.3.4), the cover should be removed from each test chamber, and observations and records made of the number of live or dead worms on the surface of the soil therein. Thereafter, the surficial layer of the soil in each test chamber should be examined to appraise the presence and quantity of any uneaten food and to reach a decision regarding the need to skip a feeding (see Section 4.3.4). Observations and records should also be made at this time of the number of worms seen inside each jar on its glass sides or bottom (this might be an indication of an avoidance response to the soil therein).

On Day 28 of the test, the covers of any "extra" test chambers used to determine if acceptable production of progeny in these treatments has occurred by this time (see Section 4.3.1) should be removed.[62] The contents of these "extra" test chambers should be examined for the presence of cocoons or juvenile worms. If any cocoons and/or juveniles are observed in each of these treatments, the cover of each of the definitive test chambers should be removed and its contents examined (see next paragraph). If cocoons or juvenile worms are not observed in the "extra" test chambers representing each of these treatments, it is recommended that the definitive test chambers be left undisturbed for an additional seven days before their examination for and removal of adults (ESG, 2001, 2002; ESG and Aquaterra Environmental, 2002). In this instance, the contents of the "extra" test chambers (including all cocoons, and any live juvenile and adult worms) should be replaced and held under test conditions until they are re-examined.

Thereafter (i.e., on Day 28 or, in some instances, Day 35; see preceding paragraph), the cover of each definitive test chamber must be removed, as should the covers of each "extra" test chamber. The number of live and dead adult worms on the surface of the soil or against the glass on the inner sides or bottom of each test chamber should be observed and recorded. Thereafter, the contents of each test chamber must be transferred to a sorting tray or plastic sheeting, and the number of live and dead adult worms counted and recorded. Adults appearing to be dead should be touched gently on their anterior end with a glass rod or spatula; absence of any response is defined as death. Missing adults must be counted as dead. The appearance (e.g., normal or signs of discolouration or lesions) and behaviour (e.g., normally active or lethargic) of each surviving adult should be noted and recorded. Immediately after this evaluation, all adults (live and dead) must be discarded, and the test soil returned to the jar together with any cocoons and juvenile worms therein. The cover of each test chamber must be replaced. Observations of the number of juvenile worms reproduced during the test, and their *growth* (final dry weight of individuals), appearance, and behaviour, are made at the end of the test (i.e., on Day 56 or, in some instances, on Day 63).

Air temperature in the test facility (Section 3.1) must be measured daily (e.g., using a maximum/minimum thermometer) or continuously (e.g., using a continuous chart recorder).

For each treatment, the contents of one or more replicate test chambers should be examined weekly for apparent "wetness". If, for any treatment, the soil appears to be too dry at any time during the test, all replicates representing that treatment should be examined and the surface of the soil in each test chamber that appears to be too dry should be moistened by spraying it with de-ionized water using a fine-spray mister which dispenses about 1 mL of water per spray.[63] On the occasions that food is added (i.e.,

[62] These "extra" test chambers might include one or more additional jars containing negative control soil, one or more jars containing reference soil (if included in the definitive test), and one or more jars containing the lowest concentration of a test soil if a multi-concentration test.

[63] The apparent "wetness" of a soil is affected by the degree of activity of the worms in the soil, as well as by the nature of the soil and the amount of water lost from

on Days 0, 14, 28, and 42), the apparent "wetness" of each test soil is judged at that time, and water sprayed onto the surface of each replicate as necessary.

The pH and moisture content of the test soil representing each treatment (including the *negative control soil* and, if used, *reference soil*) must be measured and recorded at the beginning and end of the test. Additionally, it is recommended that conductivity be measured at the beginning and end of the test in instances where the test soil is anticipated to have a high salt content. The initial measurements should be made using a composite sample made up of subsamples of each *batch* of test soil used to set up replicates of a particular treatment (see Section 4.3.1). The final (i.e., Day-56 or, in certain instances, Day-63) measurements should be made using subsamples of the replicates of each treatment to which worms were exposed, following the end-of-test observations of worm distribution, survival, appearance, and behaviour (see Section 4.2.6). Measurements of soil pH and moisture content should be made according to the guidance in Section 4.1.5.

The test soils might be analyzed for concentrations of chemical(s) or chemical product(s) of concern. Guidance in Section 4.1.5 applies here. Sections 5.4 and 6.3 should be consulted for further advice.

4.3.6 Ending the Test

To terminate this soil toxicity test, the number of live juvenile worms observed on the surface of the soil in each definitive test chamber, or adjacent to the glass on its sides or bottom, should be determined and recorded on Day 56 (or, in some instances, on Day 63; see Section 4.3.5). Thereafter, the number of live juveniles within each test chamber must be counted and recorded.

As part of the developmental studies associated with the standardization of this biological test method, Stantec and Aquaterra Environmental (2004) designed a novel "heat-extraction" procedure for recovering surviving juveniles from the soil in each test chamber, at the end of the test. This heat-extraction procedure proved efficient and effective, and is recommended for routine use with this test method.[64] To perform this recovery procedure, the test jars are transferred sequentially to a heated (40–45 °C) water bath, while ensuring that the level of the water in the bath does not rise higher than half of the height of the soil in each jar. Each test jar is left in the water bath for no more than 15 minutes. Thereafter, the surficial 2-cm layer of soil in the jar is carefully removed, and placed onto a sorting tray or plastic sheeting. This subsample of the test soil is sorted using blunt-nosed forceps, while recovering all surviving juvenile worms. Worms appearing to be dead should be touched gently on their anterior end with a glass rod or spatula; absence of any response is defined as death. Dead worms are discarded.

For each test chamber, the appearance (e.g., normal or signs of discolouration or lesions) and behaviour (e.g., normally active or lethargic) of the surviving worms should be noted and recorded. Thereafter, the surviving juveniles recovered from each test chamber must be rinsed briefly in de-ionized or distilled water to remove any adhering soil. The rinsed worms from each test chamber must then be transferred as a group to a damp paper towel or blotting paper. They should then be transferred to a clean aluminum weighing pan

---

test chambers due to evaporation. Typically, as much as 3 mL per test chamber can be lost weekly due to evaporation. However, weekly additions of water frequently result in the soil being too wet at test end. Any decision as to whether or not to spray water onto the surface of the soil in each test chamber should be made based on apparent "wetness" of the test soil, during each weekly period of observation (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). To assess this, a qualitative "squeeze test" (see Section 5.3) should be applied to a small quantity (i.e., a "pinch") of the surficial soil within the test chamber. If no water appears, the soil is likely too dry. In this instance, the surface of the soil in the test chamber should be lightly misted.

[64] The use of heat, to improve the efficiency of recovering juvenile earthworms from test chambers at the end of an eight-week test, was recommended by Dr. Kees van Gestel (Institute of Ecological Science, Amsterdam, Netherlands). This procedure causes the surviving worms to move to the cooler soil above the height of the water in the water bath, enabling their efficient recovery. The advantage of using this procedure is that the amount of soil to sort through is appreciably less than the full quantity (~350 mL) in the test chamber. Stantec and Aquaterra Environmental (2004) found that the time required to recover surviving earthworms from each test chamber, while manually sorting the soil, was reduced from about 30–40 minutes per test chamber (if the entire contents were sorted) to only 15–20 minutes per test chamber using the heat-extraction procedure. The percent recovery of earthworms using this procedure was typically 100%, and in all instances ≥94% of the surviving worms were recovered by the heat-extraction technique (Stantec and Aquaterra Environmental, 2004).

that has been previously numbered, weighed, and held in a desiccator.[65]

Observations of numbers of unhatched or hatched cocoons at test end, although not included as a test endpoint (see Section 4.3.7), might prove useful for certain tests in discerning adverse effects on (delayed) development of earthworms or early survival of young produced. To enable these (optional) observations, the number of hatched or unhatched cocoons found within the contents of each test chamber should be counted and recorded.[66] The heat-extraction procedure for recovering test organisms, described previously, is not suitable for any test involving the recovery of cocoons. In this instance, the entire contents of the soil within each test chamber must be carefully sorted while recovering and counting the numbers of hatched and unhatched cocoons as well as the number of live juvenile earthworms.

Separate weighing pans, each containing the group of surviving juveniles recovered from each test chamber, are placed into an oven and dried at 90 °C until a constant weight is achieved (this usually takes a minimum of 48 h) (Aquaterra Environmental and ESG, 2000). Upon removal from the oven, the weighing pans are moved immediately to a desiccator. Following cooling, each weighing pan should be individually and randomly removed from the desiccator, and weighed immediately[67] to the nearest 0.1 mg on a balance that measures accurately to this limit. Mean dry weight per surviving juvenile worm is calculated for each group (Section 4.3.7).

During the series of dry-weight determinations for the groups of surviving juveniles from a test, the first weighing pan should be replaced in the desiccator and weighed again at the end of all weighings. This serves as a check on any sequential gain of water by the weighing pans in the desiccator over time, which can occur when each weighing pan is removed for its weight determination. The change in weight of the first weighing pan over time should not be >5%; if it is, redrying of all weighing pans for ≥2 h and reweighing should be carried out.

Following the recovery of juvenile worms from each test chamber, subsamples of each test soil (including the *negative control soil* and, if included in the test, *reference soil*) should be taken for determinations of pH and moisture content (Section 4.3.5). Analyses for other chemical constituents (i.e., concentrations of contaminants) should also be made at this time using representative subsamples of each test soil (Section 4.3.5).

4.3.7 Test Endpoints and Calculations

For each test, the mean (± SD) percent survival for all replicate groups (n = 10) of adult worms exposed to each treatment for 28 days must be calculated and reported. For a single-concentration test (see Sections 5.3 and 6.2), the mean (± SD) value for the percent survival of adults on Day 28, as determined for each treatment, is compared with that for the sample(s) of *reference soil* or, as necessary and appropriate, compared with that for the *negative control soil*. For a multi-concentration test (see Sections 5.3 and 6.2), the 28-day LC50 for the adult worms must be calculated and reported if sufficient data are available to enable this. Environment Canada (2004b) provides guidance for calculating LC50s, which should be followed; Section 6.4.1 gives further guidance in this regard.

The reproductive endpoint for this test is based on the number of surviving juveniles produced in each replicate and each treatment during the 56-day test period. A significant reduction in this number is considered indicative of an adverse toxic effect of the treatment on the reproductive success of the adult worms. For a single-concentration test (see Sections 5.3 and 6.2), the mean (± SD) value for number of surviving juveniles in the test soil on Day 56 is determined and compared to that for the sample(s) of *reference soil* or, as necessary and appropriate, compared to that for the *negative control soil*. A *Student's t-test* or other appropriate statistic (EC, 2004b) should be used for this comparison. For a multi-concentration test (see Sections 5.3 and 6.2), the

---

[65] If any deposits (e.g., wax) associated with the weighing pans are cause for concern with respect to providing weighing errors, the weighing pans should be oven-dried for at least 48h to achieve a constant weight (EC, 1997a, b, 2001).

[66] Hatched cocoons are hollow with a translucent appearance, and are easily dented or collapsed by pressing gently with forceps. Conversely, unhatched cocoons usually are turgid, have an opaque appearance, and their exterior covering is not easily dented with gentle forceps pressure although they will pop open with greater pressure.

[67] The dried worms can take up water vapour readily, so weighing should be rapid and the time standardized among weighing pans.

56-day *ICp* for reproductive inhibition must be calculated and reported (data permitting).[68] Environment Canada (2004b) provides direction and advice for calculating ICps, which should be followed; Section 6.4.2 (including Appendix J) gives further guidance in this regard. Initially, regression techniques (see Section 6.4.2.1) must be applied to multi-concentration data intended for calculation of an ICp.[69] In the event that the data do not lend themselves to calculating the 56-day ICp for reproductive inhibition using the appropriate regression analysis (see Appendix J), linear interpolation of these data using the program ICPIN should be applied in an attempt to derive an ICp (see Section 6.4.2.2).

The *growth* endpoint for this test is based on the mean dry weight of individual juvenile worms produced in each treatment which survived the 56-day test period. A significant reduction in this weight is considered indicative of an adverse toxic effect of the treatment on the growth of surviving juveniles produced by the adult worms used to start the test. For a single-concentration test (see Sections 5.3 and 6.2), the mean (± SD) value for dry weight of individual juveniles surviving in the test soil on Day 56 is determined and compared to that for the sample(s) of *reference soil* or, as necessary and appropriate, compared to that for the *negative control soil*. A *Student's t-test* or other appropriate statistic (EC, 2004b) should be used for this comparison. For a multi-concentration test (see Sections 5.3 and 6.2), the 56-day ICp for growth inhibition (i.e., decreased mean dry weights of individual juveniles) must be calculated and reported (data permitting).[68] The direction and advice in Section 6.4.2 for calculating ICps is applicable and should be followed. In this regard, the approach described in the preceding paragraph for calculating a 56-day ICp for reproductive inhibition applies equally here, when calculating a 56-day ICp for growth inhibition.

4.3.8 Tests with a Reference Toxicant

The routine use of a reference toxicant is necessary to assess, under standardized test conditions, the relative sensitivity of a portion of the population of adult earthworms within a particular culture (Section 2.3.9) or a particular acclimation chamber (Section 2.4.9) from which test organisms are selected for use in a test for effects of prolonged exposure on the survival, reproduction, and growth of earthworms. Tests with a reference toxicant also serve to demonstrate the *precision* and reliability of data produced by the laboratory personnel for that reference toxicant, under standardized test conditions.

A seven-day acute lethality test with a reference toxicant, conducted according to the procedures and conditions described in Section 4.1.8 (and summarized in Table 5), must be initiated at the start of each 56-day test for effects of contaminated or potentially contaminated soil on the survival, reproduction, and growth of earthworms. This reference toxicity test must be performed using adult *E. andrei* taken from the same population as those used to start the 56-day test (see Sections 2.3.9 and 2.4.9). Boric acid is the recommended reference toxicant for this seven-day test (see Section 4.1.8).

Besides performing acute lethality tests with a reference toxicant, it is recommended that any laboratory performing eight-week tests with samples of contaminated (field-collected or chemical-spiked) soil conduct one or more eight-week tests with their culture(s) of *E. andrei* using a reference toxicant. In keeping with the guidance in OECD (2000), these tests should either be performed at least twice a year or, where the eight-week testing of contaminated soil is carried out at a lesser frequency, in parallel with each

---

[68] Historically, investigators have frequently analyzed *quantitative* sublethal data from multi-concentration tests by calculating the *no-observed-effect concentration (NOEC)* and the *lowest-observed-effect concentration (LOEC)*. Disadvantages of these statistical endpoints include their dependence on the test concentrations chosen and the inability to provide any indication of *precision* (i.e., no 95% or other confidence limits can be derived) (NERI, 1993; EC, 2004b). Given these disadvantages, ICp is the required statistical endpoint for reproductive and growth data derived in a multi-concentration test using *E. andrei*.

[69] Regression is the method of choice for estimating ICp. It involves fitting the data mathematically to a selected model and then calculating the statistical endpoint using the model that best describes the exposure-concentration response relationship. Nonlinear regression techniques were originally recommended by Stephenson *et al.* (2000b) for several reasons including: the relationship that exists between exposure concentration and earthworm growth or reproduction responses is typically nonlinear; the heteroscedasticity of the data cannot be reduced by transformation; the more standard bootstrap simulation technique has several limitations for these types of data; and nonlinear regression can fit effect distributions showing hormesis. By using standard mathematical techniques, a regression can be well described in terms that convey useful information to others, effects at high and low concentrations can be predicted, and confidence intervals can be estimated. Deficiencies of the smoothing and interpolation method can be largely remedied (EC, 2004b).

definitive soil toxicity test. The procedures and conditions to be applied to these eight-week reference toxicity tests should be consistent with those described in Section 4.3 herein. Any endpoint data (i.e., 28-day LC50 and/or ICp; see Section 4.3.7) should be compared with values obtained in the past by that laboratory for the reference toxicant. This testing and comparison is useful to provide assurance that the laboratory's test conditions and procedures when performing an eight-week test are adequate, and to verify that the eight-week response of the earthworms to the reference toxicant has not changed appreciably from that for earlier eight-week tests with this chemical performed at the testing facility. Boric acid spiked in
artificial soil is the recommended reference toxicant for this eight-week test.[70, 71]

---

[70] Results for a number of eight-week (or, in some instances, nine-week) reference toxicity tests with boric acid, performed by Stantec and Aquaterra Environmental (2004) according to the biological test method described herein in Section 4.3, demonstrated similar findings for *E. andrei*. In two tests initiated using adult earthworms from asynchronous or synchronous cultures, the 35-day LC50s for these worms were 2706 or 3207 mg boric acid/kg dry wt of artificial soil, respectively (Stantec and Aquaterrra Environmental, 2004). Data for number of live progeny generated during these and two additional tests performed with worms from asynchronous or synchronous cultures according to Section 4.3 yielded IC50s ranging from 270–568 mg boric acid/kg dry wt of artificial soil, and IC20s ranging from 163–425 mg/kg. Data for dry weights of surviving progeny generated during these four tests yielded IC50s ranging from 147–948 mg boric acid/kg dry wt, and IC20s ranging from 23–414 mg/kg. Results for side-by-side tests performed using worms from asynchronous or synchronous cultures showed that the differences among respective statistical endpoints were, in each instance, not large and the 95% confidence limits overlapped (Stantec and Aquaterra Environmental, 2004). The following concentrations of boric acid were used by Stantec and Aquaterra Environmental (2004) to calculate both lethal and sublethal endpoints during prolonged-exposure tests with this reference toxicant: 0, 7, 14, 28, 56, 113, 225, 450, 900, 1800, and 3600 mg/kg soil (dry wt). An expanded range (based on a logarithmic series of concentrations; see Appendix I) that includes one or two higher test concentrations is recommended for future tests intended to calculate both lethal and sublethal endpoints. For tests restricted to sublethal endpoints, the following concentrations of boric acid proved adequate when calculating IC50s and IC20s for number of live progeny and their dry weights at test end: 0, 10, 16, 30, 50, 100, 300, 560, and 1000 mg/kg soil (Stantec and Aquaterra Environmental, 2004.

[71] Section 6.2 "*Preparing Text Mixtures*" includes an example showing the amounts of de-ionized or distilled water and boric acid to be added to dry artificial soil, to prepare a given treatment for a reference toxicity test with a specific concentration of boric acid in artificial soil. The calculations in this example show the amount of water necessary to adjust the moisture content of the artificial soil to a fixed percentage (i.e., 70%) of the soil's water-holding capacity, while taking into account the volume of the stock solution of boric acid as part of the overall adjustment for soil moisture content.

Specific Procedures for Testing Field-Collected Soil or Similar Particulate Material

This section provides specific instructions for preparing and testing samples of field-collected (site) soil or similar particulate material, in addition to the procedures discussed in Section 4.

Detailed guidance for the collection, handling, transport, storage, and analyses of field-collected soil is given in a number of reports specific to these subjects (e.g., van Ee *et al.*, 1990; Webster and Oliver, 1990; USEPA, 1991; Keith, 1992; Klute, 1986; Carter, 1993; OMAFRA, 1999). In the absence of guidance specific to these subjects from Environment Canada, such reports should be consulted and followed (in addition to the guidance provided here), when collecting samples of field-collected soil and preparing them for toxicity tests with earthworms using any of the biological test methods herein.

5.1 Sample Collection

Crépin and Johnson (1993) provide a useful summary of field-sampling design and appropriate techniques for sample collection. Field surveys of soil toxicity using biological tests with earthworms and/or other suitable, soil-associated test organisms (e.g., EC, 2004d, e) are frequently part of more comprehensive surveys (e.g., Callahan *et al.*, 1991; van Gestel, 1992; Heimbach, 1993; Holmstrup, 2000; Saterbak *et al.*, 2000; van Gestel *et al.*, 2001). Such surveys could include a *battery of toxicity tests* to evaluate the toxicity of soil together with tests for bioaccumulation of contaminants, chemical analyses, biological surveys of epifaunal and/or infaunal organisms, and perhaps the compilation of geological and hydrographic data. Statistical correlation can be improved and costs reduced if the samples are taken concurrently for these tests, analyses, and data acquisitions.

Samples of soil to be used in one or more of the three biological test methods herein (Section 4), might be taken quarterly, semiannually, or annually from a number of contaminated or potentially contaminated sites for *monitoring* and *compliance* purposes. Samples of soil might also be collected on one or more occasions during field surveys of sites for spatial (i.e., horizontal or vertical) or temporal definition of soil quality. One or more sites should be sampled for *reference* (presumably clean) *soil* during each field collection.[72]

The number of stations to be sampled at a study site and the number of *replicate samples* per station will be specific to each study. This will involve, in most cases, a compromise between logistical and practical constraints (e.g., time and cost) and statistical considerations. Webster and Oliver (1990), Crépin and Johnson (1993), and OMAFRA (1999) provide guidance on the sampling design; van Ee *et al.* (1990) and USEPA (1991) address issues related to quality assurance and quality control.

For certain monitoring and regulatory purposes, multiple replicates (i.e., separate samples from different grabs or cores taken at the same site) should be taken at each *sampling station*, including one or more reference stations.[73] Each of these field replicates should be tested for its toxicity to earthworms using five or more test chambers per replicate sample if conducting an acute lethality test (Section 4.1) or an acute avoidance test (Section 4.2), and ten or more test chambers per replicate sample if performing an eight-week test for effects of prolonged exposure (Section 4.3). The use of power analysis (see Section 5.5.2) with endpoint data obtained in previous tests of the same type, performed with previous samples from the

[72] Ideally, a *reference soil* is collected near the site(s) of concern. It possesses geochemical characteristics (e.g., texture, total organic carbon content, organic matter content, pH) similar to those of the field-collected test soil(s) but without anthropogenic contaminants. It is not unusual for nearby reference sites to have some degree of contamination from anthropogenic chemicals, and in some instances, reference soil might be toxic or otherwise unacceptable for use in a soil toxicity test, because of naturally occurring physical, chemical, or biological properties.

[73] ASTM (1999b) recommends the collection and testing of three field replicates from each sampling station, as a minimum.

same or similar sites, will assist in determining if additional laboratory replicates need to be tested with each field replicate. Also, some of the statistical tests have requirements for a minimum number of replicates. For certain other purposes (e.g., preliminary or extensive surveys of the spatial distribution of toxicity), the survey design might include only one sample from each station, in which case the sample would normally be homogenized and split between 5–10 replicate test chambers. The latter approach precludes any determination of mean toxicity at a given sampling location (station), and completely prevents any conclusion on whether a station is different from the control or reference, or from another location. It does, however, allow a statistical comparison of the toxicity of that particular sample with the reference or control, or with one or more samples from other locations. It is important to realize that any conclusion(s) about differences, which arise from testing single field samples lacking replication, cannot be extended to make any conclusion(s) about the sampling locations.

Sites for collecting *reference soil* should be sought where the geochemical properties of the soil are similar to soil characteristics encountered at the test sites. Matching of total organic carbon content (%) or organic matter content (%) might not be warranted in cases where *pollution* (e.g., from or within sewage or industrial sludge) is responsible for the high content of organic carbon and organic matter in test soils. Preliminary surveys to assess the toxicity and geochemical properties of soil within the region(s) of concern and at neighbouring sites are useful for selecting appropriate sites at which to collect reference soil.

Samples of municipal or industrial sludge (e.g., sewage sludge, dewatered mine tailings, or biosolids from an industrial clarifier or settling pond) might be collected for the assessment of their toxic effect(s) on earthworms, and for geochemical and contaminant analyses. Other particulate wastes being considered for land disposal might also be collected for toxicity and physicochemical evaluation.

Guidance for various soil sampling plans and procedures is available in the technical literature (e.g., Petersen and Calvin, 1986; Keith, 1992; Crépin and Johnson, 1993). Procedures used for sample collection (i.e., core, grab, or composite) will depend on the study objectives and the nature of the soil or other particulate material being collected. A shovel, auger, or soil corer (preferably stainless steel) is frequently used for collecting soil samples (ASTM, 1999b).

The surface of the location where each sample is to be collected should be cleared of debris such as twigs, leaves, stones, thatch, and litter. If the location is an area of grass or other plant material, the plants should be cut to ground level and removed before the sample is collected. Removal of the vegetation should be done such that removal of soil particles with the roots is minimal. Dense root masses (e.g., grasses) should be removed and then shaken vigorously to remove soil particles adhering to the roots. The soil sample to be collected for toxicity evaluation and chemistry should be taken from one or more depths that represent the layer(s) of concern (e.g., a surficial layer of soil, or one or more deeper layers of soil or subsoil if there is concern about historical deposition of contaminants).

The required volume of soil per sample should be calculated, before a sampling program is initiated. This calculation should take into account the quantity of soil required to prepare laboratory replicates for soil toxicity tests, as well as that required for particle size characterization, total organic carbon content (%), organic matter content (%), moisture content (%), and specific chemical analyses. A volume of at least 5–7 L of soil per sample is normally required, although this will depend on the study objectives/design (e.g., single-concentration or multi-concentration test) and the nature of the chemical analyses to be performed, and possibly also on the nature of the soil (e.g., need for removal of excess water and/or debris in the laboratory, which can reduce the sample volume). To obtain the required sample volume, it is frequently necessary to combine subsamples retrieved using the sampling device. The same collection procedure should be used at all field sites sampled.

5.2 Sample Labelling, Transport, Storage, and Analyses

Containers for transport and storage of samples of field-collected soil or similar particulate material must be made of nontoxic material. The choice of container for transporting and storing samples depends on both sample volume and the potential end uses of the sample. The containers must either be new, thoroughly cleaned, or lined with high-quality plastic. Thick (e.g.,

4 mil) plastic bags are routinely used for sample transport and storage (ASTM, 1999b). If plastic bags are used, it is recommended that each be placed into a second clean, opaque sample container (e.g., a cooler or a plastic pail with a lid) to prevent tearing and to support the weight of the sample and to maintain darkened conditions during sample transport.

Following sample addition, the air space in each container used for sample transport and storage should be minimized (e.g., by collapsing and taping a filled or partially filled plastic bag). Immediately after filling, each sample container must be sealed, and labelled or coded. Labelling and accompanying records made at this time must include at least a code or description that identifies sample type (e.g., grab, core, composite), source, precise location, land use information, replicate number, and date of collection; and should include the name and signature of sampler(s). Persons collecting samples of soil should also keep records that describe details of:

- the nature, appearance, and volume of each sample;
- the sampling procedure and apparatus;
- any procedure used to composite or subsample grabs or cores in the field;
- the number of replicate samples taken at each sampling station;
- the sampling schedule;
- the types and numbers of containers used for transporting samples;
- any field measurements (e.g., temperature, pH, soil moisture content) of the soil at the collection site;
- procedures and conditions for cooling and transporting the samples;
- observations of environmental conditions at the time of sampling (e.g., raining); and
- observations of soil fauna and vegetation at the collection site.

Soil samples should not freeze or become overheated during transport or storage. It is recommended that samples be kept in darkness (i.e., held in light-tight, opaque transfer containers such as coolers or plastic pails with lids) during transport, especially if they might contain PAHs or other chemicals or chemical products that could be photoactivated or otherwise altered due to exposure to sunlight. As necessary, gel packs, regular ice, or other means of refrigeration should be used to assure that the temperature of the sample(s) remains cool (e.g., 7 ± 3 °C) during transit.

The date the sample(s) is received at the laboratory must be recorded. Sample temperature upon receipt at the laboratory should also be measured and recorded. Samples to be stored for future use must be held in airtight containers. If volatile contaminants are in the soil or of particular concern, any air "headspace" in the storage container should be purged with nitrogen gas, before capping tightly. Samples must not freeze or partially freeze during transport or storage (unless they are frozen when collected), and must not be allowed to dehydrate. If, however, one or more samples are saturated with excess water upon arrival at the laboratory (e.g., sampling occurred during a significant rainfall event), the sample(s) may be transferred to plastic sheeting for a brief period (e.g., one or more hours) to enable the excess water to run off or evaporate. Thereafter, the sample(s) should be returned to the transport container(s) or transferred to one or more airtight containers for storage.

It is recommended that samples be stored in darkness at 4 ± 2 °C. These storage conditions must be applied in instances where PAHs or other light-sensitive contaminants are present or if the samples are known to contain unstable volatiles of concern. It is also recommended that samples of soil or similar particulate material be tested as soon as possible after collection. The soil toxicity test(s) should begin within two weeks of sampling, and preferably within one week. The test must begin within six weeks, unless it is known that the soil contaminants are aged and/or weathered and therefore considered stable.

Dry sieving (i.e., press sieving; not wet sieving) of samples through a coarse-mesh sieve is desirable to remove large particles (see Section 5.3). This procedure may be performed in the field. Undesirable coarse material (e.g., large gravel or stones, large debris, large indigenous macroinvertebrates, or large plant material) may also be removed in the field before sample transport. In the laboratory, each sample of field-collected soil should be thoroughly mixed (Section 5.3), and representative subsamples taken for physicochemical characterization. Each sample (including all samples of *negative control soil* and *reference soil*) must be characterized by analyzing subsamples for at least the following:

- particle size distribution (% sand, % silt, and % clay),
- total organic carbon content (%),
- organic matter content (%),

- moisture content (%),
- water-holding capacity (%, based on dry wt of soil),
- pH, and
- conductivity.

Additionally, the following analyses should be performed:

- texture,
- cation exchange capacity,
- major cations,
- organophosphorous insecticides, and
- organochlorine insecticides.

Other analyses could include:

- bulk density,
- fertility,
- C:N ratio,
- total inorganic carbon,
- total volatile solids,
- biochemical oxygen demand,
- chemical oxygen demand,
- oxidation-reduction potential,
- total nitrogen,
- total phosphorous,
- metals,
- petroleum hydrocarbons (including PAHs), and
- a suite of herbicides.

Unless indicated otherwise, identical chemical, physical, and toxicological analyses should be performed with subsamples representative of each replicate sample of field-collected soil (including *reference soil*) taken for a particular survey of soil quality, together with one or more subsamples of *negative control soil*.

5.3 Preparing Sample for Testing

Field-collected soil or similar particulate waste material must not be sieved with water, as this would remove contaminants present in the interstitial water or loosely sorbed to particulate material. Large gravel or stones, debris, indigenous macroinvertebrates, or plant material should normally be removed using forceps or a gloved hand. If a sample contains a large quantity of debris (e.g., plant material, wood chips, glass, plastic, large gravel) or large macroinvertebrates, these may be removed by pressing the soil through a coarse sieve (e.g., mesh size of ~6 mm; ASTM, 1999b).

Qualitative descriptions of each sample of field-collected test soil should be made and recorded at the testing laboratory, including information on sample colour, texture, and the presence and description of roots, leaves, and macroscopic soil organisms. Unless research or special study objectives dictate otherwise, each sample of field-collected test material should be homogenized in the laboratory before use (USEPA, 1989; ASTM, 1999b).[74] Mixing can affect the concentration and bioavailability of contaminants in the soil, and sample homogenization might not be desirable for all purposes.

As indicated in Section 3.5, one or more samples of field-collected test soil might either be tested at a single concentration only (typically, 100%), or evaluated for toxicity in a multi-concentration test whereby a series of concentrations are prepared by mixing measured quantities with either negative control soil or reference soil. When performing a multi-concentration test, the following series of concentrations of test soil (mixed in negative control soil or reference soil), which spans the range of 100–1% test soil using eight concentrations, might prove suitable: 100%, 80%, 50%, 30%, 15%, 7.5%, 3%, 1%, and 0%. Guidance on other concentration series that might prove as or more suitable is found in Section 6.2, along with that for preparing test mixtures which might apply equally when performing a multi-concentration test with one or more samples of field-collected soil. Refer to the introductory comments in Section 4, for additional guidance when selecting test concentrations. In each instance, the test must include a treatment comprised solely of negative control soil (see Section 3.2).

To achieve a homogeneous sample, transfer it to a clean, rigid mixing container (e.g., a large stainless steel or plastic bowl). The sample should be mixed manually (using a gloved hand or a nontoxic device such as a stainless steel spoon) or mechanically (e.g., using a domestic hand-held mixer with beaters at low speed, or a hand-held wire egg beater) until its texture and colour are homogeneous. While mixing, care should be taken to ensure that the impact of mixing on soil structure is minimal and that the structure is not

[74] One of the reasons for routinely homogenizing samples is to mix into the soil, any pore water which rises to the surface during sample shipment and storage. Homogenization is also necessary to redistribute the sample constituents that have compacted and layered according to particle size during transport and storage.

destroyed entirely. As soon as the texture and colour of the sample appears to be homogeneous, mixing should be discontinued.

For each sample included in a test, mixing conditions including duration and temperature must be as similar as possible. If there is concern about the effectiveness of sample mixing, subsamples of the soil should be taken after mixing, and analyzed separately to determine the homogeneity of particle sizes, chemical(s) of interest, etc. Any moisture that separates from a sample during its transport and/or storage must be remixed into it.

Soil structure is an important factor influencing the survival, reproduction, and growth of earthworms, and moisture content plays an important role in the determination of soil structure. A qualitative procedure, informally known as a "squeeze test", is useful when determining if the optimal moisture content of a sample of test soil has been achieved. Investigators might find it useful to apply this procedure when adjusting the moisture content of each sample of test soil to a particular percentage of the sample's water-holding capacity (see following paragraphs), in preparation for a toxicity test. The squeeze test can also be applied as a test proceeds (see Sections 4.1.5 and 4.3.5).[75] To perform this test, a small, representative subsample of the test soil (e.g., a "pinch" of soil) is randomly taken using a gloved hand, and gently compressed between the thumb and forefinger. If a small quantity of water can be squeezed from the soil with gentle pressure, then the soil's moisture content is acceptable. If, however, no water appears, the soil is likely too dry. Conversely, if a substantial amount of water can be squeezed from the subsample of soil, it is likely too wet.

The moisture content of a given sample of field-collected test soil should be standardized by determining its *water-holding capacity* (WHC) and then hydrating the soil to an optimal moisture content based on a percentage of this value. The optimal percentage of the WHC for each sample of field-collected soil must be determined before sample preparation and test initiation. In order to do so, the moisture content of each homogenized sample (i.e., each sample of test soil, including the negative control soil) must be determined first (Sections 4.1.1, 4.1.5, 4.2.1, and 4.3.1). Thereafter, the WHC of each sample must be determined using a recognized standard procedure (see following three paragraphs). A subsample of each soil sample is then hydrated to a homogeneous, crumbly consistency with clumps approximately 3–5 mm in diameter.[76] Based on the initial moisture content of the sample, the WHC of the sample, and the amount of water added to achieve the desired soil consistency, the sample's optimal moisture content can be calculated and expressed as a percentage of the WHC for each soil. Once this target (or optimal) percentage of the WHC has been determined, the moisture content of each sample of test soil (including the negative control soil) can be standardized to the selected (sample-specific) moisture content. *Test water* (i.e., de-ionized or distilled water) should be added to each sample with a moisture content that is less than the pre-determined optimal percentage of its WHC, until this moisture content is achieved [77] (Aquaterra

---

[75] A "squeeze test" can be applied during a 14-day lethality test with earthworms, when performing an optional 7-day check for the survival, appearance, and behaviour of test organisms (see Section 4.1.5). The squeeze test is also useful when making weekly observations of soil "wetness" during an eight- or nine-week test for effects of prolonged exposure to samples of test soil (see Section 4.3.5).

[76] An unpublished study, carried out by Environment Canada (J. Princz, personal communication, Biological Methods Division, Environmental Technology Centre, Ottawa, ON, 2004), determined the optimal moisture content for each of the diverse types of soil used while developing each of the biological test methods described herein (see Section 3.2 and Appendix H), based on a percentage of each sample's WHC. The optimal percentage of the WHC of these soils ranged from approximately 45–50% for the silt and sandy loam soils to 60% for the clay loam soil. These values were considered optimal since, at these levels of saturation, the soil mixed well, had an adequate moisture content according to the "squeeze test", and formed an acceptable structure (i.e., the resulting macro-aggregation of soil particles was conducive to healthy earthworms). Experience indicates that the actual moisture content of the test soils hydrated to optimal conditions can vary greatly (e.g., 20% for sandy loam soil to 50% for clay loam soil), depending on the bulk density and the WHC of the sample(s) of field-collected soil being tested (ESG and Aquaterra Environmental, 2002; Becker-van Slooten *et al*., 2003).

[77] An alternate approach sometimes used by certain investigators is to standardize (and adjust) the moisture content of each sample of field-collected soil to a fixed concentration, such as 35–45% of its dry weight (ASTM, 1999b; EC, 2000b). However, a disadvantage of this approach is that certain samples of field-collected soil can appear to be very wet and have standing water on the surface after hydration to only 35–45% of their dry weight; whereas other site soils can appear considerably dryer after the same level of hydration (ASTM, 1999b; EC, 2000b).

Environmental, 1998a). If a sample is too wet, it should be spread as a thin layer on a clean sheet of plastic (e.g., a new plastic garbage bag) or a clean, non-reactive (e.g., stainless steel or plastic) tray, and allowed to dry by evaporation at ambient (~20 °C) room temperature. Rehydration to the pre-determined optimal percentage of its WHC might be necessary. Upon adjustment of a sample's moisture content to the desired percentage of its WHC, the moisture content (%) of the hydrated soil must be determined and the percent WHC and percent moisture content recorded and reported.

The WHC of a particular soil is generally unique to each soil type, and is ultimately the result of the interaction of many variables associated with soil structure (e.g., micro/macro-aggregation, pore space, bulk density, texture, organic matter content). There are a number of methods that can be used to determine WHC; however, most of these methods require measurements to be made on an intact soil sample (e.g., soil core) where characteristics (structural aggregations, pore space, bulk density, texture, and OM) are preserved during collection. The USEPA (1989) has described an appropriate method for toxicity testing using unconsolidated materials (such as samples of field-collected soils that have been dried, sieved, and homogenized; or samples of soil formulated in the laboratory from constituents).[78] This method is outlined here.

For this method, ~130 g (wet wt) of sample is placed in an aluminum pan or petri dish (15 × 1 cm), and dried at 105 °C until a constant weight is achieved (this usually takes a minimum of 24 h). Thereafter, 100 g of the oven-dried soil is placed into a 250-mL glass beaker with 100 mL of distilled or de-ionized water. The resulting slurry is mixed thoroughly with a glass stir rod. A folded filter paper (185-mm diameter Fisherbrand P8 coarse porosity, qualitative creped filter paper; catalogue no. 09-790-12G) is placed into a glass funnel (with a top inside diameter of 100 cm and a stem length of 95 cm). The folded filter paper should be level with the top of the glass funnel. Using a pipette, up to 9 mL of distilled or de-ionized water is slowly added to the filter paper to wet the entire surface. The funnel and hydrated filter paper are then weighed. To obtain the initial weight for the mass of the funnel plus hydrated filter paper plus dried soil (see "I" in the following Equation 1), the weight of the dried soil (100 g) is added to the weight of the funnel and the wet filter paper.

The funnel is then placed into a 500-mL Erlenmeyer flask and the soil slurry is slowly poured onto the hydrated filter paper held in the funnel. Any soil remaining on the beaker and stir rod is rinsed into the funnel with the least amount of water necessary to ensure that all of the solid material has been washed onto the filter. The funnel is then tightly covered with aluminum foil and allowed to drain for three hours at room temperature. After three hours, the funnel containing the hydrated filter paper and wet soil is weighed. This weighing represents the final weight for the mass of the funnel plus hydrated filter paper plus (wet) soil (see "F" in the following Equation 1).

The water-holding capacity for the subsample of soil in the funnel, expressed as percentage of soil dry mass, is then calculated using the following equation:

$$WHC = \frac{F - I}{D} \times 100 \qquad \text{(Equation 1)}$$

where: WHC = water-holding capacity (%)
F = mass of funnel + hydrated filter paper + wet mass of soil
I = mass of funnel + hydrated filter paper + dry mass of soil
D = 100 g (i.e., dry mass of soil)

The WHC of each sample of test soil should be determined in triplicate, using three subsamples.

The percentage of water (i.e., P_w) that is added to a sample of field-collected soil to achieve the desired

---

Accordingly, the use of this alternate approach is not recommended here.

[78] Certain participants at a soil toxicity testing workshop sponsored by EC in Vancouver, BC (February 2003) considered the determination of WHC and a percentage of that capacity to be the most appropriate way of expressing soil moisture content (EC, 2004c). This led to a testing program to compare two different methods for estimating the WHC of soil (i.e., as per Annex C in ISO, 1999 or according to USEPA, 1989) as well as a somewhat different method for expressing soil moisture content, as a percentage of the soil's water-filled pore space (WFPS). The results of this investigation showed that each method had distinct advantages and disadvantages; however, the USEPA (1989) method for measuring WHC was recommended for use in EC's soil toxicity test methods when adjusting (if and as necessary) the moisture content of soil samples (Becker-van Slooten, *et al.*, 2004).

hydration (i.e., the optimal percentage of the WHC) can be calculated as follows:[79]

[79] The following example provides calculations that pertain to the hydration of samples of a contaminated field-collected soil and a negative control soil, when preparing a test concentration of 25% for use in an acute lethality test with earthworms involving three replicates per treatment.

Assumptions:

Soil #1: Negative Control (nc) Soil

W_{nc}	= 2.3934 g	P_{Wnc}	= 22.92%
D_{nc}	= 1.9108 g	M_{Dnc}	= 468.75 g dry wt
WHC_{nc}	= 80.30%	V_{Wnc}	= 107.44 mL
P_{WHCnc}	= 60.00%	V_{Wnc}	= 107.44 mL
MC_{nc}	= 25.26%	M_{Wnc}	= 587.14 g wet wt

Soil #2: Contaminated (c) Soil

W_c	= 7.0575 g	P_{Wc}	= 1.20%
D_c	= 5.6174 g	M_{Dc}	= 156.25 g dry wt
WHC_c	= 67.10%	V_{Wc}	= 1.88 mL
P_{WHCc}	= 40.00%	M_{Wc}	= 196.31 g wet wt
MC_c	= 25.64%		

MC	= [(W – D) / D] × 100	[Equation 1]
P_W	= [WHC × (P_{WHC}/100)] – MC	[Equation 2]
V_W	= (P_W × M) / 100	[Equation 3]
M_W	= (M_D × W) / D	

W	= wet mass of substrate (g)
D	= dry mass of substrate (g)
WHC	= water-holding capacity (% of dry mass)
P_{WHC}	= percentage of WHC desired (%)
MC	= initial moisture content of substrate (%)
P_W	= percentage of water to add to soil (%)
M_D	= total mass of soil required for experiment (expressed as dry wt)
V_W	= volume of water to add to soil (mL)
M_W	= total mass of soil required for experiment (expressed as wet wt based on initial MC)

Calculations for a 25% concentration of a contaminated soil in negative control soil:

For an acute lethality test using this example, it is assumed that a total mass of 625.00 g dry weight (wt) of soil is sufficient to satisfy the requirement for each treatment (i.e., 200.00 g dry wt per replicate × 3 replicates + 25.00 g dry wt extra soil for pH and conductivity etc.). To simplify the calculations, this example assumes that 200 g (dry wt) of either type of soil is sufficient to provide the 350-mL aliquot of soil to be added to each of three replicate test chambers per treatment, when performing an acute lethality test using *Eisenia* sp. (see Section 4.1.1).

$P_w = [WHC \times (P_{WHC}/100)] - MC$ (Equation 2)

where: P_w = percentage of water to add to the soil (%)
WHC = water-holding capacity (%)
MC_i = initial moisture content of the soil

The volume of water (i.e., V_w) that should be added to a sample of field-collected soil to achieve the desired hydration (i.e., the optimal percentage of the sample's water-holding capacity) can be calculated as follows:[79]

$V_W = (P_W \times M)/100$ (Equation 3)

where: V_w = volume of water to add to the soil (mL)
P_W = percentage of water to add to the soil (%)

For a 25% concentration of contaminated soil in negative control soil, 25% of the total mass of soil, on a dry-wt basis, must consist of the contaminated soil:

= 625.00 g dry wt × (25/100)
= 156.25 g dry wt of contaminated soil

The remainder of the test soil required to prepare this treatment (i.e., 75 %) will consist of the negative control soil:

= 625.00 g dry wt × (75/100)
[or 625.00 g dry wt – 156.25 g dry wt]
= 468.75 g dry wt of negative control soil

Therefore, the final total mass of soil required, based on wet weight, is 694.58 g [587.14 g wet wt at the soil's initial moisture content (i.e., M_{Wnc}) + 107.44 mL of water] for the negative control soil, and 198.19 g [196.31 g wet wt at the soil's initial moisture content (i.e., M_{Wc}) + 1.88 mL of water] for the contaminated soil.

The final moisture content for each soil would be 48.18% {[(694.58 – 468.75)/468.75] × 100} for the negative control soil, and 26.84% {[(198.19 – 156.25)/156.25] × 100} for the contaminated soil.

The final moisture content of the negative control soil (i.e., 48.18% moisture) represents 60% of that soil's water-holding capacity (48.18 ÷ 80.30 = 0.60). The final moisture content of the contaminated soil (i.e., 26.84% moisture) represents 40% of that soil's water-holding capacity (26.84 ÷ 67.10 = 0.40).

M = total mass of soil required for test (expressed as dry wt)[80]

Except for research-oriented toxicity tests intended to determine the influence of pH on sample toxicity, the pH of samples of field-collected soil must not be adjusted. Studies intending to investigate the effect of pH on sample toxicity should conduct two side-by-side tests, whereby one or more sets of treatments is adjusted to a fixed pH value using calcium carbonate or a suitable acid or base, and the pH of one or more duplicate sets of treatments is not adjusted.

Immediately following sample hydration (or dehydration) and mixing, subsamples of test material required for the toxicity test and for physicochemical analyses must be removed and placed into labelled test chambers (see Sections 4.1.1, 4.2.1, and 4.3.1), and into the labelled containers required for the storage of subsamples for subsequent physicochemical analyses. Any remaining portions of the homogenized sample that might be required for additional toxicity tests using earthworms or other test organisms (e.g., according to EC 2004d,e) should also be transferred to labelled containers at this time. All subsamples to be stored should be held in sealed containers with minimal air space, and must be stored in darkness at 4 ± 2 °C (Section 5.2) until used or analyzed. Just before being analyzed or used in the toxicity test, each subsample must be thoroughly remixed to ensure that it is homogeneous.

5.4 Test Observations and Measurements

A qualitative description of each field-collected test material should be made at the time the test is set up. This might include observations of sample colour, texture, and homogeneity, and the presence of plants or macroinvertebrates. Any changes in the appearance of the test material observed during the test or upon its termination, should be noted and reported.

Sections 4.1.5, 4.2.5, and 4.3.5 provide test-specific guidance and requirements for the observations and measurements to be made during or at the end of each test. These observations and measurements apply and must be made when performing any of the soil toxicity tests described herein using one or more samples of field-collected (site) soil.

Depending on the test objectives and experimental design, additional test chambers might be set up at the beginning of the test (Sections 4.1.1, 4.2.1, and 4.3.1) to monitor soil chemistry. These would be destructively sampled during and at the end of the test. Test organisms might or might not be added to these extra test chambers, depending on the study's objectives. Measurements of chemical concentrations in the soil within these chambers may be made by removing aliquots of the soil for the appropriate analyses (see Section 5.2).

5.5 Test Endpoints and Calculations

The common theme for interpreting the results of tests with one or more samples of field-collected test soil, is a comparison of the biological effects for the test (site) soil(s) with the effects found in a *reference soil*. The reference sample should be used for comparative purposes whenever possible or appropriate, because this provides a site-specific evaluation of toxicity (EC, 1997a, b, 2001). Sometimes the reference soil might be unsuitable for comparison because of toxicity or atypical physicochemical characteristics. In such cases, it would be necessary to compare the test soils with the *negative control soil*. Results for the *negative*

---

[80] For tests with samples of field-collected soil, the amount of soil added to each test chamber is based on the wet weight of soil that is equivalent to a volume of ~350 mL (see Sections 4.1.1, 4.2.1, and 4.3.1). However, "M" (i.e., the total mass of soil required for the test) is expressed as dry weight in the formula used to calculate the volume of water to be added to a sample of field-collected soil to achieve the desired hydration (see Equation 3). To calculate the amount of soil required per test chamber (by dry wt), a subsample of "wet" soil is placed into a test chamber (e.g., 500-mL glass jar) to determine the correct volume of soil required on a wet-weight basis. For example, assume that (for a given sample) this volume is equivalent to 270 g wet wt and, that the wet and dry weights of a subsample of this soil, previously determined for the purpose of calculating the sample's water-holding capacity, are 4.1507 g and 2.7813 g, respectively. The dry weight equivalent to a 350-mL volume of this sample of soil (which has a wet weight of 270 g) can be calculated as follows:

$$(270\ g \times 2.7813\ g) \div 4.1507\ g = 181\ g$$

This mass of soil can be rounded up to 200 g dry weight, thereby providing a little extra soil, if necessary. Therefore, for the example provided here, the mass of this sample of soil required for each replicate (expressed as dry wt) is 200 g. The total mass ("M") can then be calculated simply by multiplying the dry mass required for each replicate (in this instance, 200 g dry wt) by the number of replicates to be used in the test (i.e., for this example, three replicates).

control soil will assist in distinguishing contaminant effects from non-contaminant effects caused by soil physicochemical properties such as particle size, total organic carbon content (%), and organic matter content (%). Regardless of whether the *reference soil* or *negative control soil* is used for the statistical comparisons, the results from *negative control soil* must be used to judge the validity and acceptability of the test (see Sections 4.1.3, 4.2.3, and 4.3.3).

Analysis of the results will differ according to the purposes and particular designs of the test. This section covers the analytical procedures, starting with the simplest design and proceeding to the more complex designs. Standard statistical procedures are generally all that is needed for analyzing the results. Investigators should consult EC (2004b) for guidance on the appropriate statistical endpoints and their calculation. As always, the advice of a statistician familiar with *toxicology* should be sought for the design and analysis of tests.

Analysis of variance (ANOVA) involving multiple comparisons of endpoint data derived for single-concentration tests involving field replicates of field-collected soil from more than one sampling location is commonly used for statistical interpretation of the significance of findings from soil toxicity tests. This hypothesis-testing approach is subject to appreciable weaknesses. Notably, any increased variability within the test will weaken its power to distinguish toxic effects (i.e., less toxicity is concluded). Similarly, use of only a few replicates instead of many replicates will weaken the discrimination of a test and will lead to a conclusion of less apparent toxicity, other things being equal (see Section 5.5.2). There is no alternative to hypothesis testing, when comparing toxicity data for multiple samples of field-collected soil (i.e., field replicates of soil from more than one sampling location) that use only one concentration (usually full strength, i.e., 100% sample). There are alternatives for comparing point estimates of toxicity if various concentrations of each sample of field-collected soil are tested and multiple endpoint values for ICp or EC50 are determined (see Section 6.4). Section 9 in EC (2004b) should be consulted for guidance when comparing multiple ICps or EC50s (including multiple LC50s).

The parametric analyses involving ANOVA for comparative data from single-concentration tests with multiple samples of field-collected soil (i.e., field replicates of soil from more than one sampling location) assume that the data are normally distributed, that the treatments are independent, and that the variance is homogeneous among the different treatments. As the first step in analysis, these assumptions should be tested using the *Shapiro-Wilk's Test* for normality and *Bartlett's Test for Homogeneity of Variance* (Eisenhart *et al.*, 1947; Sokal and Rohlf, 1969). If the data satisfy these assumptions, analysis may proceed. If not, data could be transformed (e.g., as square roots, logarithms, or as arcsine square root for *quantal* data which are to be used in statistical analysis; Mearns *et al.*, 1986). The tests for normality and homogeneity might then show conformance with normality and homogeneity and, in fact, that is a likely outcome of a transformation. Assumptions should be re-tested following any transformation of data. Parametric tests are reasonably robust in the face of moderate deviations from normality and equality of variance; therefore, parametric analysis (e.g., ANOVA and multiple comparison) should proceed, even if moderate nonconformity continues after transformation. Excluding a data set for minor irregularities might lose a satisfactory and sensitive analysis and forgo the detection of real effects of toxicity.[81] Analysis by nonparametric statistical procedures should also proceed in parallel, with the more sensitive (lower endpoint) of the two analyses providing the final estimates of toxicity. Section 3 in EC (2004b) should be consulted for guidance when comparing the findings for single-concentration tests involving field replicates of samples from multiple locations, using parametric or non-parametric tests.

Guidance in Section 6 (including that in Section 6.2 for performing range-finding tests, and that in Section 6.4 for calculating test endpoints) should be followed if a multi-concentration test is performed using one or more samples of field-collected soil diluted with

[81] Tests for normality and homogeneity become less meaningful with the small number of samples from individual sampling stations typically collected in studies of environmental toxicology. Plotting and examining the general nature of the distribution of toxicity and its apparent deviations can be more revealing and is recommended (EC, 2004b). Equality in sample sizes and the magnitude of variation are probably more important factors for the outcome of parametric analysis, but they have received scant attention in toxicology. The robustness of ANOVA is shown by its ability to produce realistic probabilities if the distribution of data is reasonably symmetrical, and if treatment variances are within threefold of each other (Newman, 1995).

negative control soil or *clean* reference soil. Section 9 in EC (2004b) should be consulted when comparing such point estimates of toxicity for multiple samples of field-collected soil.

5.5.1 Variations in Design and Analysis

A very preliminary survey might have only one sample of *test soil* (i.e., contaminated or potentially contaminated *site soil*) and one sample of *reference soil*, without replication. Simple inspection of the results might provide guidance for designing more extensive studies.

If there is a single test sample and a single reference sample, with equal replication for each, a standard *Student's t-test* would be suitable for analysis (Paine and McPherson, 1991; EC, 1997a, b, 2001). The t-test is fairly robust, and handles unequal numbers of replicates in the test and reference samples, as well as moderately unequal variances in the two groups (Newman, 1995; USEPA, 1995).

A preliminary evaluation might conceivably be conducted with samples from many stations, but without either field replicates or laboratory (within-sample) replicates. The objective might be to identify a reduced number of sampling stations deserving of more detailed and further study. Opportunities for statistical analysis would be limited. The non-replicated test data could be compared with the reference data using outlier detection methods (USEPA, 1994a; Newman, 1995; EC, 1997a, b, 2001, 2004b). A sample would be considered toxic if its result was rejected as an extreme value when considered as part of the data for the *reference soil* and/or the *negative control soil*.

A more usual survey of soils would involve the collection of replicate samples from several places by the same procedures, and their comparison with replicate samples of a single *reference soil* and/or *negative control soil*. There are several pathways for analysis, depending on the type and quality of data, but often there would be an analysis of variance (ANOVA) followed by one of the multiple-comparison tests. In the ANOVA, the *reference soil* would also be treated as that from a "location".

In these multi-location surveys, the type of replication would influence the interpretation of results. If field replicates were collected at each of the sampling locations, and no laboratory replicates were used, a one-way ANOVA would evaluate the overall difference in test results with respect to sampling location, over and above the combined variability of sampling the location and running the test. It would be unusual but much more powerful, to have field replicates for all sampling locations and also laboratory replicates of each field replicate. If that were done, the laboratory replicates would become the replicates in a nested one-way ANOVA, and would be the base of variability for comparing differences in the samples. The ANOVA could be used to see (a) if there was an overall difference in test results for samples with respect to their sampling location, and (b) whether there was an overall difference in replicates taken at the various locations. After an ANOVA, the analysis would proceed to one or more types of multiple-comparison test, as described in the following text.

If only laboratory replicates and no field replicates were tested, there could be no conclusions about differences due to sampling location (see also Section 5.1). The laboratory replicates would only show any differences in the samples that were greater than the baseline variability in the within-laboratory procedures for setting up and running the test. Sample variability due to location would not really be assessed in the statistical analysis, except that it would contribute to any difference in test results associated with sampling location.

If it were desired to compare the test results for the replicate samples from each sampling location with those for the reference soil, to see if the toxicity of the two sources of soil (locations) differed, *Dunnett's test* should be used. It assumes normality and equal variance, and is based on an experiment-wise value of α (the probability of declaring a significant difference when none actually exists). If replication was unequal, investigators could use the *Dunn-Sidak modification of the t-test*, or alternatively the *Bonferroni adjustment of the t-test* (p. 189 in Newman, 1995; Appendix D in USEPA, 1995; Section 7.5.1 in EC, 2004b).

In a multi-location survey, an investigator might wish to know which of the samples from various sampling locations showed results that differed statistically from the others, as well as knowing which ones were different from the reference and/or negative control sample(s). Such a situation might involve sampling from a number of locations at progressively greater distances from a point source of contamination, in which instance the investigator might want to know

which sampling locations provided samples that had significantly higher toxicity than others, and thus which locations were particularly deserving of cleanup. *Tukey's test* is designed for such an analysis; this test is commonly found in statistical packages and can deal with unequal sample sizes.[82]

If it were desired to compare the toxicity of the samples from each sampling location with that for the reference sample(s), but the data do not conform to requirements of normality and equal variance, the ANOVA and subsequent tests would be replaced by nonparametric tests. *Steel's Many-One Rank test* would be used if replication were equal, while unequal replication would require use of the *Wilcoxon Rank Sum test* with Bonferroni's adjustment.

5.5.2 Power Analysis

An important factor to consider in the analysis of the results for toxicity tests with soil is the potential for declaring false positives (i.e., calling a *clean* site contaminated; Type I error) or false negatives (i.e., calling a contaminated site *clean*; Type II error). Scientists are usually cautious in choosing the level of significance (α) for tolerating false positive results (Type I error), and usually set it at P = 0.05 or 0.01. Recently, toxicologists have been urged to report both α and statistical power (1 - β), i.e., the probability of correctly rejecting the null hypothesis (H_0) and not making a Type II error. There are several factors that influence statistical power, including:

- variability of replicate samples representing the same treatment;
- α (i.e., the probability of making a Type I error);
- *effect size (ES)*, (i.e., the magnitude of the true effect for which you are testing); and
- *n* (i.e., the number of samples or replicates used in a test).

Environment Canada's guidance document on statistical methods for environmental toxicity tests (EC, 2004b) provides further information and guidance on errors of Types I and II.

Power analysis can be used *a priori* to determine the magnitude of the Type II error and the probability of false positive results. It can also be used to ascertain the appropriate number of field and laboratory replicates for subsequent surveys involving this test, or to assist in the selection of future sampling sites. It is always prudent to include as many replicates in the test design as is economically and logistically warranted (see Section 5.1); power analysis will assist in this determination. A good explanation of the power of a test, and how to assess it, can be found in USEPA (1994a). Guidance on power analysis is provided in EC (2004b).

Many investigators have difficulty with power analysis, and do not apply it due to its perceived complexity and the differing formulae specific to various statistical tests. In view of this complexity, the *Minimum Significant Difference* may be applied as an alternative approach (i.e., as an "index of power"; see EC, 2004b for guidance).

---

[82] An alternative approach is available (EC, 1997a, b, 2001). For equal replicates, *Fisher's Least Significant Difference (LSD)* is recommended. It is based on a smaller "pairwise error rate" for α in comparing data for samples from any given location with those for samples from another location, but holds the overall value of α to the pre-selected value (usually 0.05). *LSD* is seldom included in software packages for toxicity, but it is described in some textbooks (e.g., Steel and Torrie, 1980). Instead, *Tukey's test* is recommended here, partly because *LSD* might declare significant differences too readily. *LSD* is also intended for only a few of all the possible comparisons in a set of data, and those comparisons would have to be specified in advance.

Specific Procedures for Testing Chemical-Spiked Soil

This section gives guidance and instructions for preparing and testing *negative control soil* spiked experimentally with chemical(s) or chemical product(s). These recommendations and instructions apply to each of the three biological test methods described in Section 4. Guidance in EC (1995) for *spiking* negative control sediment with chemical(s) and conducting toxicity tests with chemical/sediment mixtures is also relevant here, for chemical-spiked soil. Further evaluation and standardization of procedures for preparing chemical-spiked soil provided herein (Section 6.2) might be required before soil toxicity tests with earthworms or other appropriate soil organisms are applied to evaluate specific chemical/soil mixtures for regulatory purposes.

The cause(s) of soil toxicity and the interactive toxic effects of chemical(s) or chemical product(s) in association with otherwise *clean* soil can be examined experimentally by spiking *negative control soil* (Section 3.2) with these substances. The spiking might be done with one or more chemicals or chemical products. Other options for toxicity tests with earthworms, performed using the procedures described herein, include the spiking of chemical(s) or chemical product(s) in *reference soil* (Section 3.4) or *test soil* (Section 3.5). Toxicity tests using soil spiked with a range of concentrations of test chemical(s) or chemical product(s) can be used to generate data that estimate LC50s (see Section 6.4.1), and can determine other statistical endpoints based on threshold concentrations causing specific sublethal effects (see Sections 6.4.1 and 6.4.2).

In Section 6.2, procedures are described for preparing test mixtures of chemical-spiked soil. Section 6.3 describes procedures for making observations and measurements during and at the end of the toxicity test, and Section 6.4 provides procedures for estimating test endpoints for multi-concentration tests. These procedures also apply to the mixing of multiple concentrations of field-collected *test soil* (including particulate waste material such as sludge or other dredged material intended for land disposal) in *negative control soil* or *reference soil*, and to performing multi-concentration tests and determining statistical endpoints for these mixtures (see Section 5, and especially 5.5). Multi-concentration tests with *positive control soil* (Section 3.3) or one or more reference toxicants spiked in *negative control soil* (Sections 4.1.8, 4.2.8, and 4.3.8) are also performed using the procedures and statistical guidance described in this section. Additionally, the influence of the physicochemical characteristics of natural or artificial negative control soil on chemical toxicity can be determined with spiked-soil toxicity tests according to the procedures and statistical guidance described in this section.

6.1 Sample Properties, Labelling, and Storage

Information should be obtained on the properties of the chemical(s) or chemical product(s) to be spiked experimentally in the *negative control soil*.[83] Information should also be obtained for individual chemicals or chemical products (e.g., pesticides or other commercial formulations) on concentration of major or "active" ingredients and impurities, water solubility, vapour pressure, chemical stability, dissociation constants, adsorption coefficients, toxicity to humans and terrestrial organisms, and biodegradability. Where aqueous solubility is in doubt or problematic, acceptable procedures previously used for preparing aqueous solutions of the chemical(s) should be obtained and reported. If an acceptable procedure for solubilizing the test chemical(s) in water is not available, preliminary testing for its solubility in test water or a non-aqueous solvent should be conducted and confirmed analytically. Other available

[83] Some studies might require the spiking (mixing) of one or more concentrations of chemical(s), chemical product(s), or *test soil* (e.g., contaminated or potentially contaminated field-collected soil or waste sludge) in either *negative control soil* or *reference soil*. Other applications could include the spiking of chemical(s) or chemical product(s) in one or more samples of *test soil*. For such studies involving samples of contaminated soil or similar particulate material (e.g., domestic or industrial sludge), instructions on sample characterization given in Section 5.2 should be followed. Sample(s) of field-collected negative control soil, reference soil, contaminated soil, or particulate waste to be evaluated in spiked-soil toxicity tests should be collected, labelled, transported, stored, and analyzed according to instructions provided in Sections 5.1 and 5.2.

information such as the structural formulae, nature and percentage of significant impurities, presence and amounts of additives, and n-octanol:water partition coefficient, should be obtained and recorded. Any pertinent Material Safety Data Sheets (MSDSs) should be obtained and reviewed.

Chemical(s) to be tested should be at least reagent grade, unless a test on a formulated commercial product or technical grade chemical(s) is required. Chemical containers must be sealed and coded or labelled upon receipt. Required information (chemical name, supplier, date received, person responsible for testing, etc.) should be indicated on the label and/or recorded on a separate datasheet dedicated to the sample, as appropriate. Storage conditions (e.g., temperature, protection from light) are frequently dictated by the nature of the chemical.

6.2 Preparing Test Mixtures

On the day preceding the start of the toxicity test (i.e., Day -1), the mixture(s) of chemical(s) or chemical product(s) spiked in *negative control soil* should be prepared, transferred to test chambers, and held overnight before adding the test organisms the next day (i.e., Day 0) (see Sections 4.1.1, 4.2.1, and 4.3.1). Each *batch* of test soil representing a particular treatment (concentration) should be prepared in a quantity sufficient to enable all test replicates of that treatment (concentration) to be set up along with any additional replicates or quantities required for physicochemical analyses (Section 6.3) or the performance of other soil toxicity tests using earthworms or other soil organisms (e.g., those performed according to EC 2004d or EC 2004e).

The use of artificial soil (Section 3.2.2) to prepare each test mixture is recommended since it offers a consistent, standardized approach for comparing results for other chemicals or chemical products tested similarly in the same laboratory or by others (e.g., according to OECD, 1984; USEPA, 1989; ISO, 1993, 1998; ASTM, 1999b; or OECD, 2000). If used, the formulation for artificial soil provided in Section 3.2.2 should be followed. The quantity of artificial soil required for the test(s) should be prepared, hydrated to ~20% moisture content, adjusted if and as necessary to a pH within the range of 6.0–7.5[84], aged for a minimum three-day period, and stored at 4 ± 2 °C until required (see Section 3.2.2). The final moisture content (including that due to the addition of a measured aliquot of a test chemical or chemical product dissolved in *test water*, with or without an organic solvent) of any chemical-spiked soil prepared using artificial soil should be ~70% of the water-holding capacity of the final mixture (Section 3.2.2), for each treatment (concentration).[85] The final

---

[84] If, however, the test chemical(s) or chemical product(s) are anticipated to modify soil pH and the intent of the study is to nullify this influence, the (aqueous) pH of each *batch* (concentration) should be adjusted to a standard value (e.g., pH 6.5). Studies for determining the extent to which an acidic or basic test substance modifies the toxicity of soil spiked with a range of concentrations of this substance, due to the influence of pH *per se*, should involve two side-by-side tests. One test adjusts the pH of each test concentration to a standard value (e.g., pH 6.5) using the required (differing, depending on concentration) quantity of calcium carbonate, and the other test uses an identical quantity of calcium carbonate for each treatment sufficient to attain the "standard" pH (e.g., pH 6.5) in the *negative control* treatment.

[85] The following example provides calculations that show the volume of both water (de-ionized or distilled) and a stock solution of a reference toxicant (boric acid) to be added to a sample of artificial soil with an existing moisture content, to create a treatment with a moisture content that is 70% of the WHC for the artificial soil. The calculations take into account the volume of a stock solution of boric acid added when preparing the treatment, as part of the overall adjustment for soil moisture content. To simplify the calculations, this example assumes that 200 g (dry wt) of artificial soil (AS) is sufficient to provide the 350-mL aliquot of soil to be added to each test chamber when performing an acute lethality test with *Eisenia* sp. involving three replicate test chambers per treatment (see Section 4.1.1).

The equations shown in Section 5.3 for calculating WHC and adjusting soil moisture content to a certain percentage of this value apply equally here. For this example, assume that the following assumptions apply (see Section 5.3 for equations and associated definitions of these terms).

Assumptions:
Wet mass of artificial soil (AS) = 3.2486 g
Dry mass of AS = 2.6924 g
Moisture content (MC) of AS = [(3.2486 – 2.6924)/ 2.6924] × 100
= 20.66% (initial moisture content)
Water-holding capacity (WHC) of AS = 72.10%
Desired percentage of WHC (P_{WHC}) = 70.00%
Dry mass of AS required for test (M_D) = [200.00g per rep × 3 reps] + 25.00 g extra
= 625.00 g dry wt

mixture (treatment) included in a test should be as similar as possible.

Investigators may choose to use natural control soil (Section 3.2.1) rather than artificial control soil (Section 3.2.2) as the *negative control soil* to be spiked with chemical(s) or chemical product(s) and for the corresponding replicates of control soil to be included in the test. Procedures described herein for artificial soil apply equally if natural soil is used. An exception is that the final moisture content of each *batch* of chemical-spiked soil (including control batches) prepared using field-collected soil should be adjusted to the optimal percentage of its WHC using guidance in Section 5.3.

The procedure to be used for experimentally spiking soil is contingent on the study objectives and the nature of the test substance to be mixed with *negative control soil* or other soil. In many instances, a chemical/soil mixture is prepared by making up a stock solution of the test chemical(s) or chemical product(s) and then mixing one or more measured volumes into artificial or natural *negative control soil* (Section 3.2). The preferred solvent for preparing *stock solutions* is *test water* (i.e., de-ionized or distilled water); use of a solvent other than 100% test water should be avoided unless it is absolutely necessary. For test chemical(s) or chemical product(s) that do not dissolve readily in test water, a suitable water-miscible organic solvent of low toxicity (e.g., acetone, methanol, or ethanol) may be used in small quantities to help disperse the test substance(s) in water. Surfactants should not be used.

If an organic solvent must be used, the test must be conducted using a series of replicate test chambers containing only *negative control soil* (i.e., 100% artificial or natural *clean* soil containing no solvent and no test substance), as well as a series of replicate test chambers containing only solvent control soil (OECD, 1984, 2000; ISO, 1993, 1998; ASTM, 1999b; EC, 2000b). For this purpose, a *batch* of *solvent control soil* must be prepared containing the concentration of solubilizing agent that is present in the highest concentration of the test chemical(s) or chemical product(s) in soil. Solvent from the same *batch* used to make the stock solution of test substance(s) must be used. Solvents should be used sparingly, since they might contribute to the toxicity of the prepared test soil. The maximum concentration of solvent in the soil should be at a concentration that does not affect the performance of earthworms during the test. If this information is unknown, a preliminary *solvent only* test, using various concentrations of solvent in *negative control soil*, should be conducted to determine the threshold-effect concentration of the particular solvent being considered for use in the definitive test.

For tests involving the preparation of concentrations of chemical spiked in artificial soil, in which the chemical

---

/ 2.6924
= 754.11 g wet wt

Calculations to prepare a treatment comprised of 2000 mg boric acid per kg artificial soil (dry wt):

The stock solution consists of 25 g of H_3BO_3 in 1 L of de-ionized water

The amount of boric acid required, on a dry-mass basis is:
H_3BO_3 = (2 g H_3BO_3 / 1000 g soil dry wt) × 625.00 g dry wt
= 1.25 g H_3BO_3

The amount of stock solution required, on a volume basis, is:

H_3BO_3 = 1.25 g H_3BO_3 / (25 g H_3BO_3 / 1000 mL of water)
= 50.00 mL stock solution

The percentage of water (P_W) required for addition to this treatment to achieve the desired percentage of WHC (70%) is:

P_W = [WHC x (P_{WHC} / 100)] – MC
= [72.10 x (70.00 / 100)] – 20.66
= 29.81%

The volume of water (V_W) required for addition to this treatment to achieve the desired percentage of WHC (70%) is):

V_W = (P_W x M_D) / 100
= (29.81 x 625.00 g dry wt) / 100
= 186.31 mL of water required

However, as part of this required volume, 50.00 mL of the stock solution is to be added for dosing; therefore, an additional volume of water of only 136.31 mL will be required (186.31 mL of water – 50.00 mL of stock solution).

Accordingly, the final total mass of soil required, based on wet weight, would be 940.42 g [754.11 g wet wt at the soil's initial moisture content (i.e., M_W) + 136.31 mL of water + 50.00 mL of stock solution], and the final moisture content of the soil, based on dry weight, would be 50.47 % {[(940.42 – 625.00) / 625.00] x 100}.

The final moisture content of this test treatment (i.e., 50.47% moisture) represents 70% of the test soil's water-holding capacity (50.47 ÷ 72.10 = 0.70).

is insoluble in water but soluble in an organic solvent, the quantity of test substance needed to prepare a required volume of a particular test concentration should be dissolved in a small volume of a suitable organic solvent (e.g., acetone). This chemical-in-solvent mixture should then be sprayed onto or mixed into a small portion of the final quantity of fine quartz sand that is required when preparing each test concentration comprised of a measured amount of a particular chemical-in-solvent mixture spiked in artificial soil (see Section 3.2.2). The solvent is then removed by evaporation by placing the container under a fume hood for at least one hour, and until no residual odour of the solvent can be detected. Thereafter, the chemical-in-sand mixture (with solvent evaporated) is mixed thoroughly with the remaining quantity of pre-moistened sand and other ingredients required to make up artificial soil (Section 3.2.2). An amount of de-ionized water necessary to achieve a final moisture content of approximately 70% of the maximum water-holding capacity for this artificial soil is then added and mixed with the soil/sand/peat mixture. The chemical-spiked soil can then be added to the test chambers (OECD, 2000).

The sample of *solvent control soil* to be included in the test must be prepared using the same procedure but without the addition of the test chemical. Additionally, the solvent control soil must contain a concentration of solvent that is as high as that in any of the concentrations of chemical-spiked soil included in a test.

If the test chemical to be spiked in artificial soil is insoluble in both water and any suitable (non-toxic) organic solvent, a mixture should be prepared comprised of 10 g finely ground industrial quartz sand and the quantity of the test chemical necessary to achieve the desired test concentration in the soil. This mixture should then be mixed thoroughly with the remaining constituents of the pre-moistened artificial soil. An amount of de-ionized water necessary to achieve a final moisture content of ~70% of the maximum water-holding capacity is then added and mixed in. The resulting mixture of chemical-spiked soil can then be added to the test chambers (OECD, 2000).

Concentrations of chemical(s) or chemical product(s) in soil are usually calculated, measured, and expressed as mg test substance/kg soil (or µg substance/g soil) on a dry-weight basis (OECD, 1984, 2000; ISO, 1993. Endpoint results are similarly expressed on a dry-weight basis (Section 6.4).

Mixing conditions, including solution:soil ratio, mixing and holding time, and mixing and holding temperature, must be standardized for each treatment included in a test. Time for mixing a *spiked soil* should be adequate to ensure homogeneous distribution of the chemical, and may be from minutes up to 24 h. During mixing, temperature should be kept low to minimize microbial activity and changes in the mixture's physicochemical characteristics. Analyses of subsamples of the mixture are advisable to determine the degree of mixing and homogeneity achieved.

For some studies, it might be necessary to prepare only one concentration of a particular mixture of negative control (or other) soil and chemical(s) or chemical product(s), or a mixture of only one concentration of contaminated soil or particulate waste in negative control or other soil. For instance, a single-concentration test might be conducted to determine whether a specific concentration of chemical or chemical product in *clean* soil is toxic to the test organisms. Such an application could be used for research or regulatory purposes.

A multi-concentration test, using a range of concentrations of chemical added to *negative control soil* (or other soil) under standardized conditions, should be used to determine the desired endpoint(s) (i.e., LC50, EC50, ICp; see Section 6.4) for the chemical/soil mixtures. A multi-concentration test using *negative control soil* spiked with a specific particulate waste might also be appropriate. At least five test concentrations plus the control(s) must be prepared for each multi-concentration test performed to estimate an acute LC50 (see Sections 4.1.2 and 4.1.8) or an EC50 (or other ECp) for acute avoidance (see Section 4.2.2); and more (i.e., six to ten plus controls) are recommended. For a 56-day (or longer) test, at least seven test concentrations plus the appropriate control treatment(s) must be prepared for each multi-concentration test, and more (i.e., $\geq$10 plus controls) are recommended (see Sections 4.3.1 and 4.3.2). When selecting the test concentrations, an appropriate geometric dilution series may be used in which each successive concentration of chemical(s) or chemical product(s) in soil is at least 50% of the previous one (e.g., 10, 5, 2.5, 1.25, 0.63 mg/kg). Test concentrations may also be selected from other appropriate logarithmic dilution series (see Appendix I); or may be derived based on the findings of preliminary “range-finding” toxicity tests. The reader is referred to the introductory comments in Section 4 for additional guidance when selecting test concentrations.

To select a suitable range of concentrations, a preliminary or range-finding test covering a broader range of test concentrations might prove worthwhile. The number of replicates per treatment (see Sections 4.1.1, 4.2.1, and 4.3.1) could be reduced or eliminated altogether for range-finding tests and, depending on the expected or demonstrated (based on earlier studies with the same or similar test substance) variance among test chambers within a treatment, might also be reduced for nonregulatory screening *bioassays* or research studies.

Based on the objectives of the test, it might be desirable to determine the effect of substrate characteristics (e.g., particle size or organic matter content) on the toxicity of chemical/soil mixtures. For instance, the influence of soil particle size on chemical toxicity could be measured by conducting concurrent multi-concentration tests with a series of mixtures comprised of the test chemical(s) or chemical product(s) mixed in differing fractions (i.e., segregated particle sizes) or types of natural or artificial *negative control soil* (Section 3.2). Similarly, the degree to which the total organic carbon content (%) or organic matter content (%) of soil can modify chemical toxicity could be examined by performing concurrent multi-concentration tests using different chemical/soil mixtures prepared with a series of organically enriched negative control soils. Each fraction or formulation of natural or artificial *negative control soil* used to prepare these mixtures should be included as a separate control in the test.

Depending on the study objectives and design, certain soil toxicity tests using earthworms might be performed with samples of negative control soil or reference soil to which chemical(s) or chemical product(s) are applied to the soil surface, rather than mixing it with the soil. Surface applications can be applied in the field or the laboratory. Procedures for chemical application include the use of a calibrated track sprayer to achieve a uniform distribution of the chemical over a specific area. Concentration of chemical(s) or chemical product(s) in the soil can be determined based on the penetration depth, the surface area or swathe width, the nozzle size, the pressure, and the speed of coverage of the sprayer (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001). The OECD (2000) provides guidance for applying test substances to the soil surface, in preparation for earthworm reproduction tests.

6.3 Test Observations and Measurements

A qualitative description of each mixture of *chemical-spiked soil* should be made when the test is being established. This might include observations of the colour, texture, and visual homogeneity of each mixture of chemical-spiked soil. Any change in appearance of the test mixture during the test, or upon its termination, should be recorded.

Sections 4.1.5, 4.2.5, and 4.3.5 provide test-specific guidance and requirements for the observations and measurements to be made at the beginning, during, and at the end of the test. These observations and measurements apply and must be made when performing any of the soil toxicity tests described herein using one or more samples of chemical-spiked soil.

Depending on the test objectives and experimental design, additional test chambers might be set up on Day -1 of the test (see Sections 4.1.1, 4.2.1, and 4.3.1) to monitor soil chemistry. These would be destructively sampled during (i.e., on Day 0 and, in certain instances, other days as the test progresses) and at the end of the test. Test organisms might or might not be added to these extra test chambers, depending on study objectives. Measurements of chemical concentrations in the soil within these test chambers could be made by removing aliquots of soil for the appropriate analyses, at the beginning of the test, as it progresses, and/or at its end.

Measurements of the quality (including soil pH and moisture content) of each mixture of spiked soil being tested (including the *negative control soil*) must be made and recorded at the beginning and end of the test, as described in Sections 4.1.5, 4.2.5, and 4.3.5. If analytical capabilities permit, it is recommended that the stock solution(s) be analyzed together with one or more aliquots of each spiked-soil mixture, to determine the chemical concentrations, and to assess whether the soil has been spiked satisfactorily. These should be preserved, stored, and analyzed according to suitable, validated procedures.

Unless there is good reason to believe that the chemical measurements are not accurate, toxicity results for any test in which concentrations are measured for each spiked-soil mixture included in the test should be calculated and expressed in terms of these measured values. As a minimum, sample aliquots should be

taken from the high, medium, and low test concentrations at the beginning and end of the test[86]; in which instance, endpoint values calculated (Section 6.4) would be based on nominal ones. Any such measurements of concentrations of the test chemical(s) or chemical product(s) should be compared, reported, and discussed in terms of their degree of difference from nominal strengths. If nominal concentrations are used to express toxicity results, this must be explicitly stated in the test-specific report (see Section 7.1.6).

6.4 Test Endpoints and Calculations

Multi-concentration tests with mixtures of spiked soil are characterized by test-specific statistical endpoints (see Sections 4.1.7, 4.2.7, and 4.3.7). Guidance for calculating the LC50 (Sections 4.1.7 and 4.3.7) or EC50 (Section 4.2.7) is provided in the following Section 6.4.1, whereas that for calculating an ICp (based on data showing reproductive inhibition or growth inhibition; see Section 4.3.7) is given in Section 6.4.2. Section 5.5 provides guidance for calculating and comparing endpoints for single-concentration tests using samples of field-collected soil, which applies equally to single-concentration tests performed with mixtures of spiked soil. For further information on these or other appropriate parametric (or nonparametric) statistics to apply to the endpoint data, the investigator should consult the Environment Canada report on statistics for the determination of toxicity endpoints (EC, 2004b).

For any test that includes *solvent control soil* (see Section 6.2), the test results for earthworms held in that soil must be compared statistically with that for test organisms held in *negative control soil*. If any of the endpoints for these two control soils used to establish test validity (see Sections 4.1.3, 4.2.3, and 4.3.3) differ significantly according to *Student's t-test*, only the *solvent control soil* may be used as the basis for comparison and calculation of results. If the results for the two controls are the same, the data from both controls should be combined before using it to calculate results or to assess test validity.

6.4.1 LC50 or EC50

When a multi-concentration test with spiked soil mixtures is conducted (Section 6.2), the quantal mortality data for a specific period of exposure must be used to calculate (data permitting) the appropriate *median lethal concentration (LC50)*, together with its 95% confidence limits. For an acute lethality test using earthworms, a multi-concentration test should determine the seven-day LC50 (i.e., that based on mortality data mid-way through the test), and must determine the 14-day LC50 (at test end) (see Section 4.1.7). For a test of effects of prolonged exposure to multiple concentrations of spiked soil on earthworms, the 28-day LC50 for the adult (first generation) worms must be calculated and reported, data permitting[87] (see Section 4.3.7).

For a multi-concentration acute avoidance test using earthworms, the *median effective concentration (EC50)* must be calculated (together with its 95% confidence limits), data permitting. This calculation is based on the percent avoidance responses for each test concentration (Section 4.2.7). The guidance for calculating LC50 that follows applies equally to calculating an EC50, in which instance the word "mortality" should be replaced with the word "avoidance", and "mortalities" replaced with "avoidance responses". To estimate an LC50, mortality data at the specified period of exposure are combined for all replicates at each concentration. If mortality is not $\geq$50% in at least one concentration, the LC50 cannot be estimated. If there are no mortalities at a specific concentration, that information is used as an effect of 0% mortality. However, if successive concentrations yield a series of 0% mortalities, only the highest concentration of the series should be used in estimating the LC50 (i.e., the zero-effect that is "closest to the middle" of the distribution of data). Similarly, if there were a series of successive complete mortalities at the high concentrations in the test, only one value of 100% effect would be used, i.e., the one at the lowest concentration. Use of only one 0% and one 100% effect applies to any form of statistical analysis and to plotting on a graph.

[86] Certain chemicals might be known to be stable under the defined test conditions, and unlikely to change their concentration over the test duration. In this instance, an investigator might choose to restrict their analyses to samples taken only at the beginning of the test.

[87] Depending on the study objectives and the associated experimental design, an eight-week test for effects on survival, reproduction, and growth of earthworms (*Eisenia* sp.) might be focused on sublethal effects. In this instance, the test might not include a sufficient number of high (lethal) concentrations to enable calculation of the 28-day LC50.

Environment Canada (2004b) provides guidance on the choice of statistical test methods to be applied to quantal (e.g., LC50 or EC50) data, which should be consulted when choosing the statistical test to be applied to such data for toxicity tests using earthworms. Probit and/or logit regressions are the "preferred" methods (EC, 2004b), provided that the data include two or more concentrations showing partial effects. The probit analysis also gives the slope of the line, which should be reported. If probit or logit do not work because of only one partial effect, use the Spearman-Kärber method with no trim. If no partial effect is evident, use the binomial method. The binomial estimate might differ somewhat from the others, and this estimate should only be used as a last resort. Formal confidence limits are not estimated using the binomial method; instead, outer limits of a range are provided, within which the LC50 and the true confidence limits would lie.

Various computer programs may be used to calculate the LC50. Stephan (1977) developed a program to estimate LC50s using probit, moving average, and binomial methods, and adapted it for the IBM-compatible personal computer. Use of this program, which was modified in 1989 to include estimates using the Spearman Kärber method with no "trimming" (i.e., with no deletion of data from the calculations), is available on diskette[88] from Environment Canada (see Appendix C). Other satisfactory computer and manual methods may be used (e.g., SAS 1988 or version 3.5 of TOXSTAT 1996; see EC 2004b for additional information). Programs using the trimmed Spearman-Kärber method are available for personal computers; however, this method (with trimming) should be applied cautiously to LC50 (or EC50; see Section 4.2.7) estimates according to EC (2004b), because divergent results might be obtained by operators who are unfamiliar with the implications of trimming ends of the concentration-response data. However, there are situations where application of the trimmed Spearman-Kärber method is warranted (see EC, 2004b for guidance).

Any computer-derived LC50 should be checked by examining a plot, on logarithmic-probability scales, of percent mortalities at a defined period of exposure for the various test concentrations (EC, 2004b). Any major disparity between the estimated LC50 derived from this plot and the computer-derived LC50 must be resolved. A hand-plotted graph is recommended for this check (EC, 2004b). A computer-generated plot (e.g., SigmaPlot™; Version 8.0.2 or later)[89] could be used if it were based on logarithmic-probability scales. If there has been an error in entering the data, however, a computer-generated plot would contain the same error as the mathematical analysis, and so the investigator should carefully check for correct placement of points (EC, 2004b).

A manual plot of mortality/concentration data to derive an estimated LC50 is illustrated in Figure 3. This (hypothetical) figure is based on test concentrations of 1.8, 3.2, 5.6, 10, and 18 mg chemical/kg soil (dry-wt basis) causing mortalities of 0, 20, 40, 90, and 100% of test organisms exposed to the respective concentrations for a specified period of time. The concentration expected to be lethal to 50% of the earthworms can be read by following across from 50% (broken line) to the intersection with the best-fit line, then down to the horizontal axis for an estimated LC50 (5.6 mg/kg, dry wt).

In fitting a line such as that in Figure 3, more emphasis should be assigned to points that are near 50% mortality. Logarithmic-probability paper (*log-probit*, as in Figure 3) can be purchased in good technical bookstores, ordered through them, or photocopied (see blank graph in EC, 2004b).

For the regular set of data in Figure 3, computer programs gave very similar estimates to the graphic one. Some of the computed LC50s (and 95% confidence limits) were:

Stephan (1977) method:

• probit:	5.58 (4.24 and 7.37)
• moving average:	5.58 (4.24 and 7.33)
• binomial:	6.22 (between 1.8 and 10)

[88] Through the courtesy of Dr. Charles E. Stephan (USEPA, Duluth, MN).

[89] Available for purchase from SYSTAT Software, Inc., 501 Canal Boulevard, Suite C, Point Richmond, CA 94804-2028, USA, phone no. 1-800-797-7401; see Web site www.systat.com/products/SigmaPlot/.

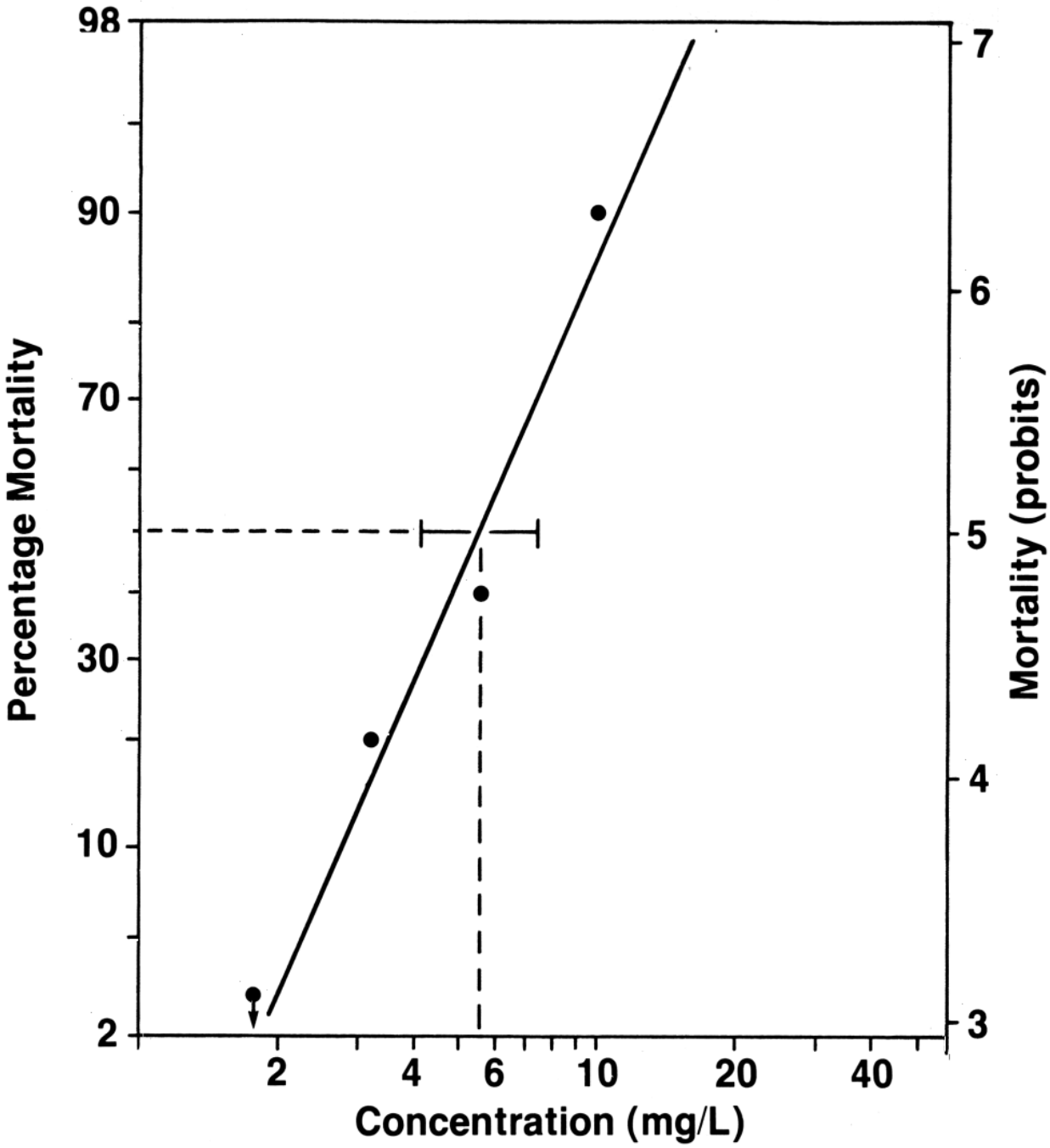


Figure 3 Estimating a Median Lethal Concentration by Plotting Mortalities on Logarithmic-Probability Paper

SAS (1988) probit analysis: 5.58 (4.26 and 7.40)

TOXSTAT (1996) method (version 3.5):
- probit: 5.58 (4.38 and 7.12)
- Spearman-Kärber, zero trim: 5.64 (4.40 and 7.23)
- logit: 5.63 (4.39 and 7.22)

Table 4.2 in EC (2004b) provides additional examples of computed data for acute quantal tests using various computer programs.

6.4.2 ICp

When a multi-concentration test for effects of prolonged (56-day) exposure of earthworms to spiked soil mixtures is conducted, the *quantitative* data representing reproductive inhibition and growth inhibition must be used to calculate the *ICp* (*inhibiting concentration for a specified percent effect*) for each of these endpoints, data permitting (see Sections 4.3.7 and 6.2). The ICp is a quantitative estimate of either:

(1) the concentration causing a fixed percent reduction in the mean number of juveniles produced by the adult worms during the test; or

(2) the concentration causing a fixed percent reduction in the mean dry weight of juvenile worms at test end.

It is calculated as a specified percent reduction for each endpoint (e.g., the IC25 and/or IC20, which represent 25% and 20% reduction). The desired value of *p* is selected by the investigator, and 25% or 20% is currently favoured. Any ICp that is calculated and reported must include the 95% confidence limits.

In the analyses of reproductive performance, the number of juveniles produced in each replicate is used to calculate the average number of surviving juveniles produced per treatment (concentration) in relation to the average number produced in the *negative control* replicates. A value of zero is assigned for a number of juveniles in a replicate, if the adult earthworms in that replicate died before producing progeny. If one or both of the adult worms died during the test, after producing young, the number of juveniles produced is still to be used in the analyses. If there are no surviving juveniles in a replicate (test chamber), it contributes a value of zero to the calculation used to obtain the average number of survivors for that treatment (concentration). If there are no surviving juveniles in all replicates at a given concentration, that concentration is still included in the analysis, using an average value of zero juveniles.

The mean weight of individual juveniles in each replicate (test chamber) is calculated as the total dry weight of the juveniles that survived in the test chamber divided by the number of juveniles that survived in that chamber to the end of the test (Section 4.3.7). The mean weights from all the replicates of a given treatment (concentration) are used to calculate the average for the treatment; this is the average individual dry weight of surviving juveniles per concentration. It is compared to the average individual dry weight in the *negative control*, obtained by the same procedure. If there are no surviving juveniles in a replicate, that replicate does not contribute to the average for the treatment. If there are no survivors in all replicates at a given concentration, that concentration does not have an average weight of surviving juveniles and cannot be used in the analysis for comparison with the average weight in the negative control.

As indicated in Section 4.3.7, separate ICps for mean number of surviving progeny produced in each treatment, and their mean dry weights, must be calculated and reported (data permitting) upon completion of a 56-day multi-concentration test with *E. andrei*. These calculations must be made using the appropriate linear or nonlinear regression analyses (see Section 6.4.2.1). If, however, regression analyses fail to provide meaningful ICps for the mean number of live progeny produced and/or their mean dry weights, the ICPIN analyses described in Section 6.4.2.2 should be applied to the corresponding data.

6.4.2.1 Use of regression analysis. Upon completion of a definitive 56-day (or, in certain instances, 63-day) multi-concentration test with *E. andrei*, separate ICps (including their respective 95% confidence limits) for the mean number of surviving progeny produced in each treatment, and their individual mean dry weights, must be calculated using linear and/or nonlinear regression procedures. These values may be calculated using a series of linear and nonlinear regression models (data permitting) proposed by Stephenson *et al*. (2000b) that have been re-parameterized, based on techniques applied by van Ewijk and Hoekstra (1993), to automatically generate the ICp and its 95% confidence limits for any value of 'p' (e.g., IC25 or IC50). The proposed models for application consist of one linear model, and the following four nonlinear

regression models: exponential, Gompertz, logistic, and logistic adjusted to accommodate hormesis[90]. Further guidance on the use of these linear and nonlinear regression models for calculating ICps is provided by Stephenson (2003a) and Stephenson *et al*. (2000b). The reader is also strongly advised to consult EC (2004b) for additional guidance on the general application of linear and non-linear regression for the analysis of quantitative toxicity data. Instruction for the appropriate application of linear and non-linear regression, using Version 11.0 of the statistical program SYSTAT[91], is provided in Appendix J. However, any statistical software capable of linear and nonlinear regression may be used when calculating the respective ICps and their associated 95% confidence limits. Appendix J provides instruction on the use of regression models to derive the most appropriate ICps for reduced numbers of surviving progeny and their reduced mean dry weight.

The five models recommended for application follow. Further information on these specific models is presented in Appendix J.

Exponential model: $Y = a \times (1 - p)^{(C \div ICp)}$

where:

- Y = number or mass of progeny
- a = the y-intercept (i.e., the control response)
- p = desired value for 'p' (e.g., 0.25 for a 25% inhibition)
- C = the test concentration as a logarithm
- Icp = the ICp for the data set

Gompertz model: $Y = t \times \exp[\log(1 - p) \times (C \div ICp)^b]$

where:

- Y = number or mass of progeny
- t = the y-intercept (i.e., the control response)
- p = desired value for 'p' (e.g., 0.25 for a 25% inhibition)
- C = the test concentration as a logarithm
- ICp = the ICp for the data set
- b = a scale parameter (estimated to be between 1 and 4) that defines the shape of the equation

Hormesis model: $Y = t \times [1 + h \times C] \div \{1 + [(p + h \times ICp) / (1 - p)] * (C/ICp)^b\}$

where:

- Y = number or mass of progeny
- t = the y-intercept (i.e, the control response)
- h = describes the hormetic effect (estimated to be small, usually between 0.1 and 1)
- C = the test concentration as a logarithm
- p = desired value for 'p' (e.g., 0.25 for a 25% inhibition)
- ICp = the ICp for the data set
- b = a scale parameter (estimated to be between 1 and 4) that defines the shape of the equation

Linear model: $Y = [(-b \times p) \div ICp] \times C + b$

where:

- Y = number or mass of progeny
- b = the y-intercept (i.e., the control response)
- p = desired value for 'p' (e.g., 0.25 for a 25% inhibition)
- ICp = the ICp for the data set
- C = the test concentration as a logarithm

Logistic model: $Y = t \div \{1+[p \div (1-p)] \times (C \div ICp)^b\}$

where:

- Y = number or mass of progeny
- t = the y-intercept (i.e, the control response)
- p = desired value for 'p' (e.g., 0.25 for a 25% inhibition)
- C = the test concentration as a logarithm
- ICp = the ICp for the data set
- b = a scale parameter (estimated to be between 1 and 4) that defines the shape of the equation

---

[90] A hormetic response (i.e., hormesis) might be found in sublethal observations at the lowest concentration(s), i.e., performance at such concentration(s) is enhanced relative to that in the negative control. For instance, there might be more progeny produced in low concentrations than in the control, or the weights of individuals might be higher than in the control. This is not a flaw in the testing, but it is a real biological phenomenon. To calculate the ICp when this phenomenon occurs, the data should be analyzed using the hormesis model. The hormetic effects are included in the regression, but do not bias the estimate of the ICp. An estimated IC25 would still represent a 25% reduction in performance from that of the control.

[91] The latest (e.g., Version 11.0) version of SYSTAT™ is available for purchase by contacting SYSTAT Software, Inc., 501 Canal Boulevard, Suite C, Point Richmond, CA 94804-2028, USA, phone no. 1-800-797-7401; see Web site www.systat.com/products/Systat/.

The general process for the statistical analysis and selection of the most appropriate regression model (linear or non-linear) for quantitative toxicity data is outlined in Figure 4. The selection process begins with an examination of a scatter plot or line graph of the test data to determine the shape of the concentration-response curve. The shape of the curve is then compared to available models so that one or more appropriate model(s) that best suits the data is (are) selected for further examination (see Figure J.1, Appendix J, for an example of five potential models).

Once the appropriate model(s) is (are) selected for further consideration, assumptions of normality and *homoscedasticity* of the *residuals* are assessed. If the regression procedure for one or more of the examined models meets the assumptions, the data (and regression) are examined for the presence of outliers. If an outlier has been observed, the test records and experimental conditions should be scrutinized for human error. If there are one or more outliers present, the analysis should be performed with and without the outlier(s), and the results of the analyses compared to examine the effect of the outlier(s) on the regression. Thereafter, a decision must be made as to whether the outlier(s) should be removed from the final analysis. The decision should take into consideration natural biological variation, and biological reasons that might have caused the apparent anomaly. Additional guidance on the presence of outliers and unusual observations is provided in Appendix J (Section J.2.4) as well as in EC (2004b). If there are no outliers present or none are removed from the final analysis, the model that demonstrates the smallest residual mean square error is selected as the model of best choice. Additional guidance from a statistician familiar with dealing with outlier data is also advised.

Normality should be assessed using the *Shapiro-Wilk's test* as described in EC (2004b). A normal probability plot of the residuals may also be used during the regression procedure, but is not recommended as a stand-alone test for normality as the detection of a 'normal' or 'non-normal' distribution depends on the subjective assessment of the user. If the data are not normally distributed, then the user is advised to try another model, consult a statistician for further guidance on model selection or to perform the less-desirable linear interpolation (using ICPIN, see Section 6.4.2.2) method of analysis.

Homoscedasticity of the residuals should be assessed using *Levene's test* as described in EC (2004b), and by examining the graphs of the residuals against the actual and predicted (estimated) values. Levene's test provides a definite indication of whether the data are homogeneous (e.g., as in Figure J.2A of Appendix J) or not. If the data (as indicated by Levene's test) are *heteroscedastic* (i.e., not homogeneous), then the graphs of the residuals should be examined. If there is a significant change in the variance and the graphs of the residuals produce a distinct fan or 'V' pattern (refer to Figure J.2B, Appendix J for an example), then the data analysis should be repeated using weighted regression. Before choosing the weighted regression, the standard error of the ICp is compared to that derived from the unweighted regression. If there is a difference of greater than 10% between the two standard errors[1], then the weighted regression is selected as the regression of best choice. However, if there is less than a 10% difference in the standard error between the weighted and unweighted regressions, then the user should consult a statistician for the application of additional models, given the test data, or the data could be re-analyzed using the less-desirable linear interpolation (using ICPIN, see Section 6.4.2.2) method of analysis. This comparison between weighted and unweighted regression is completed for each of the selected models while proceeding through the process of final model selection (i.e., model and regression of best choice). Some non-divergent patterns might be indicative of an inappropriate or incorrect model (refer to Figure J.2C, Appendix J, for an example), and the user is again urged to consult a statistician for further guidance on the application of additional models.

---

[92] The value of 10% is only a "rule-of-thumb" based upon experience. Objective tests for the improvement due to weighting are available, but beyond the scope of this document. Weighting should be used only when necessary, as the procedure can introduce additional complications to the modeling procedure. A statistician should be consulted when weighting is necessary.

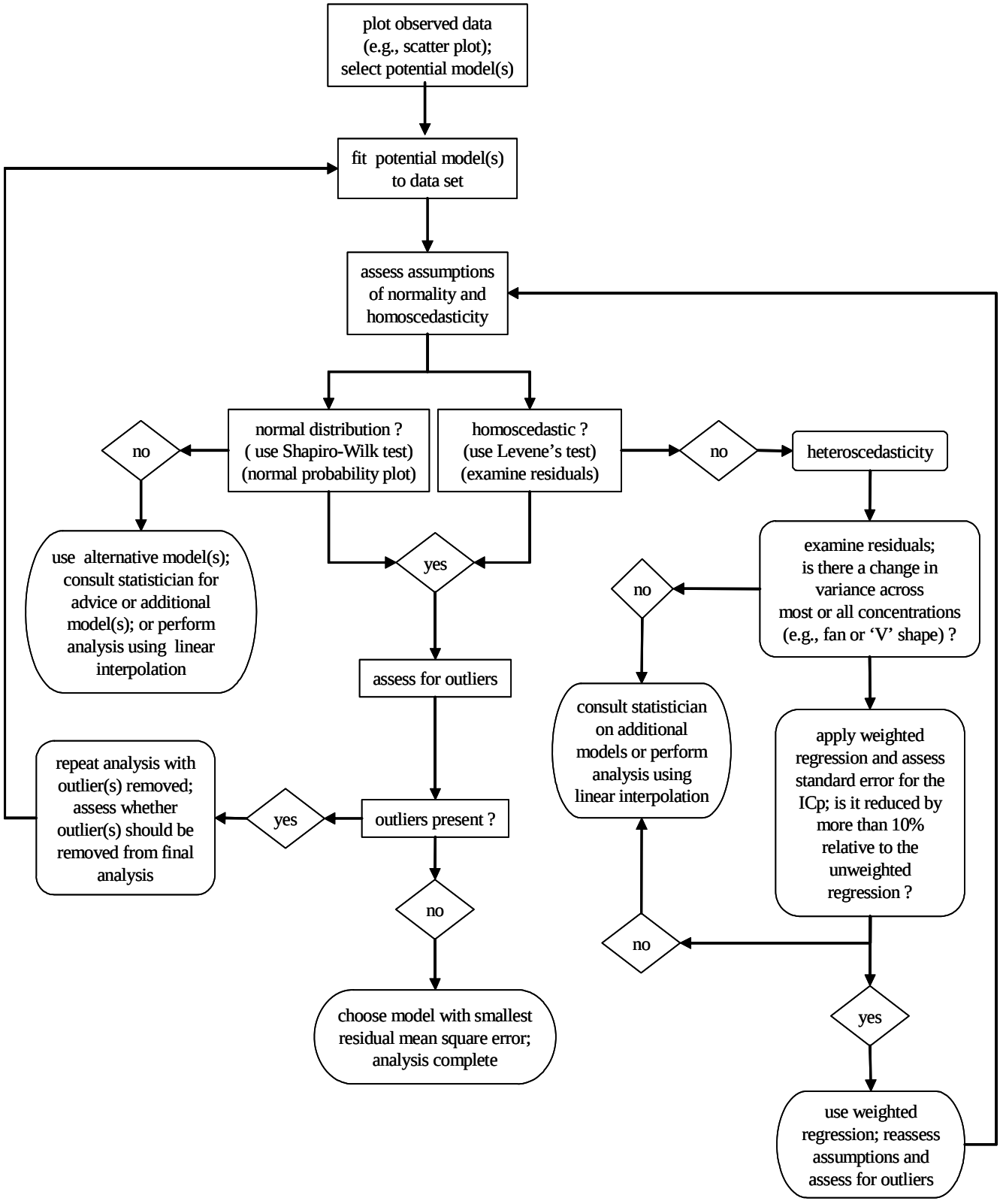


Figure 4 **The General Process for the Statistical Analysis and Selection of the Most Appropriate Model for Quantitative Toxicity Data (adapted and modified from Stephenson *et al.* 2000b)**

6.4.2.2 Linear interpolation using ICPIN. If regression analyses of the endpoint data (see Section 6.4.2.1) fail to provide an acceptable ICp for growth and/or reproductive inhibition, linear interpolation using the computer program called *ICPIN* should be applied. This program (Norberg-King, 1993; USEPA, 1994b, 1995) is not proprietary, is available from the USEPA, and is included in most computer software for environmental toxicology, including TOXSTAT. The original instructions for ICPIN from the USEPA are clearly written and make the program easy to use (Norberg-King, 1993).[93] An earlier version was called BOOTSTRP.

Analysis by ICPIN does not require equal numbers of replicates in different concentrations. The ICp is estimated by smoothing of the data as necessary, then using the two data-points adjacent to the selected ICp (USEPA, 1994b, Appendix L; USEPA, 1995, Appendix L). The ICp cannot be calculated unless there are test concentrations both lower and higher than the ICp; both those concentrations should have an effect reasonably close to the selected value of p, preferably within 20% of it. At present, the computer program does not use a logarithmic scale of concentration, and so Canadian users of the program must enter the concentrations as logarithms. Some commercial computer packages have the logarithmic transformation as a general option, but investigators should make sure that it is actually retained when proceeding to ICPIN. ICPIN estimates confidence limits by a special "bootstrap" technique because usual methods would not be valid. Bootstrapping performs many resamplings from the original measurements. The investigator must specify the number of resamplings, which can range from 80–1000. At least 400 is recommended here, and 1000 would be beneficial.[94]

If there are several adjacent high concentrations with no surviving juveniles, only the lowest of that string of concentrations should be used in analysis (i.e. the concentration closest to the middle of the series of concentrations used in the test). Normally, there is no particular benefit to including the additional concentrations, because they offer nothing to the analysis (i.e., the data consist only of zero progeny and provide no mean weights thereof) .

Besides determining and reporting the computer-derived ICps for reproduction and weight of juveniles at test end, a graph of percent reduction in number of live juveniles produced or percent reduction of dry weight should be plotted against the logarithm of concentration, to check the mathematical estimations and to provide visual assessments of the nature of the data (EC, 2004b).

If the ICPIN program is used when there is a hormetic effect, an inherent smoothing procedure could change the control value and bias the estimate of ICp. Accordingly, before statistical analysis, hormetic values at low concentration(s) should be arbitrarily replaced by the control value. This is considered a temporary expedient until a superior approach is established (EC 2004b). The correction is applied for any test concentration in which the average effect (i.e., the geometric average of the replicate means) is higher ("better") than the average for the control. To apply this correction, replace the observed mean weights (or mean numbers of progeny) of the replicates in the hormetic concentration(s), with the means of replicates in the control. The geometric average for that/those concentration(s) will then be the same as that for the control.

---

[93] The instructions in Norberg-King (1993) are sometimes misleading on the identity of "replicates". The term is used in such a way that it would apply to numbers or weights of individual organisms within the same chamber. This slip of wording does not affect the functioning of the program. Some commercial programs have been less user-friendly for entry of data and analysis.

[94] ICPIN has some deficiencies, which is why it is recommended only in cases where use of regression fails to provide an acceptable ICp. Its interpolation method is an inefficient use of data, sensitive to peculiarities of the two concentrations used. The program fails to adopt logarithm of concentration, which would introduce a slight bias towards a higher value of ICp. A modification of the bootstrap method has now remedied a problem of overly narrow confidence limits; however, regression analyses provide better methods of estimating the ICp and its 95% confidence limits (EC, 2004b) (see Section 6.4.2.1).

Reporting Requirements

Each test-specific report must indicate if there has been any deviation from any of the *must* requirements delineated in Sections 2 to 6 and, if so, provide details of the deviation(s). The reader must be able to establish from the test-specific report whether the conditions and procedures preceding and during the test rendered the results valid and acceptable for the use intended.

Section 7.1 provides a list of items that must be included in each test-specific report. A list of items that must either be included in the test-specific report, provided separately in a general report, or held on file for a minimum of five years, is found in Section 7.2. Specific monitoring programs, related test protocols, or regulations might require selected test-specific items listed in Section 7.2 (e.g., details about the test material and/or explicit procedures and conditions during sample collection, handling, transport, and storage) to be included in the test-specific report, or might relegate certain test-specific information as *data to be held on file*.

Procedures and conditions common to a series of ongoing tests (e.g., routine toxicity tests for monitoring or compliance purposes) and consistent with specifications in this document, may be referred to by citation or by attachment of a general report that outlines standard laboratory practice.

Details on the procedures, conditions, and findings of the test, which are not conveyed by the test-specific report or general report, must be kept on file by the laboratory for a minimum of five years so that the appropriate information can be provided if an audit of the test is required. Filed information might include:

- a record of the chain-of-continuity for field-collected or other samples tested for regulatory or monitoring purposes;
- a copy of the record of acquisition for the sample(s);
- chemical analytical data on the sample(s) not included in the test-specific report;
- bench sheets for the observations and measurements recorded during the test;
- bench sheets and warning chart(s) for the reference toxicity tests;
- detailed records of the source of the test organisms, their taxonomic confirmation, and all pertinent information regarding their culturing and/or holding and acclimation and health; and
- information on the calibration of equipment and instruments.

Original data sheets must be signed or initialled, and dated by the laboratory personnel conducting the tests.

7.1 Minimum Requirements for a Test-Specific Report

The following items must be included in each test-specific report.

7.1.1 Test Substance or Material

- brief description of sample type (e.g., waste sludge, reference or contaminated field-collected soil, negative control soil) or coding, as provided to the laboratory personnel;
- information on labelling or coding of each sample; and
- date of sample collection; date and time sample(s) received at test facility.

7.1.2 Test Organisms

- species and source of brood stock and test organisms;
- wet weight (mean ± SD), at start of test; and
- any unusual appearance, behaviour, or treatment of the organisms, before their use in the test.

7.1.3 Test Facilities

- name and address of test laboratory; and
- name of person(s) performing the test (or each component of the test).

7.1.4 Test Method

- citation of biological test method used (i.e., as per this document);
- design and description if specialized procedure(s) (e.g., preparation of mixtures of spiked soil; preparation and use of solvent and, if so, solvent control) or modification(s) of the standard test method described herein;
- brief description of frequency and type of all measurements and all observations made during test; and
- name and citation of program(s) and methods used for calculating statistical endpoints.

7.1.5 Test Conditions and Procedures

- design and description of any deviation(s) from, or exclusion of, any of the procedures and conditions specified in this document;
- number of discrete samples per treatment; number of replicate test chambers for each treatment; number and description of treatments in each test including the control(s); test concentrations (if applicable);
- volume of soil in each test chamber;
- number of organisms per test chamber and treatment;
- dates when test was started and ended;
- feeding regime and ration, for 56-day test;
- date when adults were removed from test chambers, for 56-day (or longer) test;
- for each soil sample — any measurements of soil particle size, moisture content, water-holding capacity, and pH; and
- for each composite sample of subsamples taken at the same time from all replicates of each treatment — all measurements of temperature, pH, moisture content, and water-holding capacity.

7.1.6 Test Results

- *for an acute lethality test* — mean (± SD) percent survival in each treatment on Days 0, 7 (if determined), and 14;
- *for an acute avoidance test* — mean (± SD) number of surviving worms in replicates of each treatment representing *clean* soil and test soil, at 48 h if *E. andrei* or *E. fetida* or at 72 h if *L. terrestris*;
- *for a test of effects due to prolonged exposure* — mean (± SD) percent survival of adult worms in each treatment on Day 28; mean (± SD) number of surviving juveniles in each treatment on Day 56; mean (± SD) dry weight of individual juveniles surviving in each treatment on Day 56; mean (± SD) number of surviving juveniles produced by each adult worm in *negative control soil* (and in *positive control soil* and/or *solvent control soil*, if used), on Day 56;
- any LC50 or EC50 (including the associated 95% confidence limits and, if calculated, the slope) determined; any additional LCx or ECx (e.g., LC20 or EC20) calculated;
- any ICp (together with its 95% confidence limits) determined for the data on reproductive success (i.e., number of surviving juvenile worms in each treatment at test end); any ICp (together with its 95% confidence limits) determined for the data on growth (i.e., dry weight of individual juveniles surviving at test end); details regarding any transformation of data, and indication of quantitative statistical method used or procedures applied to the data;
- for a multi-concentration test with chemical-spiked soil, indication as to whether results are based on nominal or measured concentrations of chemical(s) or chemical product(s); all values for measured concentrations;
- results for any seven-day LC50 (including its 95% confidence limits) performed with the reference toxicant in conjunction with the definitive soil

toxicity test, using the same lot (group) of test organisms; *geometric mean* value (± 2 SD) for the same reference toxicant and test species, as derived at the test facility in previous seven-day LC50s using the procedures and conditions for reference toxicity tests described herein; and

- anything unusual about the test, any problems encountered, any remedial measures taken.

7.2 Additional Reporting Requirements

The following list of items must be either included in the test-specific report or the general report, or held on file for a minimum of five years.

7.2.1 Test Substance or Material

- identification of person(s) who collected and/or provided the sample;
- records of sample chain-of-continuity and log-entry sheets; and
- conditions (e.g., temperature, in darkness, in sealed container) of sample upon receipt and during storage.

7.2.2 Test Organisms

- name of person(s) who identified the organisms and the taxonomic guidelines used to confirm species;
- history and age of brood stock, for any culture used to provide test organisms;
- description of culture conditions and procedures, including temperature, lighting, type and amount of substrate and details on its periodic renewal, measurements of substrate quality, density of worms, feeding regime and quantity, records of health and performance indices;
- history of any population of test organisms obtained from an outside source, including specifics related to the period(s) of holding and acclimation before their use in the test, type and amount of substrate and details on its periodic renewal, measurements of substrate quality, density of worms, feeding regime and quantity, records of health and performance indices;
- procedures used to count, handle, sort, and transfer animals; and those to determine their mortality, condition, appearance, and behaviour; and
- source and composition of food, procedures used to prepare and store food, feeding method(s), feeding frequency, and ration.

7.2.3 Test Facilities and Apparatus

- all results for initial tests with negative control soil and reference toxicant, undertaken by the laboratory previously inexperienced with performing the biological test method described herein in advance of any reporting of definitive test results (see Section 3.1.1).
- description of systems for providing lighting and for regulating temperature within test facility;
- description of test chambers and covers; and
- description of procedures used to clean or rinse test apparatus.

7.2.4 Negative Control Soil or Reference Soil

- procedures for the preparation (if artificial soil) or *pretreatment* (if natural soil) of *negative control soil*;
- source of natural soil; history of past use and records of analysis for pesticides or other contaminants;
- formulation of artificial soil, including sources for the constituents and conditions and procedures for hydration and pH adjustment; and
- storage conditions and duration before use.

7.2.5 Test Method

- procedures used for mixing or otherwise manipulating test soils before use; time interval between preparation and testing;
- procedure used in preparing stock and/or test solutions of chemicals; description and concentration(s) of any solvent used;

- details concerning aliquot sampling, preparation, and storage before physicochemical analysis, together with available information regarding the analytical methods used (with citations); and

- use and description of preliminary or range-finding test.

7.2.6 Test Conditions and Procedures

- measurements of light intensity adjacent to surface of soil in test chambers;

- procedure for adding test organisms to test chambers;

- appearance of each sample (or mixture thereof) in test chambers; changes in appearance noted during test;

- records of each spraying of test water on the surface of the soil in each test chamber throughout an eight-week test, for increasing moisture content;

- any other physicochemical measurements (e.g., analyses of aliquots from the same *batch* to determine conductivity, homogeneity, contaminant concentration, total volatile solids, biochemical oxygen demand, chemical oxygen demand, total inorganic carbon, cation exchange capacity, oxidation-reduction potential, total nitrogen) made before and during the test on test material (including *negative control soil* and *reference soil*) and contents of test chambers, including analyses of whole soil and pore water;

- any other observations or analyses made on the test material (including samples of *negative control soil* or *reference soil*); e.g., qualitative and/or quantitative data regarding indigenous macrofauna or detritus, or results of geochemical analyses; and

- any chemical analyses of the concentration of chemical in stock solution(s) of reference toxicant and, if measured, in test concentrations.

7.2.7 Test Results

- results for any range-finding test(s) conducted;

- *for an acute lethality test* — percent survival of worms in each test chamber on Days 0, 7 (if determined), and 14;

- *for an acute avoidance test* — total number of surviving worms in *clean* soil and test soil within each test unit at 48 h if *E. andrei* or *E. fetida* or at 72 h if *L. terrestris*;

- *for an eight-week test* — number of surviving adult worms in each test chamber on Day 28; number of surviving juveniles in each test chamber on Day 56; group dry weight and mean individual dry weight of juveniles surviving in each test chamber on Day 56; for regression analyses, hold on file information indicating sample size (e.g., number of replicates per treatment), parameter estimates with variance, any ANOVA table(s) generated, plots of fitted and observed values of any models used, and the output provided by the statistical program (e.g., SYSTAT);

- warning chart showing the most recent and historic results for acute toxicity tests with the reference toxicant;

- results for eight-week test(s) performed with a reference toxicant;

- graphical presentation of data; and

- original bench sheets and other data sheets, signed and dated by the laboratory personnel performing the test and related analyses.

References

ACECSS (Agriculture Canada Expert Committee on Soil Survey), "The Canadian System of Soil Classification", Publication 1646, Second Edition, ISBN 0-660-12460-2, Research Branch, Agriculture Canada, Ottawa, ON (1987).

AESA (Alberta Environmentally Sustainable Agriculture Program), "AESA Soil Quality Benchmark Study — Soil Organic Matter", Fact Sheet FS2001-ISQ, 4 p., AESA Soil Quality Program, Conservation and Development Branch, Alberta Agriculture, Food and Rural Development, Edmonton, Alberta (2001).

Aquaterra Environmental, "Development of Earthworm Toxicity Tests for Assessment of Contaminated Soils", report prepared by Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (1998a).

———. "Development of a Reproduction Toxicity Test with *Onychiurus folsomi* for Assessment of Contaminated Soils", report prepared by Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (1998b).

———. "Development of Plant Toxicity Tests for Assessment of Contaminated Soils", report prepared by Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (1998c).

Aquaterra Environmental, "Unpublished Data", derived by Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2001).

Aquaterra Environmental and ESG (Ecological Services Group), "Assessment of the Biological Test Methods for Terrestrial Plants and Soil Invertebrates: Metals", August 2000 Final Report, prepared by Aquaterra Environmental, Ltd. (Orton, ON) and ESG International Inc. (Guelph, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2000).

ASTM (American Society for Testing and Materials), "Standard Guide for Use of Lighting in Laboratory Testing", E-1733-95, p. 1249–1259, In: *1999 Book of ASTM Standards, Volume 11.05 Biological Effects and Environmental Fate; Biotechnology; Pesticides, Vol. 11.05*, ASTM, Philadelphia, PA (1999a).

———. "Standard Guide for Conducting Laboratory Soil Toxicity or Bioaccumulation Tests with the Lumbricid Earthworm *Eisenia fetida*", E1676-97, p. 1062–1079, In: *1999 Annual Book of ASTM Standards — Biological Effects and Environmental Fate; Biotechnology; Pesticides, Vol. 11.05*, ASTM, Philadelphia, PA (1999b).

Barber, I., J. Bembridge, P. Dohmen, P. Edwards, F. Heimbach, R. Heusel, K. Romijn, and H. Rufli, "Development and Evaluation of Triggers for Earthworm Toxicity Testing with Plant Protection Products", p. 269–278, In: *Advances in Earthworm Ecotoxicology*, S.C. Sheppard, J.D. Bembridge, M. Holmstrup, and L. Posthuma (eds.), SETAC Press, Pensacola, FL (1997).

.

Bauer, C. and J. Römbke, "Factors Influencing the Toxicity of Two Pesticides on Three Lumbricid Species in Laboratory Tests", *Soil Biol. Biochem., 29*:705–708 (1997).

BBA (Biologische Bundesanstalt für Land- und Forstwirtschaft), "Auswirkungen von Pflanzenschutzmitteln auf die Reproduktion und das Wachstum von *Eisenia fetida*/*Eisenia andrei*", *Richtlinien für die Prüfung von Pflanzenschutzmitteln im Zulassungsverfahren*, Teil VI, 2-2, Januar 1994 (1994).

Becker-van Slooten, K, S. Campiche, and J. Tarradellas, "Research in Support of the Environment Canada Collembolan Toxicity Test Method with *Folsomia candida* for Assessment of Contaminated Soils", report prepared by the Laboratory of Environmental Chemistry and Ecotoxicology (CECOTOX), Ecole polytechnique fédérale de Lausanne (EPFL) ENAC-ISTE (Lausanne, Switzerland) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2003).

———, "Research in Support of the Environment Canada Soil Toxicity Test Methods: Consideration of Various Methods for Measuring Soil pH, Water Holding Capacity, and Water Filled Pore Space", report prepared by Laboratory of Environmental Chemistry and Ecotoxicology (CECOTOX), Ecole polytechnique fédérale de Lausanne (EPFL), ENAC-ISTE (Lausanne, Switzerland) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004).

Bonnell Environmental Consulting, "Assessment of Soil Toxicity Test Species for Canadian Representativeness", Technical Report TS-28, prepared for the Method Development and Applications Section, Environment Canada, Ottawa, ON (1994).

Bouché, M.B., "Earthworm Toxicological Test, Hazard Assessment and Biomonitoring — A Methodological Approach", p. 315–320, In: *Earthworms in Waste and Environmental Management*, C.A. Edwards and E.F. Neuhauser, (eds.), SPB Academic Publishing, The Hague, The Netherlands (1988).

———. "Earthworm Species and Ecotoxicological Studies", p. 20–35, In: *Ecotoxicology of Earthworms*, P.W. Greig-Smith, H. Becker, P.J. Edwards, and F. Heimbach (eds.), Intercept, Andover, UK (1992).

Brousseau, P., N. Fugere, J. Bernier, D. Coderre, D. Nadeau, G. Poirier, and M. Fournier, "Evaluation of Earthworm Exposure to Contaminated Soil by Cytometric Assay of Coelomocytes Phagocytosis in *Lumbricus terrestris* (Oligocaheta)", *Soil Biol. Biochem., 29*:681–684 (1997).

Callahan, C.A., "Earthworms as Ecotoxicological Assessment Tools", p. 295–301, In: *Earthworms in Waste and Environmental Management*, C.A. Edwards and E.F. Neuhauser (eds.), SPB Academic Publishing, The Hague, The Netherlands (1988).

Callahan, C.A., C.A. Menzie, D.E. Burmaster, D.C. Wilborn, and T. Ernst, "On-Site Methods for Assessing Chemical Impact on the Soil Environment Using Earthworms: A Case Study at the Baird and McGuire Superfund Site, Holbrook, Massachusetts", *Environ. Toxicol. Chem., 10*:817–826 (1991).

Callahan, C.A., M.A. Shirazi, and E.F. Neuhauser, "Comparative Toxicity of Chemicals to Earthworms", *Environ. Toxicol. Chem., 13 (2)*:291–298 (1994).

Carter, M.R. (ed.), *Soil Sampling and Methods of Analysis*", Lewis Publishers, CRC Press Inc., Boca Raton, Florida (1993).

Chang, L.W., J.R. Meier, and M.K. Smith, "Application of Plant and Earthworm Bioassays to Evaluate Remediation of a Lead-contaminated Soil", *Arch. Environ. Contam. Toxicol., 32*:166–171 (1997).

Christensen, O.M. and J.G. Mather, "Earthworms as Ecotoxicological Test-Organisms", Technical Report, Danish Environmental Protection Agency, Copenhagen, Denmark, 99 p., (1994).

Cikutovic, M.A., L.C. Fitzpatrick, B.J. Venables, and A.J. Goven, "Sperm Count in Earthworms (*Lumbricus terrestris*) as a Biomarker for Environmental Technology: Effects of Cadmium and Chlordane", *Environ. Pollut., 81*:123–125 (1993).

Conder, J.M. and R.P. Lanno, "Evaluation of Surrogate Measures of Cadmium, Lead, and Zinc Bioavailability to *Eisenia fetida*, *Chemosphere 41*: 1659–1668 (2000).

Cooke, A.S., P.W. Greig-Smith, and S.A. Jones, "Consequences for Vertebrate Wildlife of Toxic Residues in Earthworm Prey", p. 139–155, In: *Ecotoxicology of Earthworms*, P.W. Greig-Smith, H. Becker, P.J. Edwards, and F. Heimbach (eds.), Intercept, Hants, UK (1992).

Courchesne, F., S. Savoie, and A. Dufresne, "Effects of Air-Drying on the Measurement of Soil pH in Acidic Forest Soils of Quebec, Canada", Soil Science *160(1):*56–68 (1995).

Crépin, J. and R.L. Johnson, "Soil Sampling for Environmental Assessment", p. 5–18, Chapter 2, In: *Soil Sampling and Methods of Analysis*, M.R. Carter, (ed.), Lewis Publishers, CRC Press Inc., Boca Raton, Florida (1993).

Curry, J.P., "The Ecology of Earthworms in Reclaimed Soils and their Influence on Soil Fertility", p. 251–261, In: *Earthworms in Waste and*

Environmental Management, C.A. Edwards and E.F. Neuhauser, eds., SPB Academic Publishing, The Hague, The Netherlands (1988).

Deitzer G. "Spectral Comparisons of Sunlight and Different Lamps" *c74*, p. 197–199, In: *Proceedings of International Lighting in Controlled Environments Workshop*, Tibbits, T.W. (ed.), 1 Mar. 1994, Madison, WI (1994).

Drewes, C.D., E.P. Vining, and C.A. Callahan, "Non-Invasive Electrophysiological Monitoring: A Sensitive Method for Detecting Sublethal Neurotoxicity in Earthworms", *Environ. Toxicol. Chem., 3*:599–607 (1984).

EC (Environment Canada), "Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/30, 56 p. (1995).

———. "Biological Test Method: Test for Survival and Growth in Sediment Using the Freshwater Amphipod *Hyalella azteca*", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/33, 123 p. (1997a).

———. "Biological Test Method: Test for Survival and Growth in Sediment Using Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*)", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/32, 131 p. (1997b).

———. "Guidance Document on Application and Interpretation of Single-species Tests in Environmental Toxicology", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/34, 203 p. (1999).

———. "The Procurement and Culturing of Red Wiggler Worms (*Eisenia andrei*), Standard Operating Procedure CULTEF.SOP, 6 p., December 2000, Environmental Toxicology Section, Pacific Environmental Science Centre, North Vancouver, BC (2000a).

———. "14-Day Acute Lethal Bioassay Using the Earthworm *Eisenia andrei* ", Standard Operating Procedure LC50EF1.2, 10 p., December 2000, Environmental Toxicology Section, Pacific Environmental Science Centre, North Vancouver, BC (2000b).

———. "Biological Test Method: Test for Survival and Growth in Sediment Using Spionid Polychaete Worms (*Polydora cornuta*)", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/41, 108 p. (2001).

———. "Inter-laboratory Validation of Environment Canada's New Test Methods for Measuring Soil Toxicity Using Earthworms", October, 2004, report prepared by J. Princz for the Biological Methods Division, Environment Canada, Ottawa, ON (2004a).

———. "Guidance Document on Statistical Methods for Environmental Toxicity Tests", Method Development and Applications Section, Environment Canada, Ottawa, ON, Report EPS 1/RM/46, in preparation (2004b).

———. "Proceedings to the Workshop on Toxicity Test Methodologies for Assessing the Impacts of Contaminant Mixtures in Soil Using Terrestrial Species of Ecological Relevance to Canadian Soil Systems", Workshop held February 19–21, 2003 at the Pacific Environmental Science Centre, North Vancouver, BC, report published by the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004c).

———. "Biological Test Method: Test for Measuring Survival and Reproduction Effects in Springtails", Method Development and Applications Section, Environment Canada, Ottawa, ON, Report EPS 1/RM/47, in preparation (2004d).

———. "Biological Test Method: Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil", Method Development and Applications Section, Environment Canada, Ottawa, ON, Report EPS 1/RM/45, in preparation (2004e).

Edwards, C.A. and P.J. Bohlen, "The Effects of Toxic Chemicals on Earthworms", *Reviews of Environmental Contamination and Toxicology, 125*: 23–99 (1992).

———. *Biology and Ecology of Earthworms*, 3rd ed., Chapman and Hall, London, UK (1996).

Edwards, C.A. and J.R. Lofty, *Biology of Earthworms*, Chapman and Hall, London, UK (1977).

Eisenhart, C., M.W. Hastay, and W.A. Wallace, *Techniques of Statistical Analysis*, Chapter 15, McGraw-Hill Book Co., New York, NY (1947).

Elshayeb, M., N.C. Feisthauer, G.L. Stephenson, and R.J. Brooks, "Avoidance Response of Earthworms to Benomyl Spiked Soils", Proceedings of the Laurentian Chapter of the Society of Environmental Toxicology and Chemistry, McMaster University, Hamilton ON, May 25–26, 2001 (2001).

ESG (Ecological Services Group), "Final Report on the Acute Screening and Definitive, Chronic Toxicity Tests with Motor Gasoline", report prepared by Ecological Services Group International Inc. (Guelph, ON) for Petroleum Technology Alliance Canada, Calgary, Alberta (2000).

———. "Toxicity of Petroleum Hydrocarbons to Soil Organisms and The Effects on Soil Quality: Phase 1 Fraction-Specific Toxicity of Crude Oil", report prepared by Ecological Services Group International Inc. (Guelph, ON) for Petroleum Technology Alliance Canada, Calgary, Alberta (2001).

———. "Impacts of Metal-Contaminated Forest Soils From the Canadian Shield on Terrestrial Organisms", report prepared by Ecological Services Group International Ltd. (Guelph, ON) for Metals in the Environment Research Network Canadian Network of Toxicology Centres, Guelph, ON and the Method Development and Applications Section, Environment Canada, Ottawa, ON (2002).

ESG (Ecological Services Group) and Aquaterra Environmental, "Assessment of the Biological Test Methods for Terrestrial Plants and Soil Invertebrates: Pesticides", report prepared by Ecological Services Group International Inc. (Guelph, ON) and Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2002).

ESP (Ecological Services for Planning), "The Relative Toxicity of Four Pesticides to Two Species of Earthworm in Artificial and Natural Soils", unpublished report, prepared by Ecological Services for Planning Inc. (Guelph, ON) for the Guidelines Division, Environment Canada, Ottawa, ON (1992).

Feisthauer, N.C., G.L. Stephenson, and R.P. Scroggins, "Toxicity of Pesticides (Benomyl and Diuron) to a Battery of Terrestrial Species in Different Soil Types", Presentation No. PM066, Proceedings of the Society of Environmental Toxicology and Chemistry, Baltimore, MD, November 2001 (2001).

Fender, W.M., "Earthworms of the Western United States, Part 1. Lumbricidae", *Megadriologica, 4 (5)*:93–129 (1985).

Ferreiro, A., J. Dominguez, and A. Velado, "The Taxonomy and Reproductive Behaviour of *Eisenia fetida* and *Eisenia andrei* (Lumbricidea)", p. 245, In: *Proceedings from the 7th International Symposium on Earthworm Ecotoxicology*, 387 p., Cardiff, Wales, September 1–6, 2002 (2002).

Fischer, E. and L. Molnar, "Growth and Reproduction of *Eisenia fetida* (Oligochaeta, Lumbricidae) in Semi-Natural Soil Containing Various Metal Chlorides", *Soil Biol. Biochem., 29*:667–670 (1997).

Fitzpatrick, L.C., R. Sassani, B.J. Venables, and A.J. Goven, "Comparative Toxicity of Polychlorinated Biphenyls to Earthworms *Eisenia foetida* and *Lumbricus terrestris*", *Environ. Pollut., 77*:65–69 (1992).

Fox, C., "Frequently Asked Questions About Earthworms", FAQ Menu (Web site) http://res2.agr.ca/london/pmrc/faq/earthwor.html", Research Branch, Agriculture and Agri-Food Canada, Southern Crop Protection and Food Research Centre, London, ON (2000).

Gibbs, M.H., L.F. Wicker, and A.J. Stewart, "A Method for Assessing Sublethal Effects of Contaminants in Soils to the Earthworm, *Eisenia foetida*", *Environ. Toxicol. Chem., 15 (3)*:360–368 (1996).

Giggleman, M.A., L.C. Fitzpatrick, A.J. Goven, and B.J. Venables, "Effects of Pentachlorophenol on Survival of Earthworms (*Lumbricus terrestris*) and Phagocytosis by their Immmunoactive Coelomocytes", *Environ. Toxicol. Chem., 17 (12)*:2391–2394 (1998).

Goven, A.J., G.S. Eyambe, L.C. Fitzpatrick, B.J. Venables, and E.L. Cooper, "Cellular Biomarkers for Measuring Toxicity of Xenobiotics: Effects of

Polychlorinated Biphenyls on Earthworm *Lumbricus terrestris* Coelomocytes", *Environ. Toxicol. Chem., 12*:863–870 (1993).

Goven, A., J. S.C. Chen, L.C. Fitzpatrick, and B.J. Venables, "Lysozyme Activity in Earthworm (*Lumbricus terrestris*) Coelomic Fluid and Coelomocytes: Enzyme Assay for Immunotoxicity of Xenobiotics, *Environ. Toxicol. Chem., 13*:607–613 (1994).

Haimi, J. and S. Paavola, "Responses of Two Earthworm Populations with Different Exposure Histories to Chlorophenol Contamination", *Environ. Toxicol. Chem., 17 (6)*:1114–1117 (1998).

Hartenstein, R., "Effect of Aromatic Compounds, Humic Acids and Lignins on Growth of the Earthworm (*Eisenia foetida*), *Soil. Biol. Biochem. 14*:595–599 (1982).

Hausenbuiller, R.L., *Soil Science–Principles & Practices*, 3rd Edition, W.C. Brown Publisher, Dubuque, Iowa (1985).

Heimbach, F., "Correlations Between Three Methods for Determining the Toxicity of Chemicals to Earthworms", *Pestic. Sci., 15*:605–611 (1984).

———. "Comparison of Laboratory Methods, Using *Eisenia foetida* and *Lumbricus terrestris*, for the Assessment of the Hazard of Chemicals to Earthworms", *J. Plant Diseases and Protection, 92 (2)*:186–193 (1985).

———. "A Comparison of Laboratory Methods for Toxicity Testing with Earthworms", p. 329–335, In: *Earthworms in Waste and Environmental Management*, C.A. Edwards and E.F. Neuhauser, (eds.), SPB Academic Publishing, The Hague, The Netherlands (1988).

———. "Use of Laboratory Toxicity Tests for the Hazard Assessment of Chemicals of Earthworms Representing the Soil Fauna", p. 299–302, In: *Integrated Soil and Sediment Research: A Basis for Proper Protection*, H.J.P. Eijsackers and T. Hamers, (eds.), Kluwer Academic Publishers, Netherlands (1993).

———. "Comparison of the Sensitivities of an Earthworm (*Eisenia fetida*) Reproduction Test and a Standardized Field Test on Grassland", p. 235–245, In: *Advances in Earthworm Ecotoxicology*, S.C. Sheppard, J.D. Bembridge, M. Holmstrup, and L. Posthuma (eds.), SETAC Press, Pensacola, FL (1997).

Heimbach, F. and P.J. Edwards, "The Toxicity of 2-Chloroacetamide and Benomyl to Earthworms Under Various Test Conditions in an Artificial Soil Test", *Pestic. Sci., 14*:635–636 (1983).

Hendershot W.H., H. Lalonde, and M. Duquette, "Soil Reaction and Exchangeable Acidity", p. 141–145, In: *Soil Sampling and Methods of Analysis*, M.R. Carter, (ed.), Canadian Society of Soil Science, Lewis Publishers, Boca Raton, Florida (1993).

Holmstrup, M., "Field Assessment of Toxic Effects on Reproduction in the Earthworms *Aporrectodea longa* and *Aporrectodea rosea*", *Environ. Toxicol. Chem., 19*:1781–1787 (2000).

Hund, K., "Earthworm Avoidance Test for Soil Assessment: Alternative for Acute and Reproduction Test", p. 1039–1040, In: *Contaminated Soil '98*, Vol. 2, Proceedings of the Sixth International FZK/TNO Conference on Contaminated Soil, 17–21 May 1998, Edinburgh, UK, Thomas Telford Publishing Ltd., London, UK (1998).

Hund-Rinke, K. and H. Wiechering, "Earthworm Avoidance Test for Soil Assessments. An Alternative for Acute and Reproduction Tests", *J. Soils and Sediments, 1*:15–20 (2001).

Ingraldi, S., J. Princz, and R. Scroggins, "Comparison of *E. andrei* and *E. fetida* and their Response to Contamination", unpublished technical report prepared by S. Ingraldi (Faculty of Science, University of Waterloo, Waterloo, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004).

ISO (International Organization for Standardization), "Soil — Determination of the Effect of Chemical Substances on the Reproduction of Earthworms", proposal from the Netherlands, 9 p., April 1991, Geneva, Switzerland (1991).

———. "Soil Quality — Effects of Pollutants on Earthworms (*Eisenia fetida*). Part 1: Determination

of Acute Toxicity Using Artificial Soil Substrate", ISO 11268-1, Geneva, Switzerland (1993).

———. "Soil Quality — Determination of Organic and Total Carbon After Dry Combustion (Elementary Analysis)", ISO 10694, Geneva, Switzerland (1995).

———. "Soil Quality — Effects of Pollutants on Earthworms (*Eisenia fetida*). Part 2: Determination of Effects on Reproduction", ISO 11268-2, Geneva, Switzerland (1998).

———. "Soil Quality — Inhibition of Reproduction of Collembola (*Folsomia candida*) by Soil Pollutants", ISO 11267, Geneva, Switzerland (1999).

———. "Avoidance Test for Evaluating the Quality of Soils and the Toxicity of Chemicals — Test with Earthworms (*Eisenia fetida/ andrei*", Doc ISO/TC/190/SC 4/WG 2 N1238, Draft Document, 19 p., Geneva, Switzerland (2003).

Jaenike, J., "*Eisenia foetida* is Two Biological Species", *Megadrliogica, 4:*6–8 (1982).

Karnak, R.E. and J.L. Hamelink, "A Standardized Method for Determining the Acute Toxicity of Chemicals to Earthworms", *Ecotoxicol. Environ. Safety, 6*:216–222 (1982).

Keddy, C., J.C. Greene, and M.A. Bonnell, "Review of Whole-Organism Bioassays: Soil, Freshwater Sediment, and Freshwater Assessment in Canada", *Ecotoxicol. Environ. Safety, 30*:221–251 (1995).

Keith, L.H., *Environmental Sampling and Analysis: A Practical Guide*, 143 p., Lewis Publishers, Inc., Chelsea, Michigan (1992).

Klute, A. (ed.), *Methods of Soil Analysis, Part 1 – Physical and Mineralogical Methods*, 2nd ed., American Society of Agronomy Inc. and Soil Science Society of America Inc., Madison, Wisconisin (1986).

Komex International, "Sampling and Shipping of Reference Soil for the Terrestrial Soil Toxicity Method Development Project" report prepared for Environment Canada, Ottawa, ON, 17 p. (1995).

Kula, H., "Comparison of Laboratory and Field Testing for the Assessment of Pesticide Side Effects on Earthworms", *Acta Zool. Fennica, 196*:228–341 (1995).

Kula, H. and C. Kokta, "Side Effects of Selected Pesticides on Earthworms Under Laboratory and Field Conditions", *Soil Biol. Biochem., 24*:1711–1714 (1992).

Kula, H. and O. Larink, "Tests on the Earthworms *Eisenia fetida* and *Aporrectodea caligniosa*", p. 95–112, In: *Handbook of Soil Invertebrate Toxicity Tests*, H. Løkke and C.A.M. van Gestel (eds.), Ecological and Environmental Toxicology Series, John Wiley & Sons, Chichester, UK (1997).

Lee, K.E., *Earthworms: Their Ecology and Relationship with Soils and Land Use*, Academic Press, London, UK (1985).

Leon, C.D. and C.A.M. van Gestel, "Selection of a Set of Laboratory Ecotoxicity Tests for the Effects Assessment of Chemicals in Terrestrial Ecosystems — Discussion Paper", Report No. D94004, 134 p., August 1994, Department of Ecology and Ecotoxicology, Vrije Universiteit, Amsterdam, The Netherlands (1994).

Lofs-Holmin, A., "Measuring Growth of Earthworms as a Method of Testing Sublethal Toxicity of Pesticides—Experiments with Benomyl and Trichloroacetic Acid (TCA)", *Swed. J. Agr. Res., 10*:25–33 (1980).

Lofs-Holmin, A. and U. Bostrom, "The Use of Earthworms and Other Soil Animals in Pesticide Testing", p. 303–313, In: *Earthworms in Waste and Environmental Management*, C.A. Edwards and E.F. Neuhauser (eds.), Academic Publishing, The Hague, The Netherlands (1988).

Macdonald, D.W., "Predation on Earthworms by Terrestrial Vertebrates", p. 393–414, In: J.E. Stachell (ed.), *Earthworm Ecology: From Darwin to Vermiculture*, Chapman and Hall, New York, NY (1983).

McCann, J., "Report on the Identification of Earthworm Test Cultures (*Eisenia andrei* and *E. fetida*) and Potential Commercial Sources of Earthworms", report prepared by the Department of

Biology, University of Waterloo (Waterloo, ON) and Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004).

McElroy, T.C. and W.J. Diehl, "Heterosis in Two Closely Related Species of Earthworm (*Eisenia fetida* and *E. andrei*)", *Heredity, 87*:598–608 (2001).

Mather, J.G. and O. Christensen, "Surface Migration of Earthworms in Grassland", *Pedobiologia, 36*:51–57 (1992).

Mearns, A.J., R.C. Swartz, J.M. Cummins, P.A. Dinnel, P. Plesha, and P.M. Chapman, "Interlaboratory Comparison of a Sediment Toxicity Test Using the Marine Amphipod, *Rhepoxynius abronius*", *Marine Environ. Res., 19*:13–37 (1986).

Meharg, A.A., R.F. Shore, and K. Broadgate, "Edaphic Factors Affecting the Toxicity and Accumulation of Arsenate in the Earthworm *Lumbricus terrestris*, *Environ. Toxicol. Chem., 17(6)*:1124–1131 (1998).

Meier, J.R., L.W. Chang, S. Jacobs, J. Torsella, M.C. Meckes, and M.K. Smith, "Use of Plant and Earthworm Bioassays to Evaluate Remediation of Soil from a Site Contaminated with Polychlorinated Biphenyls", *Environ. Toxicol. Chem., 16*:928–938 (1997).

Menzie, C.A., D.E. Burmaster, J.S. Freshman, and C.A. Callahan, "*Assessment of Methods for Estimating Ecological Risk in the Terrestrial Component: A Case Study at the Baird & McGuire Superfund Site in Holbrook, Massachusetts*", *Environ. Toxicol. Chem., 11*:245–260 (1992).

NERI (National Environmental Research Institute), "Manual of SECOFASE — Development, Improvement and Standardization of Test Systems for Assessing Sublethal Effects on Chemicals on Fauna in the Soil Ecosystem", H. Lokke and C.A.M. van Gestel (eds.), report from a Workshop held in Silkeborg January 18–19, 1993, Silkeborg, Denmark 41 p. (1993).

Neuhauser, E.F. and C.A. Callahan, "Growth and Reproduction of the Earthworm *Eisenia fetida fetida* Exposed to Sublethal Concentrations of Organic Chemicals, *Soil Biol. Biochem., 22 (2)*:175–179 (1990).

Neuhauser, E.F., M.R. Malecki, and R.C. Loehr, "Growth and Reproduction of the Earthworm *E. fetida* after Exposure to Sublethal Concentrations of Metals", *Pedobiologia, 27*:89–97 (1984).

Neuhauser, E.F., R.C. Loehr, M.R. Malecki, D.L. Milligan, and P.R. Durkin, "The Toxicity of Selected Organic Chemicals to the Earthworm *Eisenia fetida*", *J. Environ. Qual., 14* (3):383–388 (1985).

Neuhauser, E.F., P.R. Durkin, M.R. Malecki, and M. Anatra, "Comparative Toxicity of Ten Organic Chemicals to Four Earthworm Species", *Comp. Biochem. Physiol., 83C (1)*:197–200 (1986).

Newman, M.C., *Quantitative Methods in Aquatic Ecotoxicology*. Lewis Publishers., Boca Raton, Florida (1995).

Norberg-King, T.J., "A Linear Interpolation Method for Sublethal Toxicity: the Inhibition Concentration (ICp) Approach (Version 2.0)", USEPA, Environ. Res. Lab.-Duluth, Duluth, MN, Tech. Report 03-93 of National Effluent Toxicity Assessment Center (1993).

OECD (Organization for Economic Cooperation and Development), "OECD Guideline for Testing of Chemicals: Earthworm, Acute Toxicity Tests", No. 207, Paris, France (1984).

———. "OECD Guideline for the Testing of Chemicals: Proposal for a New Guideline — Earthworm Reproduction Test (*Eisenia fetida/andrei*), draft document, 17 p., January 2000 (2000).

Øien, N. and J. Stenersen, "Esterases of Earthworms–III. Electrophoresis Reveals that *Eisenia fetida* (Sav.) is Two Species", *Comp. Biochem. Physiol. 78C*:277–282 (1984).

OMAFRA (Ontario Ministry of Agriculture, Food and Rural Affairs), "Field Crop Recommendations 1999-2000", Publication 296, 142 p., Queen's Printer, Toronto, ON (1999).

OMEE (Ontario Ministry of Environment and Energy), "Guidance on Sampling and Analytical Methods for Use at Contaminated Sites in Ontario — Section 8.3.7 Inorganics", Version 1.1, ISBN-0-7778-4056-1, December 1996, Queen's Printer, Toronto, ON (1996).

Paine, M.D. and C.A. McPherson, "Phase IV Studies of 10-Day Tests for Sediment Toxicity Using Marine or Estuarine Infaunal Amphipods", unpublished report prepared for the Marine Environment Division, Environment Canada and IGATG by EVS Consultants Ltd., North Vancouver, BC (1991).

Petersen, R.G. and L.D. Calvin, "Sampling", p. 33–51, In: *Methods of Soil Analysis, Part 1 – Physical and Mineralogical Methods* (A. Klute, ed.), 2nd ed., American Society of Agronomy Inc. and Soil Science Society of America Inc., Madison, Wisconisin (1986).

Puurtinen, H.M. and E.A.T. Martikainen, "Effect of Soil Moisture on Pesticide Toxicity to an Enchytraeid Worm, *Enchytraeus* sp.", *Arch. Environ. Contam. Toxicol., 33*:34–41 (1997).

Reinecke, A.J. and S.A. Viljoen, "A Comparison of the Biology of *Eisenia fetida* and *Eisenia andrei* (Oligochaeta)", *Biol. Fert. Soils, 11*: 295-300 (1991).

Reinecke, A.J., S.A. Viljoen, and R.J. Saayman, "The Suitability of *Eudrilus eugeniae, Perionyx excavatus* and *Eisenia fetida* (Oligochaeta) for Vermicomposting in Southern Africa in Terms of their Temperature Requirements", *Soil Biol. Biochem., 24*:1295–1307 (1992).

Reinecke, A.J. and S.A. Reinecke, "The Influence of Heavy Metals on the Growth and Reproduction of the Compost Worm *Eisenia fetida* (Oligocaheta)", *Pedobiologia, 40*:439–448 (1996).

Reynolds, J.W., "The Earthworms (Lumbricidae and Sparganophilidae) of Ontario", Life Sci. Misc. Publ., Royal Ontario Museum, Toronto, ON (1977).

———. "Status of Exotic Earthworm Systematics and Biogeography in North America, p. 1–28, In: *Earthworm Ecology and Biogeography*, P.F. Hendrix (ed.), Lewis Publishers, CRC Press Inc., Boca Raton, FL (1995).

Riepert, F. and C. Kula, "Development of Laboratory Methods for Testing Effects of Chemicals and Pesticides on *Collembola* and Earthworms", *Mitteilungen aus der Biologischen Bundesantalt für Land- und Forstwirtschaft Berlin-Dahlem, 320*:1–82 (1996).

Robidoux, P.Y., J. Hawari, S. Thiboutot, G. Ampleman, and G.I. Sunahara, "Acute Toxicity of 2,4,6-Trinitrotoluene in Earthworm (*Eisenia andrei*)", *Ecotoxicol. Environ. Safety, 44*:311–321 (1999).

Robidoux, P.Y., C. Svendsen, J. Caumartin, J. Hawari, G. Ampleman, S. Thiboutot, J.M. Weeks, and G.I. Sunahara, "Chronic Toxicity of Energetic Compounds in Soil Determined Using the Earthworm (*Eisenia andrei*) Reproduction Test", *Environ. Toxicol. Chem., 19*:1764–1773 (2000).

Robidoux, P.Y., J. Hawari, S. Thiboutot, G. Ampleman, and G.I. Sunahara, "Chronic Toxicity of Octahydro-1,3,5,7-Tetranitro-1,3,5,7-Tetrazocine (HMX) in Soil Determined Using the Earthworm (*Eisenia andrei*) Reproduction Test", *Environ. Pollut., 111*:283–292 (2001).

Rocchini, R.J., M.J.R. Clark, A.J. Jordan, S. Horvath, D.J. McLeay, J.A. Servizi, A. Sholund, H.J. Singleton, R.G. Watts, and R.H. Young, "Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish", B.C. Ministry of Environment, Victoria, BC (1982).

Römbke, J., P. Vickus, and C. Bauer, "Experiences and Problems with the OECD-Earthworm Acute Test in Routine Testing", p. 209–212, In: *Ecotoxicology of Earthworms*, P.W. Greig-Smith, H. Becker, C.A. Edwards, and F. Heimbach (eds.), Intercept Ltd., Andover, UK (1992).

Römbke, J., P, T. Knacker, B. Förster, and A. Marcinkowski, "Comparison of Effects of Two Pesticides on Soil Organisms in Laboratory Tests, Microcosms, and in the Field", p. 229–240, In: *Ecotoxicology of Soil Organisms*, M.H. Donker, H. Eijsackers and F. Heimbach (eds.), Lewis Publishers, Chelsea, UK (1994).

Sager J.C. and C. McFarlane, "Radiation", p. 1–30, In: *Plant Growth Chamber Handbook*, R.W. Langhans and T.W. Tibbits (eds.), North Central Regional Research Publication No. 340, Iowa Agriculture and Home Economics Experiment Station Special Report No. 99, Iowa State University of Science and Technology, Ames Iowa (1997).

SAH (Soil Analysis Handbook), *Reference Methods for Soil Analysis*, Soil and Plant Analysis Council

Inc., Georgia University Station, Athens, Georgia, 202 p. (1992).

SAS (SAS Institute Inc.), "SAS Procedures Guide, Release 6.03, *and* Additional SAS/STAT Procedures, Release 6/03 (SAS Technical Report P-179). SAS Institute Inc., Cary, NC (1988).

Saterbak, A., R.J. Toy, D.C.L. Wong, B.J. McMain, M.P. Williams, P.B. Dorn, L.P. Brzuzy, E.Y. Chai, and J.P. Salanitro, "Ecotoxicological and Analytical Assessment of Hydrocarbon-contaminated Soils and Application to Ecological Risk Assessment", *Environ. Toxicol. Chem., 18(7)*:1591–1607 (1999).

Saterbak, A., R.J. Toy, B.J. McMain, M.P. Williams, and P.B. Dorn, "Ecotoxicological and Analytical Assessment of Effects of Bioremediation on Hydrocarbon-containing Soils", *Environ. Toxicol. Chem., 19*:2643–2652 (2000).

Sauvé, S., N. Cook, W.H. Hendershot, and M.B. McBride, "Linking Plant Tissue Concentrations and Soil Copper Pools in Urban Contaminated Soils", *Environ. Pollution 94(2)*:153–157 (1996).

Sauvé, S., A. Dumestre, M.B. McBride, and W.H. Hendershot, "Derivation of Soil Quality Criteria Using Predicted Chemical Speciation of Pb and Cu", *Environ. Toxicol. Chem., 17(8)*:1481–1489 (1998).

Sauvé, S., W.A. Norvell, M.B. McBride, and W.H. Hendershot, "Speciation and Complexation of Cadmium in Extracted Soil Solutions", *Environ. Science & Technol., 34*:291–296 (2000).

Schaefer, M., "Behavioural Endpoints in Earthworm Ecotoxicology: Evaluation of Different Test Systems in Soil Toxicity Assessment", *J. Soils & Sediments, 3 (2):*79–84 (2003).

Scott-Fordsmand, J.J., J.M. Weeks, and S.P. Hopkin, "Importance of Contamination History for Understanding Toxicity of Copper to Earthworm *Eisenia fetida* (Oligochaeta: Annelida), Using Neutral-Red Retention Assay", *Environ. Toxicol. Chem., 19*:1774–1780 (2000).

Sheppard, P.S., "Specific Differences in Cocoon and Hatchling Production in *Eisenia fetida* and *E. andrei*", p. 83–92, In: *Earthworms in Waste and Environmental Management*, C.A. Edwards and E.F. Neuhauser, eds., SPB Academic Publishing, The Hague, The Netherlands (1988).

Sheppard, S.C., "Development of Bioassay Protocols for Toxicants in Soil", Report PIBS 2838, ISBN 0-7778-1440-X, Environmental Research, Research and Technology Branch, Ontario Environment, Queen's Printer, Toronto, ON, February 1994 (1994).

Sheppard, S.C. and W.G. Evenden, "Optimized Design for Earthworm Survival Tests in Soil", *Bull. Environ. Contam. Toxicol., 49*:648–655 (1992).

———. "An Approach to Defining a Control or Diluent Soil for Ecotoxicity Assays", pp. 215-226, In: *Environmental Toxicology and Risk Assessment*, Seventh Volume, ASTM STP 1333, E.E. Little, A.J. DeLonay, and B.M. Greenberg (eds.), American Society for Testing and Materials, Philadelphia, PA (1998).

Sims, R.W. and B.M. Gerard, "Earthworms — Synopses of the British Fauna (New Series)". Report No. 31, the Linnaean Society of London and the Estuarine and Brackish-Water Sciences Association, London, UK (1985).

Slimak, K.M., "Avoidance Response as a Sublethal Effect of Pesticides on *Lumbricus terrestris* (Oligochaeta)", *Soil Biol. Biochem., 29*:713–715 (1997).

Sokal, R.R. and F.J. Rohlf, *Biometry*, W.H. Freement and Co., San Francisco, Calif. (1969).

Spurgeon, D.J. and S.P. Hopkin, "Extrapolation of the Laboratory-Based OECD Earthworm Toxicity Test to Metal-Contaminated Field Sites", *Ecotoxicology, 4*:190–205 (1995).

———. "Effects of Metal-Contaminated Soils on the Growth, Sexual Development, and Early Cocoon Production of the Earthworm *Eisenia fetida*, with Particular Reference to Zinc", *Ecotoxicol. Environ. Safety, 35*:86–95 (1996a).

———. "Effects of Variations of the Organic Matter Content and pH of Soils on the Availability and Toxicity of Zinc to the Earthworm *Eisenia fetida*", *Pedobiologia,,40*:80–96 (1996b).

Spurgeon, D.J., S.P. Hopkin, and D.T. Jones, "Effects of Cadmium, Copper, Lead and Zinc on Growth, Reproduction and Survival of the Earthworm *Eisenia fetida* (Savigny): Assessing the Environmental Impact of Point-Source Metal Contamination in Terrestrial Ecosystems", *Environ. Pollut., 84*:123–130 (1994).

Stantec and Aquaterra Environmental, "Developmental Studies in Support of Environment Canada's Biological Test Methods for Measuring Soil Toxicity Using Earthworms", report prepared by Stantec Consulting, Ltd. (Guelph, ON) and Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004).

Steel, R.G.D. and J.H. Torrie, *Principles and Procedures of Statistics*, McGraw-Hill Book Co., New York, NY (1980).

Stephan, C.E., "Methods for Calculating an LC50"", p. 65–84 , In: *Aquatic Toxicology and Hazard Evaluation*, F.L. Mayer and J.L. Hamelink (eds). ASTM STP 634, American Society For Testing and Materials, Philadelphia, PA (1977).

Stephenson, G.L. Unpublished Data. ESG International Inc., Guelph, ON (2002).

———. "Terrestrial Test Methods for Plants and Soil Invertebrates", 230 p., Ph.D. Thesis, Department of Environmental Biology, University of Guelph, Guelph, ON (2003a).

———. Unpublished Data. Stantec Consulting, Ltd., Guelph, ON (2003b).

Stephenson, G.L., C.D. Wren, I.C.J. Middelraad, and J.E. Warner, "Exposure of the Earthworm, *Lumbricus terrestris*, to Diazinon, and the Relative Risk to Passerine Birds", *Soil Biol. Biochem., 29 (3/4)*:717–720 (1997).

Stephenson, G.L., A. Kaushik, N.K. Kaushik, K.R. Solomon, T. Steele, and R.P. Scroggins, "Use of an Avoidance-Response Test to Assess the Toxicity of Contaminated Soils to Earthworms", p. 67–81, In: *Advances in Earthworm Ecotoxicology*, S.C. Shepard, J.D. Bembridge, M. Holmstrup, and L. Posthuma, eds., SETAC Press, Society of Environmental Toxicology and Chemistry, Philadelphia, PA (1998).

Stephenson, G.L., N. Koper, J. McCann, and Z. Wang, "Draft Report on the Acute Screening and Definitive, Chronic Toxicity Tests with Whole Federated Crude Oil", unpublished (draft) report, October 1999, prepared by ESG International (Guelph, ON) for Petroleum Technology Alliance Canada (Calgary, AB) and the Canadian Association of Petroleum Producers (Calgary, AB), 52 p. + app. (1999a).

———. "Draft Report on the Acute Screening and Definitive, Chronic Toxicity Tests with Fraction 3 Derived From Federated Crude Oil", unpublished (draft) report, December 1999, prepared by ESG International (Guelph, ON) for Petroleum Technology Alliance Canada (Calgary, AB) and the Canadian Association of Petroleum Producers (Calgary, AB), 48 p. + app. (1999b).

Stephenson, G.L., J. McCann, P. Jokuty, and Z. Wang, "Draft Report on the Acute Screening and Definitive, Chronic Toxicity Tests with Fraction 2 Derived From Federated Crude Oil", unpublished (draft) report, January 2000, prepared by ESG International (Guelph, ON) for Petroleum Technology Alliance Canada (Calgary, AB) and the Canadian Association of Petroleum Producers (Calgary, AB), 32 p. + app. (2000a).

Stephenson, G.L., N. Koper, G.F. Atkinson, K.R. Solomon, and R.P. Scroggins, "Use of Nonlinear Regression Techniques for Describing Concentration-Response Relationships of Plant Species Exposed to Contaminated Site Soils", *Environ. Toxicol. Chem., 19*:2968–2981 (2000b).

Stephenson, G.L., J.I. Princz, N. Koper, and P.G. Miasek, "Terrestrial Toxicity Testing with Volatile Substances", pp. 236–252, In: *Environmental Toxicology and Risk Assessment: Science, Policy, and Standardization – Implications for Environmental Decisions*, 10th Vol., ASTM STP 1403, BM Greenberg, RN Hull, MH Roberts, Jr., and RW Gensemer (eds.), American Society for Testing and Materials, West Conshohocken, PA (2001).

Stephenson, G.L., R.G. Kuperman, G.L. Linder, and S. Visser, "Toxicity Tests for Assessing Contaminated Soils and Ground Water", p. 25–43, In: *Environmental Analysis of Contaminated Sites*, G.I. Sunahara, A.Y. Renoux, C. Thellen, C.L. Gaudet, and A. Pilon (eds.), Ecological and Environmental Toxicology Series, John Wiley & Sons Ltd., Chichester, UK (2002).

Suzuki, M.M., E.L. Cooper, G.S. Eyambe, A.J. Goven, L.C. Fitzpatrick, and B.J. Venables, "Polychlorinated Biphenyls (PCBs) Depress Allogeneic Natural Cytotoxicity by Earthworm Coelomocytes", *Environ. Toxicol. Chem., 14*:1697–1700 (1995).

Tomlin, A.D., ""Behaviour as a Source of Earthworm Susceptibility to Ecotoxicants", p. 116–125, In: *Ecotoxicology of Earthworms*, P.W. Greig-Smith, H. Becker, P.J. Edwards, and F. Heimbach (eds.), Intercept, Hants, UK (1995).

Tomlin, A.D., R. Protz, R.R. Martin, D.C. McCabe, and R.J. Lagace, "Relationships Amongst Organic Matter Content, Heavy Metal Concentrations, Earthworm Activity, and Soil Microfabric on a Sewage Sludge Disposal Site", *Geoderma, 57*:89–103 (1993).

TOXSTAT, "Version 3.5", Lincoln Research Associates, Inc., PO Box 4276, Bisbee, Arizona 85603, USA, phone 520 432-4092, email danlra@msn.com. [Programs on disk, with printed user's manual.] (1996).

Tsukamoto, J. and H. Watanabe, "Influence of Temperature on Hatching and Growth of *Eisenia foetida* (Oligochaeta, Lumbricidae)", *Pedobiologia, 17(8)*:338–342 (1977).

USEPA (United States Environmental Protection Agency), "Protocols for Short-term Toxicity Screening of Hazardous Waste Sites", report EPA/600/3-88/029, prepared by J.C. Greene, C.L. Bartels, W.J. Warren-Hicks, B.R. Parkhurst, G.L. Linder, S.A. Peterson, and W.E. Miller, February 1989, Corvallis Environ. Research Lab., Corvallis, OR (1989).

———. "ASSESS Users Guide", Report EPA/600/8-91001, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada (1991).

———. "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates", report EPA 600/R-94/024, 133 p., USEPA, Duluth, MN (1994a).

———. "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms (Second Edition)", Report EPA/600/4-91/003, 483 p., Office of Research and Development, USEPA, Cincinnati, OH (1994b).

———. "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms", Report EPA/600/R-95/136, 661 p., Office of Research and Development, USEPA, Washington, DC (1995).

———. "Ecological Effects Test guidelines — Earthworm Subchronic Toxicity Tests", draft Report EPA 712-C-96-167, April 1996, Office of Prevention, Pesticides, and Toxic Substances (OPPTS) 850.6200, Washington, DC (1996).

van Ee, J.J., L.J. Blume, and T.H. Starks, "A Rationale for the Assessment of Errors in the Sampling of Soils", Report EPA/600/4-90/013, 57 p., United States Environmental Protection Agency, Las Vegas, Nevada (1990).

van Ewijk, P.H. and H.A. Hoekstra, "Calculation of the EC50 and its Confidence Interval when Subtoxic Stimulus is Present", *Ecotoxicol. Environ. Saf., 25*: 25–32 (1993).

van Gestel, C.A.M., "Earthworms in Ecotoxicology", Ph.D. Thesis, 197 p., Utrecht University, Utrecht, The Netherlands (1991).

van Gestel, C.A.M., "Validation of Earthworm Toxicity Tests by Comparison with Field Studies: A Review on Benomyl, Carbendazim, Carbofuran and Carbaryl", *J. Ecotoxicol. Environ. Safety, 23*:221–236 (1992).

van Gestel, C.A.M., "Scientific Basis for Extrapolating Results from Soil Ecotoxicity Tests to Field Conditions and the Use of Bioassays", p. 25–50, In: *Ecological Risk Assessment of Contaminants in Soil*, N.M. Van Straalen and H. Løkke (eds.), Chapman and Hall, London, UK (1997).

van Gestel, C.A.M. and W.A. van Dis, "The Influence of Soil Characteristics on the Toxicity of Four Chemicals to the Earthworm *Eisenia fetida andrei* (Oligochaeta)", *Biol. Fertil. Soils, 6*:262–265 (1988).

van Gestel, C.A.M., W.A. van Dis, E.M. van Breemen, and P.M. Sparenburg, "Comparison of Two Methods for Determining the Viability of Cocoons Produced

in Earthworm Toxicity Experiments", *Pedobiologia, 32*:367–371 (1988).

van Gestel, C.A.M., W.A. van Dis, E.M. van Breemen, and P.M. Sparenburg, "Development of a Standardized Reproduction Toxicity Test with the Earthworm Species *Eisenia fetida andrei* Using Copper, Pentachlorophenol, and 2,4-Dichloroaniline", *Ecotoxicol. Environ. Safety, 18*: 305–312 (1989).

van Gestel, C.A.M., E.M. Dirven-van Breemen, and R. Baerselman, "Influence of Environmental Conditions on the Growth and Reproduction of the Earthworm *Eisenia andrei* in an Artificial Soil Substrate", *Pedobiologia, 36*:109–120 (1992a).

van Gestel, C.A.M., E.M. Dirven-Van Breemen, R. Baerselman, H.J.B. Emans, J.A.M. Jannsen, R. Postuma, and P.J.M. Van Vliet, "Comparison of Sublethal and Lethal Criteria for Nine Different Chemicals in Standardized Toxicity Tests Using the Earthworm *Eisenia andrei*, *Ecotoxicol. Environ. Safety, 23*:206–220 (1992b).

van Gestel, C.A.M., J.J. van der Waarde, J.G.M. Derksen, E.E. van der Hoek, M.E.X.W. Veul, S. bouwens, B. Rusch, R. Kronenburg, and G.N.M. Stokman, "The Use of Acute and Chronic Bioassays to Determine the Ecological Risk and Bioremediation Efficiency of Oil-Polluted Soils", *Environ. Toxicol. Chem., 20*:1438–1449 (2001).

Venter, J.M. and A.J. Reinecke, "Sublethal Ecotoxicological Effects of Dieldrin on the Earthworm *Eisenia foetida* (Oligochaeta)", p. 337–353, In: *Earthworms in Waste and Environmental Management* (C.A. Edwards and E.F. Neuhauser, eds.), SPB Academic Publishing, The Hague, The Netherlands (1988).

Wallwork, J.A., *Earthworm Biology*, Studies in Biology No. 161, Camelot Press Ltd., Southampton, UK, 58 p. (1983).

Webster, R. and M.A. Oliver, *Statistical Methods in Soil and Land Resource Survey*, Oxford University Press, Oxford, UK 315 p. (1990).

Wells, J.B. and R.P. Lanno, "Passive Sampling Devices (PSDs) as Biological Surrogates for Estimating the Bioavailability of Organic Chemicals in Soil", p. 253–270, In: *Environmental Toxicology and Risk Assessment: Science, Policy, and Standardization - Implications for Environmental Decisions: Tenth Volume*, B.M. Greenberg, R.N. Hull, M.H. Roberts Jr., and R.W. Gensemer (eds.), ASTM STP 1403, American Society for Testing and Materials, West Conshohocken, PA (2001).

Wentsel, R.S. and M.A. Guelta, "Avoidance of Brass Powder-contaminated Soil by the Earthworm, *Lumbricus terrestris*", *Environ. Toxicol. Chem., 7*: 241–243 (1988).

Yeardley, R.B. Jr., J.M. Lazorchak, and L.C. Gast, "The Potential of an Earthworm Avoidance Test for Evaluation of Hazardous Waste Sites", *Environ. Toxicol. Chem., 15(9)*:1532–1537 (1996).

Zoran, M.J., T.J. Heppner, and C.D. Drewes, "Teratogenic Effects of the Fungicide Benomyl on Posterior Segmental Regeneration in the Earthworm, *Eisenia fetida*", *Pestic. Sci., 17*:641–652 (1986).

Biological Test Methods and Supporting Guidance Documents Published by Environment Canada's Method Development and Applications Section*

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods			
Acute Lethality Test Using Rainbow Trout	EPS 1/RM/9	July 1990	May 1996
Acute Lethality Test Using Threespine Stickleback (*Gasterosteus aculeatus)*	EPS 1/RM/10	July 1990	March 2000
Acute Lethality Test Using *Daphnia* spp.	EPS 1/RM/11	July 1990	May 1996
Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*	EPS 1/RM/21	February 1992	November 1997
Test of Larval Growth and Survival Using Fathead Minnows	EPS 1/RM/22	February 1992	November 1997
Toxicity Test Using Luminescent Bacteria (*Photobacterium phosphoreum*)	EPS 1/RM/24	November 1992	—
Growth Inhibition Test Using the Freshwater Alga *Selenastrum capricornutum*	EPS 1/RM/25	November 1992	November 1997
Acute Test for Sediment Toxicity Using Marine or Estuarine Amphipods	EPS 1/RM/26	December 1992	October 1998
Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars)	EPS 1/RM/27	December 1992	November 1997
Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout, Coho Salmon, or Atlantic Salmon)	EPS 1/RM/28 1st Edition	December 1992	January 1995
Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout)	EPS 1/RM/28 2nd Edition	July 1998	—
Test for Survival and Growth in Sediment Using the Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*)	EPS 1/RM/32	December 1997	—

* These documents are available for purchase from Environmental Protection Publications, Environment Canada, Ottawa, Ontario, K1A 0H3, Canada. For further information or comments, contact the Chief, Biological Methods Division, Environmental Technology Centre, Environment Canada, Ottawa, Ontario K1A 0H3.

Title of Biological Test Method or Guidance Document	**Report Number**	**Publication Date**	**Applicable Amendments**
A. Generic (Universal) Biological Test Methods (cont'd.)			
Test for Survival and Growth in Sediment Using the Freshwater Amphipod *Hyalella azteca*	EPS 1/RM/33	December 1997	—
Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*	EPS 1/RM/37	March 1999	—
Test for Survival and Growth in Sediment Using Spionid Polychaete Worms (*Polydora cornuta*)	EPS 1/RM/41	December 2001	—
Tests for Toxicity of Contaminated Soil to Earthworms (*Eisenia andrei, Eisenia fetida*, or *Lumbricus terrestris*)	EPS 1/RM/43	June 2004	—
Tests for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil	EPS 1/RM/45	September 2004?	—
Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil	EPS 1/RM/47	2005	—
B. Reference Methods**			
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 1st Edition	July 1990	May 1996, December 2000
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 2nd Edition	December 2000	—
Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*	EPS 1/RM/14 1st Edition	July 1990	May 1996, December 2000
Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*	EPS 1/RM/14 2nd Edition	December 2000	—
Reference Method for Determining Acute Lethality of Sediment to Marine or Estuarine Amphipods	EPS 1/RM/35	December 1998	—
Reference Method for Determining the Toxicity of Sediment Using Luminescent Bacteria in a Solid-Phase Test	EPS 1/RM/42	April 2002	—

** For this series of documents, a *reference method* is defined as a specific biological test method for performing a toxicity test, i.e., a toxicity test method with an explicit set of test instructions and conditions which are described precisely in a written document. Unlike other generic (multi-purpose or "universal") biological test methods published by Environment Canada, the use of a *reference method* is frequently restricted to testing requirements associated with specific regulations.

Title of Biological Test Method or Guidance Document	**Report Number**	**Publication Date**	**Applicable Amendments**
C. Supporting Guidance Documents			
Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants	EPS 1/RM/12	August 1990	—
Guidance Document on Collection and Preparation of Sediment for Physicochemical Characterization and Biological Testing	EPS 1/RM/29	December 1994	—
Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant	EPS 1/RM/30	September 1995	—
Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology	EPS 1/RM/34	December 1999	—
Guidance Document for Testing the Pathogenicity and Toxicity of New Microbial Substances to Aquatic and Terrestrial Organisms	EPS 1/RM/44	March 2004	—
Guidance Document on Statistical Methods to Determine Endpoints of Toxicity Tests	EPS 1/RM/46	October 2004?	—

Members of the Inter-Governmental Environmental Toxicity Group (as of December 2003)

Federal, Environment Canada

C. Blaise
Centre St. Laurent
Montreal, Quebec

M. Bombardier
Environmental Technology Centre
Ottawa, Ontario

U. Borgmann
National Water Research Institute
Burlington, Ontario

J. Bruno
Pacific Environmental Science Centre
North Vancouver, British Columbia

C. Buday
Pacific Environmental Science Centre
North Vancouver, British Columbia

K. Doe
Atlantic Environmental Science Centre
Moncton, New Brunswick

G. Elliott
Environmental Protection Service
Edmonton, Alberta

F. Gagné
Centre St. Laurent
Montreal, Quebec

M. Harwood
Environmental Protection Service
Montreal, Quebec

D. Hughes
Atlantic Environmental Science Centre
Moncton, New Brunswick

P. Jackman
Atlantic Environmental Science Centre
Moncton, New Brunswick

N. Kruper
Environmental Protection Service
Edmonton, Alberta

M. Linssen
Pacific Environmental Science Centre
North Vancouver, British Columbia

D. MacGregor
Environmental Technology Centre
Ottawa, Ontario

L. Porebski
Marine Environment Branch
Gatineau, Quebec

J. Princz
Environmental Technology Centre
Ottawa, Ontario

G. Schroeder
Pacific Environmental Science Centre
North Vancouver, British Columbia

R. Scroggins
Environmental Technology Centre
Ottawa, Ontario

T. Steeves
Atlantic Environmental Science Centre
Moncton, New Brunswick

Federal, Environment Canada (cont'd.)

D. Taillefer
Marine Environment Branch
Gatineau, Quebec

S. Trottier
Centre St. Laurent
Montreal, Quebec

G. van Aggelen (Chairperson)
Pacific Environmental Science Centre
North Vancouver, British Columbia

B. Walker
Centre St. Laurent
Montreal, Québec

P. Wells
Environmental Conservation Service
Dartmouth, Nova Scotia

Federal, Fisheries & Oceans Canada

R. Roy
Institut Maurice Lamontagne
Mont-Joli, Quebec

Federal, Natural Resources Canada

J. McGeer
Mineral Sciences Laboratory, CANMET
Ottawa, Ontario

B. Vigneault
Mineral Sciences Laboratory, CANMET
Ottawa, Ontario

J. Beyak
Mineral Sciences Laboratory, CANMET
Ottawa, Ontario

Provincial

C. Bastien
Ministère de l'Environnement du Québec
Ste. Foy, Quebec

B. Bayer
Manitoba Environment
Winnipeg, Manitoba

M. Mueller
Ontario Ministry of Environment
Rexdale, Ontario

D. Poirier
Ontario Ministry of Environment
Rexdale, Ontario

J. Schroeder
Ontario Ministry of Environment
Rexdale, Ontario

T. Watson-Leung
Ontario Ministry of Environment
Rexdale, Ontario

Environment Canada Regional and Headquarters Offices

Headquarters
351 St. Joseph Boulevard
Place Vincent Massey
Gatineau, Quebec
K1A 0H3

Ontario Region
4905 Dufferin St., 2nd Floor
Downsview, Ontario
M3H 5T4

Atlantic Region
15th Floor, Queen Square
45 Alderney Drive
Dartmouth, Nova Scotia
B2Y 2N6

Prairie and Northern Region
Room 210, Twin Atria No. 2
4999 - 98th Avenue
Edmonton, Alberta
T6B 2X3

Quebec Region
8th Floor, 105 McGill Street
Montreal, Quebec
H2Y 2E7

Pacific and Yukon Region*
401 Burrard Street
Vancouver, British Columbia
V6C 3S5

* A computer program for calculating LC50 is available from the Environmental Toxicology Section, Pacific Environmental Science Centre, 2645 Dollarton Highway, North Vancouver, BC, V7H 1B1, by providing a computer diskette.

Appendix D

Members of the Scientific Advisory Group

SAG Members

Mr. Christian Bastien
Centre d'expertise en analyse
environnementale du Québec
Ministère de l'Environnement
2700 Einstein
Saint-Foy, Quebec G1P 3W8
Phone: (418) 643-8225
Fax: (418) 643-9023
e-mail: christian.bastien@menv.gouv.qc.ca

Dr. Clive Edwards
Ohio State University
Department of Entomology
1735 Neil Avenue
Columbus, Ohio
USA 43210
Phone: (614) 292-3786
Fax: (614) 688-4222
e-mail: soilecol@osu.edu

Dr. Roman G. Kuperman
U.S. Army Edgewood Chemical Biological Center
AMSSB-RRT-TE E5641 DR KUPERMAN
5183 Blackhawk Road
Aberdeen Proving Ground, Maryland
USA 21010-5424
Phone: (410) 436-4697
Fax: (410) 436-4846
Email: roman.kuperman@us.army.mil

Dr. Roman P. Lanno
Ohio State University
Department of Entomology
1735 Neil Avenue
Columbus, Ohio
USA 43210
Phone: (614) 292-4943
Fax: (614) 292-2180
e-mail: lanno.1@osu.edu

Dr. Frank Riepert
Biologische Bundesantalt fur Land- und
Forstwirtschaft (BBA)
Konigin-Luise-Str. 19
D-14195
Berlin, Germany
Phone: 0049 30 8304 2406
Fax: 0049 30 8304 2403
E-mail F.Riepert@bba.de

Dr. Jöerg Römbke
ECT Oekotoxikologie GmbH
Boettgerstrasse 2-14
65439 Flörsheim am Main
Germany
Phone: 49 6145 95640
Fax: 49 6145 95649 9
e-mail: J-Roembke@ect.de

Dr. Geoffrey Sunahara
National Research Council
Biotechnology Research Institute
6100 Royalmount Avenue
Montreal, Quebec H4P 2R2
Phone: (514) 496-8030
Fax: (514) 496-6265
e-mail: geoffrey.sunahara@nrc.ca

Mr. Graham van Aggelen
Environment Canada
Pacific Environmental Science Centre
2645 Dollarton Highway
North Vancouver, BC V7H 1B1
Phone: (604) 924-2513
Fax: (604) 924-2555
e-mail: graham.vanaggelen@ec.gc.ca

Dr. Kees van Gestel
Institute of Ecological Science
Vrije Universiteit
De Boelelaan 1087
1081 HV Amsterdam
The Netherlands
Phone: 31 20 444-7079/7004
Fax: 31 20 444-7123
e-mail: gestel@bio.vu.nl

Dr. Suzanne Visser
Department of Biological Sciences
University of Calgary
2500 University Drive NW
Calgary, Alberta T2N 1N4
Phone: (403) 220-6375
Fax: (403) 289-9311
e-mail: svisser@acs.ucalgary.ca

Scientific Authority

Mr. Rick Scroggins
Environment Canada
Biological Methods Division
Environmental Technology Centre
335 River Road
Ottawa, ON K1A 0H3
Phone: (613) 990-8569
Fax: (613) 990-0173
e-mail: rick.scroggins@ec.gc.ca

Consultants

Dr. Don McLeay
McLeay Environmental Ltd.
2999 Spring Bay Road
Victoria, BC V8N 5S4
Phone: (250) 472-2608
Fax: (250) 472-2609
e-mail: mcleayenvir@islandnet.com

Ms. Gladys Stephenson
Aquaterra Environmental Ltd.
RR1, Site 5936
Orton, ON L0N 1N0
Phone: (519) 836-6050
Fax: (519) 836-2493
e-mail: gstephenson@stantec.com

Procedural Variations for Culturing *Eisenia andrei/fetida*, as Described in International Guides and Test Methods for Measuring Soil Toxicity Using These Species of Earthworm

The following source documents are listed chronologically, by originating agency rather than by author(s).

OECD 1984—the standard guideline for testing the toxicity of chemicals to earthworms (*E. fetida*) during a 14-day exposure published by the Organization for Economic Cooperation and Development (Paris, France) in 1984.

USEPA 1989—the protocol for performing 14-day lethality tests with the lumbricid earthworm *Eisenia fetida*, published in February 1989 by the United States Environmental Protection Agency (co-authors, J.C. Greene, C.L. Bartels, W.J. Warren-Hicks, B.R. Parkhurst, G.L. Linder, S.A. Peterson, and W.E. Miller) as one of several protocols for short-term toxicity screening of hazardous waste sites.

ISO 1993—an international standard test method for measuring soil toxicity using a 14-day lethality test with *Eisenia fetida*, published in 1993 by the International Organization for Standardization in Geneva, Switzerland.

ISO 1998—an international standard test method for measuring soil toxicity using a test for effects on reproduction of *Eisenia fetida*, published in 1998 by the International Organization for Standardization in Geneva, Switzerland.

ASTM 1999b—the standard guide for conducting soil toxicity tests with the lumbricid earthworm *Eisenia fetida*, written for the American Society for Testing and Materials (ASTM) under the jurisdiction of ASTM Subcommittee E47.03 on sediment toxicology and published in February 1998. Procedures described in this test-method document, which are specific to performing 14-day tests of survival of *E. fetida* in multiple concentrations of contaminated soil, are summarized here.

EC 2000a—the standard operating procedure for culturing *Eisenia andrei*, prepared in December 2000 by D. Moul of Environment Canada's Pacific Environmental Science Centre, Vancouver, BC.

OECD 2000—a draft proposal for a standard guideline for assessing the effects of chemicals on the reproduction of the earthworm *Eisenia fetida*, under consideration for publication by the Organization for Economic Cooperation and Development (Paris, France).

1 Source of Brood Stock for Culture

Document[1]	Initial Source
OECD 1984	to obtain worms of standard age and weight, it is best to start the culture with cocoons; cocoons can be purchased commercially or distributed from a central source to ensure that the same strain is used
USEPA 1989	starter cultures can be obtained from Vittor & Associates, 8100 Cottage Hill Road, Mobile, AL, 36695.
ISO 1993	to obtain worms of standard age and mass, it is best to start the culture with cocoons
ISO 1998	to obtain worms of standard age and mass, it is best to start the culture with cocoons
ASTM 1999b	starter cultures might be obtained from various institutions, laboratories, and biological firms; *E. fetida* can be found in manure piles; bait farms might contain mixtures of *E. fetida fetida* and *E. fetida andrei*; it is important to ensure a pure culture; field-collected *E. fetida* should be identified (to subspecies) using adult worms[2]
EC 2000a	sexually mature worms are purchased from reputable worm farmers
OECD 2000	to obtain worms of standard age and size (mass), it is best to start the culture with cocoons

[1] See preceding page.
[2] The taxonomic key of Fender (1985) is useful for this purpose.

2 Culture Chambers and Capacity

Document	Chamber Type and Size	Number of Units and Capacity for Substrate or Worm Production
OECD 1984	wooden breeding boxes measuring ~50 × 50 × 15 cm, with tightly fitting lids are ideal for large-scale breeding	using this apparatus, >1000 worms can be produced in six weeks[2]
USEPA 1989	glass, polyethylene, or wooden containers of suitable size for handling and moving; covered with glass lid containing air holes	maintain ≥4 culture containers, to ensure a sufficient number of earthworms on a continuing basis
ISO 1993	breeding boxes (wooden or any other shallow container) of 10–50-L capacity are suitable	NI[3]
ISO 1998	breeding boxes of 10–50-L capacity are suitable	NI
ASTM 1999b	plastic trays measuring ~34 × 28 × 14 cm, covered (e.g., with plastic)	700 g (dry wt) peat moss, hydrated with ~2300 mL of reagent water
EC 2000a	plastic worm bins measuring 53 × 38 × 30 cm, with an opaque lid and holes in the lid and/or near the top of the sides[1]	1000 adult worms can be added to the worm bin to start a new culture
OECD 2000	breeding boxes of 10–50-L capacity are suitable	NI

[1] One supplier provides a double-walled bin system, with drainage holes in the bottom of the inner bin to allow excess bedding moisture to collect in the outer bin.
[2] A breeding chamber of this size will support up to 1 kg of worms which each weigh up to 1 g, in 20 kg waste.
[3] NI = not indicated.

3 Temperature and Lighting During Culturing

Document	Temperature (°C)	Lighting Conditions
OECD 1984	20	NI[1]
USEPA 1989	22 ± 2	continuous darkness
ISO 1993	20 ± 2	NI
ISO 1998	20 ± 2	NI
ASTM 1999b	22 ± 3	continuous lighting
EC 2000a	20 ± 4	continuous lighting
OECD 2000	20 ± 2	NI

[1] NI = not indicated.

4 Culturing Substrate

Document	Culturing Substrate	pH	Renewal Conditions
OECD 1984	1:1 mixture of horse or cattle manure and peat[1]	should be ~7.0	NI[9]
USEPA 1989	peat moss bedding[2,3]	adjust to pH 5–8 using $CaCO_3$ [7]	every 2–3 months[10]
ISO 1993	1:1 mixture of horse or cattle manure and peat[1,3]	should be ~5–7 [8]	NI
ISO 1998	1:1 mixture of horse or cattle manure and peat[1,3]	should be ~6–7 [8]	NI
ASTM 1999b	sphagnum peat moss bedding, hydrated[4]	should be ~7.0 [8]	renew bedding periodically[11]
EC 2000a	mixture of artificial soil[5], shredded newspaper, and peat moss[2,6]	should be 7–8 [8]	renew bedding every 2–3 months[10, 11]
OECD 2000	1:1 mixture of horse or cattle manure and peat[1,3]	should be ~6–7 [8]	transfer periodically to fresh substrate

1 Other animal wastes are also suitable. The medium should have a low conductivity (<6.0 mS) and not be contaminated excessively with ammonia or animal waste.

2 Care must be taken to avoid overwatering. Avoid standing water in the bottom of the culture container. Add water when the bedding appears dry on the surface.

3 The substrate should be moist but not too wet.

4 Hydrate with distilled or de-ionized water. Moisture should be monitored weekly. There should be no standing water in the bottom of the tray, and the surface of the bedding should not be dry.

5 See Table 3 in Appendix E.

6 Initially, the mixture is preconditioned in a moistened state 2–4 weeks before worms are introduced.

7 pH adjustment is achieved by adding up to 3% by weight of $CaCO_3$ to the peat moss bedding.

8 $CaCO_3$ is added until this pH is achieved.

9 NI = not indicated.

10 Peat tends to become waterlogged with time and anaerobic conditions develop, as indicated by the development of a strong odour and by a change in colour between the bottom of the bedding material and the upper 2.5–5 cm.

11 For renewal, prepare a new tray of bedding, and place the contents of the old tray of bedding on top of the new bedding. Allow this tray to sit uncovered in the continuously lighted culture chamber for two days, then remove and discard the old bedding. The procedure does not remove the cocoons, and some worms will still be in the old bedding.

5 Feeding During Culturing

Document	Description of Food Used	Quantity and Feeding Procedure	Feeding Frequency
OECD 1984	animal waste[1]	50% of substrate[1]	continuous[1]
USEPA 1989	alfalfa pellets[2], saturated with water and aged for two weeks before use	sprinkle on surface of culture tray	NI[6]
ISO 1993	animal waste[1]	50% of substrate[1]	continuous[1]
ISO 1998	animal waste[1]	50% of substrate[1]	continuous[1]
ASTM 1999b	alfalfa pellets[2], saturated with water[3] and aged for two weeks before use	sprinkle on surface of culture tray[5]	once or twice per week[7]
EC 2000a	alfalfa cubes[2,4], saturated with water	sprinkle on surface of culture bedding	once per week[8]
OECD 2000	animal waste[1]	50% of substrate[1]	continuous[1]

1 See Table 4, this appendix.
2 Dried alfalfa (as pellets or cubes) can be obtained from agricultural feed and supply stores.
3 Environment Canada's laboratory at the Pacific Environmental Science Centre now uses alfalfa pellets (Moul, 2001).
4 A ratio of ~1 g of dry pellets per 2 mL test water is recommended.
5 Food is sprinkled over the surface of the bedding in an amount that has been determined will be consumed by the next feeding.
6 NI = not indicated.
7 Feeding frequency depends on the number of individuals in a tray. Remove remaining food and discard, at feeding times.
8 Prior to the addition of new food, the remains of the previous week's food is removed.

6 Culture Maintenance and Developmental Rate

Document	Culture Maintenance	Developmental Rate
OECD 1984	NI[1]	worms take 3–4 weeks to hatch and 7–8 weeks to mature
USEPA 1989	crowding must be avoided[2]; splitting the bedding material in half every 3–4 months, with pH-adjusted peat moss, prevents overcrowding	worms reach maturity in 7–8 weeks[7]
ISO 1993	place adults worms in a breeding box with fresh substrate, and remove them after 14–21 days[3]; earthworms hatched from cocoons are used for testing when mature	worms become mature after 2–3 months
ISO 1998	place adults worms in a breeding box with fresh substrate, and remove them after 14–28 days[3]; earthworms hatched from cocoons are used for testing when mature	worms become mature after 2–3 months
ASTM 1999b	at feeding times, turn the bed by hand to inspect the general condition of the worms and the bedding; remove dead worms; loading density[4,5] should be $\leq$0.03 g/cm^3	at 25 °C, a life cycle of only ~52 days has been reported; *E. fetida* has a typical life expectancy of 1–2 years
EC 2000a	once a week, the bin bedding is inspected to examine the worms' health and to watch for overcrowding[6]	the life cycle of *E. andrei* varies from 50–166 days (temperature dependent)
OECD 2000	place adults worms in a breeding box with fresh substrate, and remove them after 14–21 days[3]; earthworms hatched from cocoons are used for testing when mature	worms become mature after $\geq$2 months

1 NI = not indicated.
2 Crowding decreases the growth rate and reproduction efficiency.
3 These adult worms may be used for further breeding batches.
4 For optimal reproduction, trays containing 9000 cm^3 of bedding should hold a maximum of 245 g of worm (e.g., 350 adult worms weighing 700 mg each), i.e., the loading density should be $\leq$0.03 g/cm^3.
5 To reduce the population of worms in a crowded tray, prepare a tray of new bedding and place half of this new bedding on a piece of plastic sheeting; then place half of the bedding from the overcrowded tray into the new tray, and mix the bedding by hand; mix the remaining new bedding on the plastic sheet with the old tray of bedding.
6 If the worm bins show signs of crowding, their contents should be divided. To achieve this, prepare a bin of new bedding and place half of its contents on a plastic sheet. Half of the old bedding, containing a portion of the worms, is placed into the new bin and the bedding is mixed by hand. The half of new bedding on the plastic sheet is then added to the old bin and mixed.
7 *E. fetida* is very prolific; a single worm produces 2–5 cocoons per week, each with several worms.

7 Indices of Culture Health and Acceptability—Age of Worms Used in Toxicity Tests

Document	Indices of Culture Health and Acceptability	Age of Worms Used in Toxicity Tests
OECD 1984	NI[1]	≥2 months old, with clitellum
USEPA 1989	NI	≥2 months old with clitellum
ISO 1993	NI	sexually mature adults, ≥2 months old
ISO 1998	NI	sexually mature adults, 2–12 months old
ASTM 1999b	culture trays discarded or set aside for further observation if many dead or apparently stressed worms observed during observation periods[2]	NI (use sexually mature, fully clitellate adults)
EC 2000a	if any dead worms are noted during the weekly inspection of the culture bin(s), they are removed and the bin is set aside for more frequent examinations; records of culture health are gauged by a sensitivity to a reference toxicant	NI (adult worm must weigh >300 mg)
OECD 2000	worms are considered healthy if they move through the substrate, do not try to leave it, and reproduce continuously; substrate exhaustion is indicated by worms moving very slowly and having a yellow posterior end[3]	sexually mature adults, 2–12 months old

[1] NI = not indicated.
[2] See Table 6, this appendix.
[3] In this case, the provision of fresh substrate and/or a reduction in stocking density is recommended.

Procedural Variations for 14-Day Lethality Tests of Soil Toxicity Using Earthworms (*Eisenia andrei/fetida*), as Described in International Methodology Documents

The following source documents are listed chronologically, by originating agency rather than by author(s).

OECD 1984—the standard guideline for testing the toxicity of chemicals to earthworms (*E. fetida*) during a 14-day exposure, published by the Organization for Economic Cooperation and Development (Paris, France) in 1984.

USEPA 1989—the protocol for performing 14-day lethality tests with the lumbricid earthworm *Eisenia fetida*, published in February 1989 by the United States Environmental Protection Agency (co-authors, J.C. Greene, C.L. Bartels, W.J. Warren-Hicks, B.R. Parkhurst, G.L. Linder, S.A. Peterson, and W.E. Miller) as one of several protocols for short-term toxicity screening of hazardous waste sites.

ISO 1993—an international standard test method for measuring soil toxicity using a 14-day lethality test with *Eisenia fetida*, published in 1993 by the International Organization for Standarization in Geneva, Switzerland.

ASTM 1999b—the standard guide for conducting soil toxicity tests with the lumbricid earthworm *Eisenia fetida*, written for the American Society for Testing and Materials (ASTM) under the jurisdiction of ASTM Subcommittee E47.03 on sediment toxicology and published in February 1998. Procedures described in this test-method document, which are specific to performing 14-day tests of survival of *E. fetida* in multiple concentrations of contaminated soil, are summarized here.

EC 2000b—the standard operating procedure for performing 14-day lethality tests for soil toxicity using the earthworm *Eisenia andrei*, prepared in December 2000 by D. Moul for Environment Canada's Pacific Environmental Science Centre, North Vancouver, BC.

1 Test Type and Duration—Specifics on Test Organisms at Start

Document	Species	Test Type	Test Duration (days)	Description of Organisms at Start of Test	Wet Weight of Worms at Start (mg)
OECD 1984	*E. fetid*a	static	14[1]	laboratory-cultured adults, ≥ 2 months with clitellum	300–600
USEPA 1989	*E. fetid*a	static	14	laboratory-cultured adults, ≥2 months with clitellum[3]	300–500
ISO 1993	*E. fetid*a	static	14[1]	cultured adults, ≥2 months with clitellum	300–600
ASTM 1999b	*E. fetida*	static	14[2]	laboratory-cultured, sexually mature adults with clitellum[4,5]	NI[5]
EC 2000b	*E. andrei*	static	14[1]	laboratory-cultured, sexually mature adults with clitellum[4]	>300

[1] Mortality at seven and 14 days is determined. After the seven-day assessment, all surviving worms and the test material are returned to the test chamber, and the test is continued for a further seven days.
[2] An evaluation at seven days is optional.
[3] It is preferable to obtain the earthworms from an in-house culture unit; animals should be from the same culture container.
[4] The biomass of earthworms in each test chamber should be obtained.
[5] The laboratory should use separate constant temperature areas (chambers) for culturing and testing, to reduce the possibility of contamination by test materials and other substances (especially volatile compounds).
[5] NI = Not indicated.

2 Test Chambers and Materials

Document	Test Chamber	Cover	Type of Test Soil[1,2]	Amount of Soil/Container (g)
OECD 1984	~1-L glass beaker or dish	glass lid or perforated plastic film	AS	750, wet wt
USEPA 1989	1-pint glass canning jar	ring and screw-top lid with ~0.3-cm hole	AS, SWM, and mixtures thereof	200, dry wt
ISO 1993	glass container, 1–2 L capacity	NI[3] ("not tightly closed")	AS	500, dry wt
ASTM 1999b	473-mL glass canning jar	ring and screw-top lid with 1-2 mm hole	AS, SS, RS, SAS, SSS, SRS, or mixtures thereof	200, dry wt
EC 2000b	500-mL wide-mouth glass jar	metal lid with ~4 mm hole in middle	AS, SS, SL, RS, SAS, SSS, SRS, or mixtures thereof	200, dry wt

[1] See Table 3, this appendix for a description.
[2] AS = artificial soil; SWM = solid waste material; SS = site soil; SL = sludge (industrial or domestic); RS = reference soil; SAS = spiked artificial soil; SSS = spiked site soil; SRS = spiked reference soil.
[3] NI = not indicated.

3 Description of Test Soils, Including Composition of Artificial Soil

Document	Description of Test Soil(s)	Description of Artificial Soil[1]
OECD 1984	artificial soil with added test substance (e.g., chemical in de-ionized water or mixed with fine quartz sand)	10% sphagnum peat, 20% kaolin clay with >30% kaolinite, and 70% quartz sand with >50% particles 50–200 µm
USEPA 1989	solid hazardous waste (contaminated soil) or aqueous chemical substances mixed in artificial soil	10% sieved (2.36 mm) sphagnum peat, 20% kaolinite clay, and 70% "grade 70" silica sand
ISO 1993	artificial soil with added test substance (e.g., chemical in de-ionized water or organic solvent; if insoluble, test substance mixed in fine quartz sand)	10% sphagnum peat, 20% kaolin clay with ≥30% kaolinite, 70% quartz sand with >50% particles 50–200 µm, adjust to pH 6.0 ± 0.5 using $CaCO_3$
ASTM 1999b	reference or potentially toxic site soil; soil spiked with compounds; or soil diluted with reference or artificial soil	10% sieved (2.36 mm) sphagnum peat, 20% kaolin clay with particles <40 µm, and 70% "grade 70" silica sand
EC 2000b	reference or potentially toxic site soil; domestic or industrial sludge; soil spiked with chemicals; or or soil diluted with artificial soil	10% sieved (2.36 mm) sphagnum peat, 20% kaolin(ite) clay, and 70% "grade 70" silica sand; adjust to pH 7.0 by adding $CaCO_3$

[1] Percentages are expressed on a dry-mass basis.

4 Description of Negative Control Soil and Reference Soil

Document	Description of Control Soil	Description of Reference Soil
OECD 1984	artificial soil[1], treated with the same solvent as used for the test soil	NA[2]
USEPA 1989	100% artificial soil[1]	NI[3]
ISO 1993	artificial soil[1]; use an additional control (solvent control) if solvent other than water used to dissolve test substance	NA[2]
ASTM 1999b	artificial soil[1] or "clean" reference soil from the field with characteristics similar to test soil(s); use an additional control (solvent control) if solvent other than water used to dissolve substance	test may also include a reference soil if the negative control is an artificial soil
EC 2000b	artificial soil[1]	field-collected soil from an area that has not been cultivated or treated with pesticides or fertilizers in past 25 years, with geochemical characteristics similar to test soil(s)[4]

[1] See Table 3, this appendix.
[2] NA = not applicable.
[3] NI = not indicated.
[4] The reference soil should be air dried to 10–20% moisture content, sieved (4–9 mm), and stored at ≤8 °C.

5 Storage and Characterization of Test Soil

Document	Storage Conditions	Soil Characterization
OECD 1984	NI[1], NA[2]	moisture content after hydrating[3] (dry subsample at 105 °C and reweigh)
USEPA 1989	seal in plastic; chill to 4 °C; ship on ice; hold 4 °C; initiate test within 24 h of collection	pH at start and end of test; TOC, each test concentration including control
ISO 1993	NI[1], NA[2]	moisture content after hydrating[3]
ASTM 1999b	seal in plastic; store at 4 ± 2 °C for ≤2 weeks; do not freeze	at least pH, TOC, CEC, N, SSC, and W[4]; contaminants of concern
EC 2000b	store in dark at 4 ± 2 °C, for ≤2 weeks	moisture content and pH

[1] NI = not indicated.
[2] NA = not applicable (artificial soil prepared, hydrated, and then used).
[3] See Tables 3 and 6, this appendix.
[4] pH = hydrogen ion concentration; TOC = total organic carbon; CEC = cation exchange capacity; N = total nitrogen; SSC = % sand, % silt, and % clay; W = % water (moisture content). Moisture content is determined by drying a subsample for 24 h at 100 °C and reweighing.

6 Manipulation of Soil Before Use in Test

Document	Mixing	Hydration	pH Adjustment
OECD 1984	blend dry constituents in correct proportions and mix thoroughly; mix again after hydrating	determine moisture content; add de-ionized water until moisture ~35% of dry wt[1]	adjust to 6.0 ± 0.5, using $CaCO_3$
USEPA 1989	homogenize test material; mix with artificial soil in blender; hydrate	hydrate to 75% of water-holding capacity, using de-ionized water	maybe, if pH <4 or >10[2]
ISO 1993	blend dry constituents in correct proportions and mix thoroughly; mix again after hydrating	hydrate to 40–60% of total water-holding capacity, using de-ionized or distilled water	addition of $CaCO_3$ adjusts pH to 6.0 ± 0.5
ASTM 1999b	screen (6.30-mm mesh); mix; determine moisture content; hydrate	hydrate to 35–45% of dry weight, for each test soil	tests might include pH-adjusted soil
EC 2000b	screen if required, in which case dry to 10–20% moisture if necessary; mix; hydrate	hydrate to ~35% of dry weight, for each test soil	NI[3] (testing range, pH >4 to <10

[1] With some peats, a moisture content of >35% might be suitable.
[2] If pH outside the range of 4–10, results might reflect pH toxicity. Altering the pH of the soil can increase or decrease (depending on contaminant) the toxicity of contaminants therein.
[3] Not indicated.

7 Preliminary Test — Number of Organisms per Chamber, Number of Replicates per Treatment, and, for a Multi-concentration Test, Number of Concentrations per Sample

Document	Number of Worms per Chamber	Number of Replicates per Treatment or Concentration	Number of Concentrations per Sample or Test Material	Recommended Dilution Factor
OECD 1984[1]	NI[2]	NI	~6 + control	geometric series[3]
USEPA 1989	NI	NI	NI	NI
ISO 1993	10	1	5 + control	geometric series[4]
ASTM 1999b[1]	NI	NI	≥3 + control	geometric series[5]
EC 2000b	NI	NI	NI	NI

[1] A preliminary (range-finding) test is optional.
[2] Not indicated.
[3] The concentrations could be spaced by a factor of 10 (e.g., 0.01, 0.1, 1.0, 10, 100, and 1000 mg/kg, dry-weight basis)
[4] The concentrations could be spaced by a factor of 10 (e.g., 0.1, 1.0, 10, 100, and 1000 mg/kg, dry-weight basis)
[5] The concentrations should differ by a factor of 10.

8 Definitive Test — Number of Organisms per Chamber, Number of Replicates per Treatment, and, for a Multi-concentration Test, Number of Concentrations per Sample

Document	Number of Worms per Chamber	Number of Replicates per Treatment or Concentration	Number of Concentrations per Sample or Test Material	Recommended Dilution Factor
OECD 1984	10[1]	4	5 + control	geometric series
USEPA 1989	10	3	≥5 + control	0.5 (e.g., 100%, 50%, 25%, 12.5%, 6.25%, 3.13%, 1.56%, 0.0%)
ISO 1993	10	4	5 + control	geometric series
ASTM 1999b	10	≥3	≥5 + control	NI[2]
EC 2000b	10	3	≥5 + control	NI

[1] Worms are conditioned for 24 h in an artificial soil and then washed quickly before use
[2] Not indicated.

9 Temperature, Lighting, and Feeding During Test

Document	Temperature (° C)	Lighting Conditions	Feeding?
OECD 1984	20 ± 2	continuous; intensity, 400–800 lux	NI[3]
USEPA 1989	20 ± 2[1]	continuous; intensity, ambient lighting (~540–1080 lux)	do not feed
ISO 1993	20 ± 2	intensity, 400–800 lux on containers; photoperiod between 12 L:12 D and 16 L:8 D	NI[3]
ASTM 1999b	22 ± 3[2]	continuous, using fluorescent or incandescent; 400–1080 lux	do not feed
EC 2000b	20 ± 4	continuous, full-spectrum	NI[3]

[1] Soil temperature. Continuously monitor the temperature in the environmental control chamber.
[2] Monitor for the duration of the test. The use of a continuous temperature recorder is recommended.
[3] NI = not indicated. It is assumed that the intent was to not feed during the test.

10 Measurements and Biological Observations During Test

Document	Measurements[1]	Biological Observations
OECD 1984	W of artificial soil at start and end, pH at start	mean wet wt/treatment, at start and end; number alive and dead in each replicate, at 7 and 14 days; obvious pathological symptoms or behavioural changes
USEPA 1989	pH at start and end; TOC, each concentration; temperature in test facility, continuously	number alive and dead in each replicate, at 7 and 14 days
ISO 1993	W and pH of artificial soil at start; W and pH of each replicate at end	mean wet wt/replicate, at start and end; number alive in each replicate, at 7 and 14 days; obvious pathological symptoms or distinct behavioural changes
ASTM 1999b	pH, W, and S each treatment, at start and end; temperature in test facility, continuously	number alive and dead in each replicate at 14 days (and, optionally, at 7 days); optionally, mean wet wt/replicate at start and end[2]; optionally, for obvious pathological symptoms or distinct behavioural changes (e.g., burrowing or non-burrowing at 24-h intervals)
EC 2000b	pH, each treatment, at start and end; W, each treatment, at start and (optionally) at end; temperature in test facility	worms burrowed or not burrowed at 24 h; number dead in each replicate, at 1, 7, and 14 days; mean wet wt/replicate at start

[1] W = percent water (moisture content); pH = hydrogen-ion concentration; TOC = total organic carbon; S = salinity.
[2] If weight loss is used as an endpoint, worms should be purged of their gut contents before weighing, by placing them in petri dishes with wet filter paper for 24 h.

11 Terminating Test, Biological Endpoints, and Statistical Endpoints

Document	Terminating Test	Biological Endpoints	Statistical Endpoints
OECD 1984	empty contents onto glass tray or plate, sort worms, test for reaction to mechanical stimulus at anterior end	% mortality, each concentration, on Days 7 and 14	7-day LC50 and 14-day LC50[1] (mg/kg dry weight)
USEPA 1989	count the number of live and dead worms in each test chamber on Days 7 and 14	% mortality, each concentration, on Days 7 and 14	7-day LC50 and 14-day LC50[1]
ISO 1993	count the number of live worms in each test chamber on Day 7, and the number of live and dead worms on Day 14	% mortality, each concentration, on Days 7 and 14	14-day LC50[1] (mg/kg dry weight); NOEC, based on lethal or sublethal effects, data permitting
ASTM 1999b	count the number of live and dead worms in each test chamber on Day 14	% mortality, each treatment, on Day 14	14-day LC50[1]; NOEC and LOEC, data permitting
EC 2000b	count the number of dead worms in each test chamber on Days 1, 7, and 14	% mortality, each concentration, on Days 1, 7, and 14	14-day LC50[1]; LT50[2]

[1] Including the 95% confidence limits.
[2] The estimated time to 50% mortality, based on observations of percent mortality at 24 h, 7 days, and 14 days.

12 Requirements for Valid Test—Use of Reference Toxicity Test

Document	Requirements for Valid Test	Reference Toxicant(s)[1]	Procedures and Conditions for Reference Toxicity Test
OECD 1984	control mortality ≤10% at test end	CLOAC	determine LC50 occasionally
USEPA 1989	mean control mortality ≤10% at test end	SDS, NaPCP, or $CdCl_2$	at least monthly; determine 14-day LC50; plot results on control chart
ISO 1993	control mortality <10% and decrease in control wt ≤20%	CLOAC	mix in artificial soil[3]; LC50 should be between 20 and 80 mg/kg
ASTM 1999b	mean control mortality ≤10% at test end	NI[2]	NI
EC 2000b	mean control mortality ≤10% at test end; ≤10% of controls stressed	KCl	conduct every 30–45 days; determine LC50; plot results on warning chart

[1] CLOAC = chloroacetamide; SDS = sodium dodecylsulphate; NaPCP = sodium pentachlorophenate; $CdCl_2$ = cadmium chloride; KCl = potassium chloride.
[2] NI = not indicated.
[3] See Tables 3 and 6, this appendix.

Appendix G

Procedural Variations for Tests of Effects of Contaminated Soil on the Reproduction of Earthworms (*Eisenia andrei/fetida*), as Described in International Methodology Documents

The following source documents are listed chronologically, by originating agency rather than by author(s).

ISO 1991— is a proposal from the Netherlands (C.A.M. van Gestel) for an international standard test method for measuring soil toxicity using a test for effects on reproduction of *Eisenia fetida* or *E. andrei*. This was circulated for comments in April 1991 by the International Organization for Standardization in Geneva, Switzerland.

ISO 1998—is an international standard test method for measuring soil toxicity using a test for effects on reproduction of *Eisenia fetida fetida* or *E. fetida andrei* (now recognized as separate species), published in 1998 by the International Organization for Standardization in Geneva, Switzerland.

OECD 2000—is a draft proposal for a standard guideline for assessing the effects of chemicals on the reproduction of the earthworm *Eisenia fetida* or *E. fetida andrei* (now recognized as separate species), under consideration for publication by the Organization for Economic Cooperation and Development (Paris, France).

1 Test Species, Test Type, and Test Duration

Document	Species	Test Type	Test Duration
ISO 1991	*E. fetida* or *E. andrei*	static with exchange[1]	21 days + 5 weeks[1]
ISO 1998	*E. fetida fetida* or *E. fetida andrei*	static	4 weeks + 4 weeks[2]
OECD 2000	*E. fetida fetida* or *E. fetida andrei*	static	4 weeks + 4 weeks[2]

[1] Following a 21-day period of exposure of adult worms to chemical-spiked soil, cocoons produced by the adults are transferred to untreated cocoon incubation substrate and held in this substrate for five weeks.
[2] Following a four-week period of exposure of adult worms to chemical-spiked soil, adults are removed and the test with their progeny (cocoons and/or juveniles) is continued for an additional four weeks using the same test substrates.

2 Specifics on Test Organisms at Start

Document	Description of Organisms Used to Start Test	Acclimation Conditions	Wet Weight per Worm at Test Start (mg)
ISO 1991	cultured adults, ≥2 months, with clitellum	hold for one week in 1-L glass jars containing 500 g (dry wt) AS and 5 g dry cow dung[2]; 10 worms/jar	250–600 [4]
ISO 1998	cultured adults, 2–12 months old, with clitellum	hold for 1–7 days in artificial soil containing added food[3]	300–600 [5]
OECD 2000	cultured adults, 2–12 months old, with clitellum[1]	hold for 1–7 days in artificial soil containing added food[3]	300–600 mg[5]

[1] Worms should be selected from a synchronized culture with a relatively homogeneous age structure, and individuals in a test group should not differ in age by more than four weeks.
[2] The artificial substrate (AS; see Tables 4 and 5, this appendix) and the cow dung shall have a 50–55% moisture content. The cow dung is placed in a small hole in the substrate in the jar, and covered with substrate; each glass jar is covered with a glass lid.
[3] The quantity and type of food should be the same as that added to the artificial soil spiked with substance(s) during the test (see Tables 4, 5, and 7, this appendix).
[4] The difference in weight between batches of 10 worms/container shall be ≤1 g. The total mass of worms in each glass jar is weighed at the end of the acclimation period.
[5] The weight of each group of 10 worms to be added to a test chamber is determined.

3 Test Chambers and Materials

Document	Test Chamber	Cover	Type of Test Soil[3,4]	Amount of Soil/Container
ISO 1991	1-L glass jar if adult; 18–20 cm glass petri dish if progeny[1]	glass lid if jar; NI if petri dish[2]	AS, ASC	500 g/jar, dry wt; 300 g/petri dish, wet weight[5]
ISO 1998	1–2 L capacity, with cross-sectional area of ~200 square cm	perforated transparent lid	AS	500–600 g/test chamber, dry wt; depth, 5–6 cm
OECD 2000	1–2 L capacity, with cross-sectional area	perforated transparent	AS	500–600 g/test chamber, dry wt

[1] Diameter of glass jar should not exceed 15 cm.
[2] Jars should be loosely covered with glass lids; covers for petri dishes are not indicated (NI).
[3] See Table 4, this appendix for a description.
[4] AS = artificial soil; ASC = artificial substrate for cocoons.
[5] An amount of wet artificial substrate equivalent to 500 g dry mass is added to each jar. For petri dishes, 150g (wet mass) of a formulated artificial substrate for incubating cocoons (ASC) is spread as a thin layer in the petri dish; cocoons are spread over this layer, and covered with a second layer (150 g, wet mass) of ASC.

4 Description of Test Soils, Including Composition of Artificial Soil

Document	Description of Test Soil(s)	Composition of Artificial Soil[2]
ISO 1991	artificial soil with added test substance (e.g., chemical in de-ionized water or organic solvent; if insoluble, test substance mixed in fine quartz sand)	10% sphagnum peat, 20% kaolin clay with >50% kaolinite, and 70% industrial sand with >50% particles 50–200 µm, adjust to pH 5.5 ± 0.5 using $CaCO_3$ [3]
ISO 1998	artificial soil with added test substance (e.g., chemical in de-ionized water or organic solvent; if insoluble, test substance mixed in fine quartz sand)	10% sphagnum peat[4], 20% kaolin clay with ≥30% kaolinite, and 70% industrial sand with >50% particles 50–200 µm, adjust to pH 6.0 ± 0.5 using $CaCO_3$
OECD 2000	artificial soil with added test substance (e.g., chemical in de-ionized water or organic solvent; if insoluble, test substance mixed in fine quartz sand)[1]	10% sphagnum peat[4], 20% kaolin clay with ≥30% kaolinite, and 70% industrial sand with >50% particles 50–200 µm, adjust to pH 6.0 ± 0.5 using $CaCO_3$

[1] The test substance is applied either by mixing it into the artificial soil (this procedure is recommended for the testing of chemicals in general), or by adding it to the surface of the soil following the introduction of worms to test chambers containing artificial soil.
[2] Percentages are expressed on a dry mass basis.
[3] This AS formulation is for addition to the glass jars containing adults (see Table 3 this appendix). The ASC formulation (for incubating cocoons) is the same, except it uses finely ground(<0.5 mm) sphagnum peat rather than "not too fine" or sieved (1.0 mm) sphagnum peat, and includes 1% finely ground (<0.5 mm) dried cow dung.
[4] The sphagnum peat is finely ground, and with no visible plant remains.

5 Manipulation of Artificial Soil Before Use in Test

Document	Mixing	Hydration	pH Adjustment
ISO 1991	blend dry constituents in correct proportions and mix with some de-ionized water; determine moisture content; hydrate	hydrate to 30–35% of dry mass of artificial substrate[1]; final water content of test substrate, ~50–55%[2]	add $CaCO_3$ to bring pH to 5.5 ± 0.5
ISO 1998	blend dry constituents in correct proportions; mix with food before test; hydrate with de-ionized or distilled water	hydrate to 40–60% of total water-holding capacity	add $CaCO_3$ to bring pH of wetted substrate (including food) to 6.0 ± 0.5
OECD 2000	blend dry constituents in correct proportions; moisten 1–2 days before test to ~50% of final water content, using de-ionized water	hydrate to 40–60% of total water-holding capacity[3]	add $CaCO_3$ to bring pH of wetted substrate (including food) to 6.0 ± 0.5

[1] For the artificial substrate used to incubate cocoons ("ASC"; see Tables 3 and 4, this appendix), the water content shall be adjusted to ~35% of the dry weight of the ASC.

[2] The final water content (~50–55%) of the test substance is achieved either by adding the test substance in a sufficient amount of de-ionized water, or, if the test substance is added in a dry form, by adding de-ionized water.

[3] The final water content (40–60% of maximum water-holding capacity) is achieved by the addition of the test substance in solution and/or by adding distilled or de-ionized water.

6 Negative Control Soil

Document	Description of Negative Control Soil	Number of Control Chambers (Replicates)
ISO 1991	treat with the same solvent used in the test[1]	≥4
ISO 1998	prepare the same way as for test soils, but without the test substance; use additional controls (solvent controls) if solvent other than water used to dissolve test substance	≥4
OECD 2000	prepare in the same way as for test soils, but without but without the test substance[2]; apply organic solvents, quartz sand, or other vehicles to additional controls in amounts consistent with those used in treatments	8, for NOEC; 6, for ECx; 8, for both NOEC and ECx[3]

[1] A description of the control soil was not provided. Presumably, though, it would be comprised of artificial soil (AS) for the initial 21-day period of the test involving adult worms, and of artificial substrate for cocoons (ASC) for the remaining five-week period of the test involving cocoons and juveniles. See Tables 1, 4, and 5, this appendix.

[2] See Table 4 (including footnote 1), this appendix.

[3] Three experimental designs are proposed. Depending on the selected experimental design, the endpoints would differ as would the recommended number of test concentrations and the recommended number of replicates/treatment including the control treatment(s).

7 Feeding During Test

Document	Type of Food Recommended	Feeding Quantity, Procedure, and Frequency
ISO 1991	dried cow dung[1]	for 21-day test with adults, feed 5 g (dry mass) on Days 0, 7, and 14[4]; for subsequent five-week test with progeny, start with artificial substrate containing 1% finely ground cow dung[5]
ISO 1998	optional[2]	mix the artificial soil with the food before the test[6]; add 5 g dry, finely ground food per test container to soil surface on Day 1 and weekly thereafter during period of test with adults[7]; at start of final four-week test period with progeny, carefully mix 5 g dry, finely ground food by hand into the substrate within each test chamber
OECD 2000	optional[2,3]	add ~5 g dry, finely ground food per test container to soil surface on Day 1 and weekly thereafter during period of test with adults[7]; at start of final four-week test period with progeny, carefully mix 5 g dry, finely ground food by hand into the substrate within each test chamber[8]

[1] The cow dung is dried at 105 °C and finely ground (<0.5 mm).
[2] Any food source of a quality capable of maintaining the earthworm population may be used. Dried finely ground cow manure has been shown to be suitable. Each batch of food should have been previously tested to determine that earthworms will feed on it.
[3] "Self-collected" cow manure is recommended, since experience has shown that commercially available cow manure used as garden fertilizer might have adverse effects on the worms.
[4] Before adding adult worms to test chambers, place 5 g (dry mass) moistened, finely ground cow dung in a small hole in the middle of the AS within each test chamber. The same amount is added on Days 7 and 14.
[5] Finely ground cow dung is added to the artificial substrate for cocoons (see Tables 3 and 4, this appendix) before placing this substrate in petri dishes.
[6] When dried cow manure is used, mix 5 g dry manure per 500 g dry mass of soil.
[7] Moisten with potable water by adding ~5–6 mL water per container. If food consumption is low, reduce feeding to a minimum.
[8] No further feeding takes place during the remaining four weeks of the test.

8 Preliminary Test — Number of Organisms per Chamber, Number of Replicates per Treatment, Number of Concentrations per Sample, and Recommended Dilution Factor

Document	Number of Worms per Chamber	Number of Replicates per Treatment or Concentration	Number of Concentrations per Sample or Test Material	Recommended Dilution Factor
ISO 1991[1,2]	NI[5]	1	5 + control	geometric series[6]
ISO 1998[1,3]	10	1	4 + control	geometric series[7]
OECD 2000[1,4]	10	1	5 + control	geometric series[7]

[1] A preliminary (range-finding) test is optional.
[2] The upper limit of the concentrations to be used in a definitive test can also be based on the results of an acute lethality test.
[3] If a preliminary test is necessary to determine the range of concentrations for a definitive test, perform an acute lethality test in accordance with ISO 11268-1 (see preceding Appendix E of the present document).
[4] If a preliminary test is necessary to determine the range of concentrations for a definitive test, perform an acute lethality test of two-weeks' duration.
[5] NI = not indicated.
[6] The concentrations could be spaced by a factor of 10 (e.g., 0.1, 1.0, 10, 100, and 1000 mg/kg, dry mass).
[7] For example, 0, 1, 10, 100, and 1000 mg/kg, dry mass.

9 Definitive Test — Number of Organisms per Chamber, Number of Replicates per Treatment, Number of Concentrations per Sample, and Recommended Dilution Factor

Document	Number of Worms per Chamber	Number of Replicates per Treatment or Concentration	Number of Concentrations per Sample or Test Material	Recommended Dilution Factor
ISO 1991	10	≥4	≥5 + control	geometric series[3]
ISO 1998	10	≥4	NI[2]	geometric series[4]
OECD 2000[1]	10	4, for NOEC	≥5 + control	geometric series[5]
	10	≥2, for ECx	12 + control	geometric series[6]
	10	4, for NOEC and ECx	8 + control	geometric series[5]

[1] Three experimental designs are proposed. Depending on that chosen, the statistical endpoints would differ as would the recommended number of test concentrations and the recommended number of replicates/treatment (including the number of control chambers; see Table 6 this appendix).
[2] NI = not indicated.
[3] The concentrations should be spaced by a factor not exceeding 3.2, and preferably not exceeding 1.8.
[4] The concentrations should be spaced by a factor not exceeding 2.0.
[5] The concentrations should be spaced by a factor not exceeding 1.8.
[6] The spacing factor may vary, i.e., ≤1.8 in the expected effect range and >1.8 at the higher and lower concentrations.

10 Temperature and Lighting During Test

Document	Temperature (° C)	Lighting Conditions
ISO 1991	20 ± 2	continuous light
ISO 1998	20 ± 2	12 L:12 D or 16 L:8 D; intensity, 400–800 lux on test chambers
OECD 2000	20 ± 2	controlled light:dark cycle, preferably 16 L:8 D; intensity, 400–800 lux on test chambers

11 Measurements and Biological Observations During Test

Document	Measurements[1]	Biological Observations
ISO 1991	W and pH of artificial soil at start and end of acclimation period[2], and at start and end of test as a minimum[3]	number of live worms/chamber and their total wet wt, at start and end of 21-day exposure; number of cocoons/chamber on Day 21; number of hatched cocoons, number of non-hatched cocoons with live juveniles, number of infertile cocoons, and total number of juvenile worms/hatched cocoon, at test end
ISO 1998	W and pH, each treatment, at start and end of test; reweigh test containers periodically throughout test[4]	total number and mass of living adult worms per chamber, at start and end of four-week exposure; number of offspring per chamber hatched from cocoons, at test end; obvious pathological symptoms or distinct behavioural changes (e.g., reduced feeding activity) noted for worms in each test chamber
OECD 2000	W and pH, each treatment, at start and end of test; reweigh test containers periodically throughout test[4]	total number and wet wt of living adult worms per chamber, at start and end of 28-day exposure; number of juveniles produced per chamber, at test end; obvious pathological symptoms (e.g., open wounds) or distinct behavioural changes (e.g., lethargy) noted for worms in each test chamber

[1] W = % water (moisture content); pH = hydrogen-ion concentration.

[2] See Table 2, this appendix.

[3] The water content of both the artificial substrate for adults and the cocoon incubation substrate (see Tables 4 and 5, this appendix) shall be checked at regular intervals; loss of moisture shall be replenished by the addition of de-ionized water.

[4] Maintain the water content of the soil substrate in test containers by reweighing and, if necessary, replenishing lost water. At the end of the test, the water content shall not differ by more than 10% from that at the beginning of the test.

12 Terminating Test, Biological Endpoints, and Statistical Endpoints

Document	Terminating Test	Biological Endpoints	Statistical Endpoints
ISO 1991	on Day 21, remove adults from each chamber and determine total number and mass of live adults; at test end, count number of hatched cocoons, number of non-hatched cocoons with live juveniles, number of infertile cocoons, and total number of juvenile worms per test chamber	average mass and number of live adult worms per concentration, at start and end of test; mean and SD for number of cocoons, percent hatched cocoons, and number of juvenile worms per hatched cocoon, for each concentration	NOEC; if possible, an EC50 with confidence interval and slope of dose response curve
ISO 1998	on Day 28, remove adults from each chamber and record total number and mass of live adults; at test end, count number of juveniles in each test chamber	percent adult mortality, each test chamber and each concentration including the control; percent loss or increase in biomass of adults, each concentration; number of offspring produced in each concentration	NOEC and LOEC, if possible, an LC/EC50 [1]
OECD 2000	on Day 28, remove adults from each chamber and record total number and wet wt of live adults; at test end, count number of juveniles in each test chamber test chamber	percent adult mortality, each test chamber; the control; percent loss or increase in biomass of adults, each test container; number of juveniles produced in each	NOEC and ECx [2]

[1] Indicate, in mg/kg dry mass of soil substrate, the highest concentration tested without mortality, significant changes in biomass of adults, significant reduction in numbers of offspring (NOEC), the lowest concentration showing an effect (LOEC), and, if possible, an LC/EC50.

[2] A NOEC is likely to be required by regulatory authorities for the forseeable future. More widespread use of the ECx, resulting from statistical and ecological considerations, should be adopted soon. See Tables 6 and 9, this appendix, for recommended numbers of replicates per concentration, numbers of concentrations, and numbers of control chambers, depending on whether the endpoint to be determined is NOEC, ECx, or both NOEC and ECx.

13 Requirements for Valid Test; Use of Reference Toxicity Test

Document	Requirements for Valid Test	Reference Toxicant(s)	Procedures and Conditions for Reference Toxicity Test
ISO 1991	percent mortality of adults in controls ≤10%; control reproduction ≥...cocoons/worm/week, with ≥...juveniles per cocoon[1]	NI[2]	determine EC50 ≥once/year
ISO 1998	≥30 juveniles/control chamber; CV for control reproduction ≤30%; percent mortality of adults in control(s) ≤10%	carbendazim	determine effects on reproduction ($\alpha = 0.05$) for concentrations between 1 and 5 mg per kg dry wt of substrate
OECD 2000	≥30 juveniles/control chamber; CV for control reproduction ≤30%; percent mortality of adults in control(s) ≤10%	carbendazim	determine effects on reproduction ($\alpha = 0.05$) for concentrations between 1 and 5 mg per kg dry wt of substrate; calculate NOEC and/or ECx; perform ≥twice/year or in parallel with test substance

[1] Test validity requirements based on control reproduction rates yet to be established (draft document).
[2] NI = not indicated.

Natural and Artificial Negative Control Soils Used for Methods Development and the Establishment of Test Validity Criteria

Negative control soil must be included as one of the experimental treatments in each soil toxicity test. This treatment requires a soil which is essentially free of any contaminants that could adversely affect the performance of test organisms during the test (see Section 3.2). Before applying each of the test methods described in this document as a standardized test to be conducted according to Environment Canada, it was necessary to first assess the performance of test organisms in different types of negative control soil representative of an array of *clean* soils found within Canada. Five types of negative control soils were used to develop each of the biological test methods described herein and to further assess the robustness of each test method with samples of soil that varied considerably in their physical and chemical characteristics. These soils were also used to establish reasonable criteria for valid test results, based on control performance. The five soils tested include an artificial soil (see Section 3.2.2) and four natural soils (see Section 3.2.1) (Aquaterra Environmental, 1998a; Stephenson *et al*., 1999a, b, 2000a, Aquaterra Environmental and ESG, 2000; ESG, 2001, 2002; ESG and Aquaterra Environmental, 2002; and Stantec and Aquaterra Environmental, 2004). The artificial soil was formulated in the laboratory from natural ingredients. The four natural soils included two agricultural soils from southern Ontario, a prairie soil from Alberta, and a forest soil from northern Ontario. The physicochemical characteristics of all five soils are summarized in Table H-1 .

The artificial control soil (AS) used in this series of performance evaluation studies with diverse soil types was the same as that recommended for use herein (see Section 3.2.2). It consists of 70% silica sand, 20% kaolin clay, 10% *Sphagnum* sp. peat, and calcium carbonate (10–30 g per 1 kg peat). The soil was formulated by mixing the ingredients in their dry form thoroughly, then gradually hydrating with de-ionized water, and mixing further until the soil was visibly uniform in colour, texture, and degree of wetness. This artificial soil is much the same as that described by OECD (1984, 2000) and ISO (1991, 1993, 1998), except silica sand is used in the formulation rather than quartz sand.

The four natural soils used as negative control soil while developing these biological test methods and establishing the test validity criteria herein (see Sections 4.1.3, 4.2.3, and 4.3.3) do not represent all Canadian soil types. However, they do vary greatly in their physicochemical characteristics and include agricultural soils with diverse textures as well as a forest soil (see Table H-1). The soils originated from areas that had not been subjected to any direct application of pesticides in recent years. They were collected with either a shovel or a backhoe, depending on the location and the amount of soil collected. Sampling depth depended on the nature of the soil and the site itself.

The sample of clay loam soil, classified as a Delacour Orthic Black Chernozem, was collected in May 1995 from an undeveloped road allowance east of Calgary, Alberta. The soil beneath the sod was air dried to about 10–20% moisture content, sieved (4 or 9 mm), placed into 20-L plastic pails, and shipped to the University of Guelph (Guelph, ON) where it was kept in cold storage (4 °C) until needed. The soil was determined to be virtually free of any contaminants (Komex International, 1995). The physicochemical characteristics of the soil show that it is a moderate-to-fine clay loam, with a relatively high organic content and cation exchange capacity compared to the other *clean* soils used during the development of these biological test methods and the establishment of test validity criteria (see Table H-1).

Table H-1 Physicochemical Characteristics of Candidate Artificial and Natural Negative Control Soils [1]

Parameter	Artificial Soil	Clay Loam	Sandy Loam	Silt Loam	Forest Soil	Analytical Method
Source	formulated from constituents	field-collected from Alberta	field-collected from Ontario	field-collected from Ontario	field-collected from Ontario	—
Soil Texture	Fine Sandy Loam	Clay Loam	Fine Sandy Loam	Silt Loam	Loam	as per Hausenbuiller (1985); based on grain size distribution
Sand (%)	77.3	26.6	60.8	36.6	48.6	gravimetric grain size distribution
Silt (%)	7.8	43.3	27.8	50.1	36.9	gravimetric grain size distribution
Clay (%)	14.9	30.1	11.4	13.3	14.5	gravimetric grain size distribution
Gravel (%)	—[2]	—	0	0	0	gravimetric grain size distribution
Very Coarse Sand (%)	—	—	1.5	1.2	0.6	gravimetric grain size distribution
Coarse Sand (%)	—	—	3.2	2.3	2.2	gravimetric grain size distribution
Medium Sand (%)	—	—	10.1	5.4	9	gravimetric grain size distribution
Fine Sand (%)	—	—	25.9	13.4	20.4	gravimetric grain size distribution
Very Fine Sand (%)	—	—	20.2	14.3	16.4	gravimetric grain size distribution
Water-holding capacity (%)	71.5	80.3	44	56.5	75.6	gravimetric analysis[3]
pH (units)	6	5.9	7.3	7.4	4.2	0.01 *M* $CaCl_2$ method[4]
Conductivity (mS/cm)	0.3	1.52	0.092	0.373	0.39	saturated paste method
Bulk Density (g/cm^3)	0.98	0.83	—	—	0.51	clod method
Total Carbon (%)	4.46	6.83	1.88	2.57	11.9	Leco furnace method
Inorganic Carbon (%)	—	—	0.18	0.58	< 0.05	Leco furnace method
Organic Carbon (%)	—	—	1.7	1.99	11.9	Leco furnace method
Organic Matter (%)	9	12.8	2.9	3.5	19.9	dichromate oxidation

Parameter	Artificial Soil	Clay Loam	Sandy Loam	Silt Loam	Forest Soil	Analytical Method
Cation Exchange Capacity ($Cmol^+/kg$)	18.5	34.5	16.1	21.9	20	barium chloride method
Total Nitrogen (%)	0.05	0.59	0.115	0.166	0.74	Kjeldahl method
NH_4-N (mg/kg)	—	—	0.53	10.25	260	Kjeldahl method
NO_3-N (mg/kg)	—	—	6.94	5.44	2.26	Kjeldahl method
NO_2-N (mg/kg)	—	—	0.94	< 0.1	< 0.1	Kjeldahl method
Phosphorus (mg/kg)	23	12	6	10	35	nitric/perchloric acid digestion
Potassium (mg/kg)	22	748	61	75	250	NH_4 acetate extraction, colourimetric analysis
Magnesium (mg/kg)	149	553	261	256	192	NH_4 acetate extraction, colourimetric analysis
Calcium (mg/kg)	1848	5127	1846	4380	963	NH_4 acetate extraction, colourimetric analysis
Chloride (mg/kg)	—	—	69	42	113	H_2O extraction, colourimetric analysis
Sodium (mg/kg)	67	57	33	19	38	NH_4 acetate extraction, colourimetric analysis

[1] Characteristics of the artificial and various negative control soils that have been used to develop the definitive biological test methods and associated criteria for test validity described in this test methods document (Aquaterra Environmental, 1998a; Stephenson *et al.*, 1999a, b, 2000a; Aquaterra Environmental and ESG, 2000; ESG, 2001, 2002; ESG and Aquaterra Environmental, 2002; and Stantec and Aquaterra Environmental, 2004).

[2] Not determined.

[3] Determined according to USEPA (1989) using a Fisherbrand P8 creped filter paper (see Section 5.3).

[4] Determined by Becker-van Slooten *et al.* (2004) according to Hendershot *et al.* (1993) (see Section 4.1.5).

A large (~3000 L) sample of sandy loam soil was collected in June 1999 from Beauchamp Farms, Eramosa, Ontario, from a site that had been cultivated regularly for crop production but not subjected to pesticide application. The soil was air-dried and sieved (2 or 5 mm), placed into 20-L plastic buckets, and kept in cold storage (4 °C) until needed. This soil was analyzed for common organic and inorganic contaminants, and its physicochemical characteristics established to determine if any unusual soil characteristics (e.g., high conductivity or anomalous nutrient levels) were present. The sample was found to be virtually free of both contaminants and anomalies. This soil is a fine sandy loam with a moderate organic content and a moderate cation exchange capacity compared to the other *clean* soils included in these studies (see Table H-1).

The sample of silt loam soil was collected in June 1999 from the University of Guelph Elora Research Station, in Nichol Township, Ontario. The topsoil had been removed several years ago when the research facility was built, and had been stockpiled beside a field. Soil collected for these methods development studies was removed from the interior of the pile to avoid collecting soil that might have been inadvertently contaminated with pesticide or fertilizer spray drift from the adjacent field. The soil was air-dried and sieved (2 or 5 mm), placed into 20-L plastic buckets, and kept in cold storage (4 °C) until needed. The soil was also analyzed and found to be free of both organic and inorganic contaminants and anomalies. The measured physicochemical characteristics of this silt loam soil showed that it had a moderate organic content and a moderate cation exchange capacity, compared to the other four soils included in these methods development studies (see Table H-1).

A 400-L sample of forest soil, classified as Orthic Humo-Ferric Podzols, was collected in June 2001 from a forested area located on the Canadian Shield, in Sudbury, Ontario. The leaf litter was gently raked away and a hand trowel was used to remove soil to a depth ranging from 5–10 cm. The soil was placed without sieving into 20-L plastic-lined buckets, and transported to ESG International at Guelph, Ontario. It was air-dried for 48 hours to no less than ~10% moisture content, homogenized, and then sieved through 6-mm mesh. Once the sample was sieved, it was thoroughly homogenized and stored in the same 20-L plastic buckets until used. This soil was stored at room temperature (20 °C) until use. The physicochemical characteristics of the forest soil show that it is a loam with a moderate cation exchange capacity, and the highest total organic carbon content (11.9%) and highest percentage of organic matter (19.9%) of the five soils used in the methods development studies (see Table H-1).

Appendix I

Logarithmic Series of Concentrations Suitable for Toxicity Tests*

Column (Number of concentrations between 10.0 and 1.00, or between 1.00 and 0.10)**

1	2	3	4	5	6	7
10.0	10.0	10.0	10.0	10.0	10.0	10.0
3.2	4.6	5.6	6.3	6.8	7.2	7.5
1.00	2.2	3.2	4.0	4.6	5.2	5.6
0.32	1.00	1.8	2.5	3.2	3.7	4.2
0.10	0.46	1.00	1.6	2.2	2.7	3.2
	0.22	0.56	1.00	1.5	1.9	2.4
	0.10	0.32	0.63	1.00	1.4	1.8
		0.18	0.40	0.68	1.00	1.3
		0.10	0.25	0.46	0.72	1.00
			0.16	0.32	0.52	0.75
			0.10	0.22	0.37	0.56
				0.15	0.27	0.42
				0.10	0.19	0.32
					0.14	0.24
					0.10	0.18
					0.13	
					0.10	

* Modified from Rocchini *et al.* (1982).

** A series of five (or more) successive concentrations should be chosen from a column. Midpoints between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed on a weight-to-weight (e.g., mg/kg) or weight-to-volume (e.g., mg/L) basis. As necessary, values can be multiplied or divided by any power of 10. Column 1 might be used if there was considerable uncertainty about the degree of toxicity. More widely spaced concentrations should not be used, since such usage gives poor resolution of the confidence limits surrounding any threshold-effect value calculated.

Instruction on the Derivation of ICps Using Linear and Nonlinear Regression Analyses

J.1 Introduction

This appendix provides instruction for the use of linear and nonlinear regression analyses to derive, based on the concentration-response relationships for quantitative endpoint data (in this instance, number of surviving progeny and their dry weights at the end of an eight-week test with *Eisenia andrei*; see Section 4.3), the most appropriate ICps. It represents an adaptation and modification of the approach described by Stephenson *et al.* (2000b). Instructions herein are provided using Version 11.0 of SYSTAT*; however, any suitable software may be used. The regression techniques described in this appendix are most appropriately applied to continuous data from tests designed with ten or more concentrations or treatment levels (including the negative control treatment). The test design for measuring the effects of prolonged (eight-week) exposure on *E. andrei* is summarized in Table J.1.

An overview of the general process used to select the most appropriate regression model for each data set under consideration is presented in Figure 4 (see Section 6.4.2.1).

The reader is encouraged to refer to the appropriate sections within this biological test methods document, as well as the sections on regression analyses within the "Guidance Document on Statistical Methods for Environmental Toxicity Tests" (EC, 2004b) before data analyses. Environment Canada (2004b) also contains several additional references for the statistical analysis of quantitative test data using linear and nonlinear regression procedures. Some of the related guidance from these documents has been provided in this appendix, where appropriate.

* The latest (e.g., Version 11.0) version of SYSTAT™ is available for purchase by contacting SYSTAT Software, Inc., 501 Canal Boulevard, Suite C, Point Richmond, CA 94804-2028, USA, phone no. 1-800-797-7401; Web site www.systat.com/products/Systat/.

Table J.1 Summary of Test Design for Environment Canada's Eight-Week Test for Effects of Prolonged Exposure on the Survival, Reproduction, and Growth of *E. andrei* (see Section 4.3)

Parameter	Description
Test type	– whole soil toxicity test; no renewal (static test)
Test duration	– 56 days = 8 weeks
Test species	– *Eisenia andrei*; sexually mature adults with clitellum and individual wet wts ranging from 250–600 mg
Number of replicates	– 10 replicates per treatment
Number of treatments	– negative control soil and ≥7 test concentrations as a minimum; however, ≥10 concentrations plus a negative control are strongly recommended
Statistical endpoints	Quantal: • mean percent survival of adults in each treatment, on Day 28 • 28-d LC50 calculated if dose-response observed (using appropriate statistical procedures for quantal data; the procedures outlined in this appendix are not appropriate for quantal data) Quantitative: • mean number and dry mass of live juveniles in each treatment, on Day 56 • ICp (e.g., IC50 and/or IC25) for number and dry mass of live juveniles produced

J.2 Linear and Nonlinear Regression Analyses

J.2.1 Creating Data Tables

Note: The statistical analysis must encompass the transformation of the concentrations logarithmically (e.g., $\log_{10}$ or $\log_e$). If the concentrations fall below one (1) (e.g., 0.25), then the data can be transformed by transforming the units (e.g., from mg/kg to µg/g) with a multiplication factor (e.g., 1000); the modified data are then transformed logarithmically. The transformation can be done either in the original electronic spreadsheet, or when the original data are transferred to the SYSTAT data file.

1) Open the appropriate file containing the data set in an electronic spreadsheet.

2) Open the SYSTAT program. In the main screen, go to **File**, **New**, and then **Data**. This will open up an empty data table. Insert the variable names into the column heading by double clicking on a variable name, which opens the '**Variable Properties**' window. Insert an appropriate name for the variable of interest within the '**Variable name**' box, and select the variable type; additional comments can be inserted within the '**Comments:**' box. For example, the following variable names might be used:

conc = concentration or treatment level
logconc = $\log_{10}$ transformation of concentration or treatment level

rep = replicate within a treatment level
juveniles = number of juveniles produced
jdrywt = dry mass of juveniles produced

3) The data can now be transferred. To transfer the data, copy and paste each column from the electronic spreadsheet containing the concentrations, the replicates, and associated mean values, to the SYSTAT data table.

4) Save the data by going to **File**, then **Save As**; a **'Save As**' window will appear. Use appropriate coding to save the data file. Select **Save** when the file name has been entered.

5) Record the file name of the SYSTAT data file in the electronic spreadsheet containing the original data.

6) If the data (i.e., the test concentrations) require transformation, the data can be transformed by selecting **Data**, **Transform**, and then **Let...**. Once in the **Let...** function, select the column heading containing the appropriate header for the transformed data (e.g., logconc), and then select **Variable** within the '**Add to**' box to insert the variable into the '**Variable:**' box. Select the appropriate transformation (e.g., L10 for $\log_{10}$ transformation or LOG for the natural logarithm) in the '**Functions:**' box (the '**Function Type:**' box should be **Mathematical**), and then select **Add** to insert the function into the '**Expression:**' box. Select the column heading containing the original untransformed data (i.e., 'conc' for concentration or treatment level), followed by **Expression** within the '**Add to**' box to insert the variable into the '**Expression:**' box. If a multiplication factor is required to adjust the concentration before the log-transformation, this step can be completed within the '**Expression:**' box (e.g., L10[conc*1000]). Select **OK** when all desired transformations have been completed. The transformed data will appear in the appropriate column. *Save the data* (i.e., select **File**, followed by **Save**).

   Note: The $\log_{10}$ of the negative control treatment cannot be determined (i.e., the $\log_{10}$ of zero is undefined); therefore, assign the negative control treatment level a very small number (e.g., 0.001) known or assumed to be a no-effect level, to include this treatment in the analysis and differentiate it from the other transformed treatment levels.

7) From the data table, calculate and record the mean of the negative controls for the variable under study; each measurement endpoint is statistically analyzed independently. The mean value of these control data will be required when estimating the model parameters. In addition, determine the maximum value within the data set for that particular variable and round up to the nearest whole number. This number is used as the maximum value of the y-axis (i.e., 'ymax') when creating a graph of the regressed data.

J.2.2 Creating a Scatter Plot or Line Graph

The scatter plots and line graphs provide an indication of the shape of the concentration-response curve for the data set. The shape of the concentration-response curve can then be compared to each model (Figure J.1) so that the appropriate model(s) likely to best suit the data is (are) selected. Each of the selected models should be used to analyze the data. Subsequently, each model is reviewed, and the model that demonstrates the best fit is selected.

1) Select **Graph**, **Summary Charts**, and then **Line...**. Select the independent variable (e.g., logconc), followed by **Add** to insert the variable into the '**X-variable(s):**' box. Select the dependent variable under examination, followed by **Add** to insert the variable into the '**Y-variable(s):**' box. Select **OK**. A graph will be displayed within the '**Output Pane**' of the main SYSTAT screen containing the mean values for every treatment level; to view a larger version of the graph, simply select the '**Graph Editor**' tab located below the central window. A scatter plot of the data can also be viewed by selecting **Graph**, **Plots**, and then **Scatterplot...** and following the same

instructions for inserting the x- and y-variables. The graphs will provide an indication as to the general concentration-response trend allowing the selection of the potential model(s) of best fit to be chosen, in addition to an estimation of the ICp of interest.

Note: The main SYSTAT screen is divided into three parts. The left-hand side of the screen (i.e., '**Output Organizer**' tab) provides a list of all of the functions completed (e.g., graphs) – each function can be viewed by simply selecting the desired icon. The right-hand side of the screen forms the central window in which the general output of all of the functions completed (e.g., regression, graphs) can be viewed. The tabs below this central window allow the user to toggle between the data file (i.e., '**Data Editor**'), individual graphs (i.e., '**Graph Editor**') and the output (i.e., '**Output Pane**'). The various graphs produced can be viewed individually within the '**Graph Editor**' tab by selecting the graph of interest within the left-hand side of the screen (i.e., '**Output Organizer**' tab). The bottom portion of the screen displays the command codes used to derive the desired functions (e.g., regression and graphing codes). The '**Log**' tab within this command screen displays a history of all of the functions that have been completed.

2) Visually estimate and record an estimate of the ICp of interest (e.g., IC50) for the data set. For example, for an IC50, divide the average of the controls by two, and find this value on the y-axis. Estimate a horizontal line from the y-axis until the line intercepts the data points. At this intersection point, extend a vertical line down towards the x-axis and record this concentration value as an estimate of the IC50.

3) Using the scatter plots or line graphs, select the potential model(s) that will best describe the concentration-response trend (refer to Figure J.1 for an example of each model).

J.2.3 Estimating the Model Parameters

1) Select **File**, **Open**, and then **Command**.

2) Open the file containing the command codes for the particular model chosen from Section J.2.2 (i.e., select the appropriate file, followed by **Open**):

nonline.syc	=	exponential model
nonling.syc	=	gompertz model
nonlinh.syc	=	logistic with hormesis model
linear.syc	=	linear model
nonlinl.syc	=	logistic model

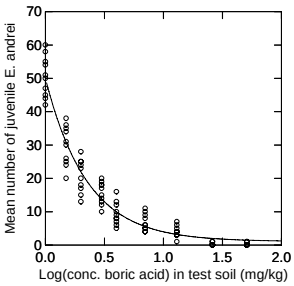
The file will provide the command codes for the selected model within the appropriate tab of the command editor box at the bottom of the main screen. All of the command codes for deriving IC50s and IC25s are provided in Table J.2; however, the equations can be formatted to derive any ICp. For example, the command codes for the logistic model to derive an IC50 would be:

```
nonlin
print=long
model jdrywt = t/(1+logconc/x)^b)
save resid1/ resid
estimate/ start = 0.5, 0.6, 2 iter=200
use resid1
pplot residual
plot residual*logconc
plot residual*estimate
```

Exponential Model

IC50: juveniles = a*exp(log((a-a*0.5-b*0.5)/a)*(logconc/x))+b
IC25: juveniles = a*exp(log((a-a*0.25-b*0.75)/a)*(logconc/x))+b

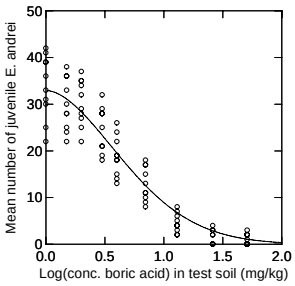
Where:
a = the y-intercept (the control response)
x = ICp for the data set
logconc = the logarithmic value of the exposure concentration
b = a scale parameter (estimated between 1 and 4)



Gompertz Model

IC50: juveniles = g*exp((log(0.5))*(logconc/x)^b)
IC25: juveniles = g*exp([(log(0.75))*(logconc/x)^b)

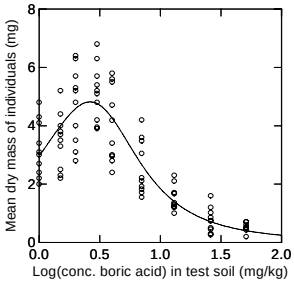
Where:
g = the y-intercept (the control response)
x = ICp for the data set
logconc = the logarithmic value of the exposure concentration
b = a scale parameter (estimated between 1 and 4)



Hormesis Model

IC50: jdrywt = (t*(1+h*logconc))/(1+((0.5+h*logconc)/0.5)*(logconc/x)^b)
IC25: jdrywt = (t*(1+h*logconc))/(1+((0.25+h*logconc)/0.75)*(logconc/x)^b)

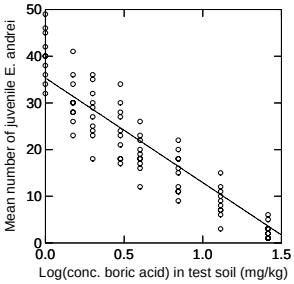
Where:
t = the y-intercept (the control response)
h = the hormetic effect (estimated between 0.1 and 1)
x = ICp for the data set
logconc = the logarithmic value of the exposure concentration
b = a scale parameter (estimated between 1 and 4)



Linear Model

IC50: juveniles = ((-b*0.5)/x)*logconc+b
IC25: juveniles = ((-b*0.25)/x)*logconc+b

Where:
b = the y-intercept (the control response)
x = ICp for the data set
logconc = the logarithmic value of the exposure concentration



Logistic Model

IC50: jdrywt = t/(1+(logconc/x)^b)
IC25: jdrywt = t/(1+(0.25/0.75)*(logconc/x)^b)

Where:
t = the y-intercept (the control response)
x = ICp for the data set
logconc = the logarithmic value of the exposure concentration
b = a scale parameter (estimated between 1 and 4)

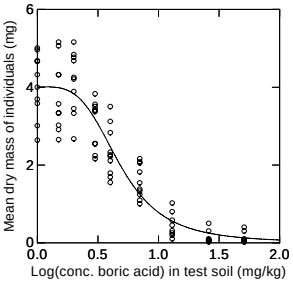


Figure J.1 SYSTAT Version 11 Equations for Linear and Nonlinear Regression Models and Example Graphs of the Observed Trends for Each Model

Table J.2 SYSTAT Command Codes for Linear and Nonlinear Regression Models

Model	Command Codes	
Exponential	nonlin print = long model juveniles = a*exp(log((a-a*0.5-b*0.5)/a)*(logconc/x))+b save resid1/ resid estimate/start = 50[a], 1[b], 0.3[c] iter=200 use resid1 pplot residual plot residual*logconc plot residual*estimate	where: [a] represents the estimate of the y-intercept (i.e., 'a') (the control response) [b] represents the scale parameter (i.e., 'b') (estimated between 1 and 4) [c] represents the estimate of the ICp for the data set (i.e., 'x')
Gompertz	nonlin print = long model juveniles = g*exp((log(0.5))*(logconc/x)^b) save resid2/ resid estimate/start = 32[a], 0.4[b], 1[c] iter=200 use resid2 pplot residual plot residual*logconc plot residual*estimate	where: [a] represents the estimate of the y-intercept (i.e., 'g') (the control response) [b] represents the estimate of the ICp for the data set (i.e., 'x') [c] represents the scale parameter (i.e., 'b') (estimated between 1 and 4)
Hormesis	nonlin print = long model jdrywt = (t*(1+h*logconc))/(1+((0.5+h*logconc)/0.5)*(logconc/x)^b) save resid3/ resid estimate/ start = 3[a], 0.1[b], 0.7[c], 1[d] iter=200 use resid3 pplot residual plot residual*logconc plot residual*estimate	where: [a] represents the estimate of the y-intercept (i.e., 't') (the control response) [b] represents the hormetic effect (i.e., 'h') (estimated between 0.1 and 1) [c] represents the estimate of the ICp for the data set (i.e., 'x') [d] represents the scale parameter (i.e., 'b') (estimated between 1 and 4)
Linear	nonlin print = long model juveniles=((-b*0.5)/x)*logconc+b save resid4/ resid estimate/ start=36[a], 0.7[b] iter=200 use resid4 pplot residual plot residual*logconc plot residual*estimate	where: [a] represents the estimate of the y-intercept (i.e., 'b') (the control response) [b] represents the estimate of the ICp for the data set (i.e., 'x')
Logistic	nonlin print = long model juveniles=t/(1+(logconc/x)^b) save resid5/ resid estimate/ start=4[a], 0.7[b], 2[c] iter=200 use resid5 pplot residual plot residual*logconc plot residual*estimate	where: [a] represents the estimate of the y-intercept (i.e., 't') (the control response) [b] represents the estimate of the ICp for the data set (i.e., 'x') [C] represents the scale parameter (i.e., 'b') (estimated between 1 and 4)

3) Type in the header of the column in the data table containing the variable of interest to be analyzed within the line entitled 'model y=' (where 'y' is the dependent variable, e.g., jdrywt).

4) The fourth line of the text should read 'save resida/ resid', where 'a' indicates a number to which the residual file is assigned. Substitute this same number into the 6th line (i.e., 'use resida') so that the same file is used to generate a normal probability plot and graphs of the residuals. The command lines that follow provide instruction for the generation of a probability plot (i.e., 'pplot residual'), the generation of a graph of residuals against the concentration or treatment level (i.e., 'plot residual*logconc'), and a graph of the residuals against the predicted and fitted values (i.e., 'plot residual*estimate'). These graphs are used to aid in the assessment of the assumptions of normality (e.g., probability plot) and homogeneity of the residuals (e.g., graphs of the residuals) when evaluating for the model of best fit (Section J.2.4).

5) Substitute the mean of the controls and the estimated ICp (e.g., IC50 and/or IC25) within the fifth line entitled 'estimate/start =' (refer to Table J.2 for details on the substitution for each model). These values were initially derived from examination of the scatter plot or line graph. The model, once it converges, will provide a set of parameters from which the ICp, and its 95% confidence limits, are reported (i.e., parameter 'x'). It is essential to provide accurate estimates for each parameter before running the model, or the iterative procedure used to derive the reported parameters might not converge. The scale parameter (Table J.2) is typically estimated to range from one to four. The number of iterations can be changed, but for this example, has been set to 200 (i.e., 'iter = 200'). Typically, 200 iterations are sufficient for a model to converge; if more iterations are required, it is likely that the most appropriate model is not being applied.

6) Select **File**, and then **Submit Window** to run the commands; alternatively, right-click the mouse and select **Submit Window**. This will generate a printout of the iterations, the estimated parameters, and a list of the actual data points with the corresponding predicted values and residuals. A preliminary graph of the estimated regression line will also be presented; this preliminary graph should be deleted. The graph can be deleted by selecting the graph in the left-hand window within the main screen. A normal probability plot and graphs of the residuals will also be presented.

J.2.4 Examining the Residuals and Test Assumptions

An examination of the residuals for each model tested helps to determine whether assumptions of normality and homoscedasticity have been met. If any of the assumptions cannot be met, regardless of the model examined, a statistician should be consulted for further guidance on assessing additional models or the data should be re-analyzed using the less desirable linear interpolation method of analysis (using ICPIN; see Section 6.4.2.2).

J.2.4.1 Assumptions of Normality. Normality should be assessed using *Shapiro-Wilk's test* as described in EC (2004b); Section J.2.4.3 provides instructions for conducting this test. The normal probability plot, displayed in the '**Output Pane**', can also be used to evaluate whether the assumption of normality is met. The residuals should form a fairly straight line diagonally across the graph; the presence of a curved line represents deviation from normality. The normal probability plot should not, however, be used as a stand-alone test for normality, since the detection of a 'normal' (e.g., straight) or 'non-normal' (e.g., curved) line depends on the subjective assessment of the user. If the data are not normally distributed, then the user should try another model, consult a statistician for further guidance, or the data should be analyzed using the less desirable linear interpolation method of analysis.

J.2.4.2 Homogeneity of Residuals. Homoscedasticity (or homogeneity) of the residuals should be assessed using *Levene's test* as described in EC (2004b) (Section J.2.4.3 provides instructions for conducting this test), and by examining the graphs of residuals against the actual and predicted (estimated) values. Homogeneity of the residuals is described by an equal distribution of the variance of the residuals across the independent variable (i.e., concentration or treatment level) (Figure J.2A). Levene's test, if significant, will indicate that the data are not homogeneous. If the data (as indicated by Levene's test) are heteroscedastic (i.e., not homogeneous), then the graphs of the residuals should be examined. If there is a significant change in the variance and the graphs of the residuals produce a distinct fan or 'V' pattern (refer to Figure J.2B for a plot of the 'residual*estimate'; a corresponding 'V' pattern in the opposite direction also occurs in the plot of the 'residual*logconc'), then the data analysis should be repeated using weighted regression. Alternatively, a divergent pattern suggestive of a systematic lack of fit (Figure J.2C) will indicate that an inappropriate or incorrect model was selected.

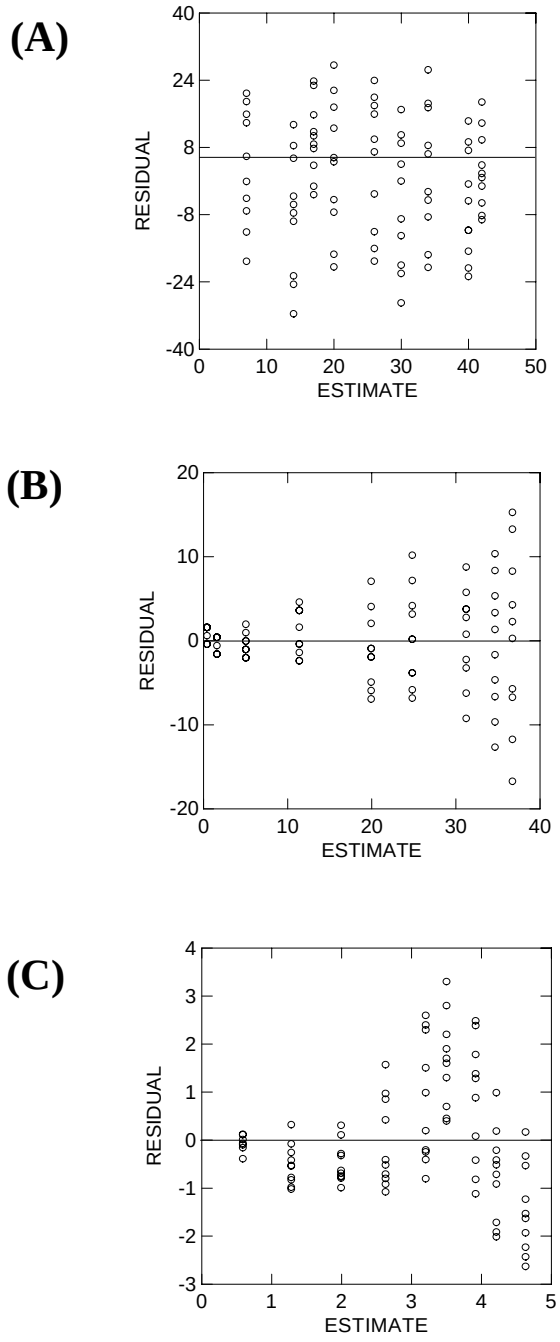


Figure J.2 **Graph of the Residuals Against the Predicted (Estimated) Values (i.e., 'residuals*estimate') Indicating Homoscedasticity (A), and Two Types of Heteroscedasticity; One Demonstrating a Fan or 'V' Shape (B) Requiring Further Examination Using Weighted Regression, and a Second Demonstrating a Systematic Lack of Fit (C) as a Result of the Selection of an Incorrect Model**

J.2.4.3 Assessing Assumptions of Normality and Homogeneity of Residuals. SYSTAT Version 11.0 can perform both Shapiro-Wilk's and Levene's tests to assess the assumptions of normality and homogeneity of residuals. Levene's test can only be performed by conducting an analysis of variance (ANOVA) on the absolute values of the residuals derived in Section J.2.3.

1) Select **File**, **Open**, and then **Data** to open the data file containing the residuals created in Section J.2.3 (e.g., resid1.syd).

2) Insert a new variable name into an empty column by double-clicking on the variable name, which opens the '**Variable Properties**' window. In this window, insert an appropriate name for the transformed residuals (e.g., absresiduals) into the '**Variable name:**' box. Transform the residuals by selecting **Data**, **Transform**, and then **Let...**. Once in the **Let...** function, select the column heading containing the appropriate header for the transformed data (e.g., absresiduals), and then select **Variable** within the '**Add to**' box to insert the variable into the '**Variable:**' box. Select the appropriate transformation (e.g., ABS for the transformation of data into its absolute form) in the '**Functions:**' box (the '**Function Type:**' box should be **Mathematical**), and then select **Add** to insert the function into the '**Expression:**' box. Select the column heading containing the original untransformed data (i.e., residuals), followed by **Expression** within the '**Add to**' box to insert the variable into the '**Expression:**' box. Select **OK**; the transformed data will appear in the appropriate column. Save the data.

3) To perform Shapiro-Wilk's test, select **Analysis**, **Descriptive Statistics**, and then **Basic Statistics...**. A '**Column Statistics**' window will appear. Select the residuals from the '**Available variable(s):**' box, followed by **Add** to insert this variable into the '**Selected variable(s):**' box. Within the '**Options**' box, select the **Shapiro-Wilk normality test**, followed by **OK**. A small table will appear within the SYSTAT Output Organizer window, where the Shapiro-Wilk critical value (i.e., 'SW Statistic') and probability value (i.e., 'SW P-Value') will be displayed. A probability value greater than the usual criterion of $p > 0.05$ indicates that the data are normally distributed.

4) To perform Levene's test, select **Analysis**, **Analysis of Variance (ANOVA)**, and then **Estimate Model...**, an '**Analysis of Variance: Estimate Model**' window will appear.

5) Select the variable within which the data are to be grouped (e.g., logconc), and place this variable into the '**Factor(s):**' box by selecting **Add**.

6) Select the transformed residuals (i.e., absresiduals), followed by **Add**, to insert the variable into the '**Dependent(s):**' box. Select **OK**. A graph of the data and a printout of the output will appear within the '**Output Pane**' tab. A probability value greater than the usual criterion of $p > 0.05$ indicates that the data are homogeneous.

J.2.5 Weighting the Data

If the residuals are heteroscedastic, as indicated by Levene's test, and there is a significant change in variance across treatment levels (i.e., the presence of a distinct fan or 'V' shape; refer to Figure J.2B), the data should be re-analyzed using weighted regression. Weighted regression involves using the inverse of the variance of observations within each concentration or treatment level as the weights. When performing the weighted regression, the standard error for the ICp (presented in SYSTAT as the asymptotic standard error ('A.S.E.'; refer to Figure J.3) is compared to that derived from the unweighted regression. If there is a difference of greater than 10% between the two standard errors, then the weighted regression is selected as the regression of best choice. However, if there is a significant change in variance across all treatment levels, and there is less than a 10% difference in the standard error between the weighted and unweighted regressions **, then the user should consult a statistician for further guidance and the application of

** The value of 10% is only a "rule-of-thumb" based upon experience. Objective tests for the improvement due to weighting are available, but beyond the scope of this document. Weighting should be used only when necessary, since the procedure can introduce additional complications to the modeling procedure. A statistician should be consulted when weighting is necessary, but the parameter estimates are nonsensical.

```
SYSTAT Rectangular file C:\SYSTAT\STATAPP.SYS,
created Tue May 25, 2004 at 13:46:14, contains variables:
          CONC        REP        LOGCONC    JUVENILES    JDRYWT


 Iteration
 No.      Loss      G           X           B
   0 .452080D+04 .340000D+02 .400000D+00 .100000D+01
   1 .184579D+04 .328003D+02 .708478D+00 .157121D+01
   2 .157417D+04 .331384D+02 .696189D+00 .197718D+01
   3 .156445D+04 .329695D+02 .702780D+00 .211068D+01
   4 .156432D+04 .329461D+02 .703292D+00 .212794D+01
   5 .156432D+04 .329427D+02 .703387D+00 .212931D+01
   6 .156432D+04 .329424D+02 .703394D+00 .212941D+01

Dependent variable is JUVENILES

    Source   Sum-of-Squares    df  Mean-Square
 Regression       41208.683     3    13736.228
   Residual        1564.317    87       17.981 ------------> residual mean square error

      Total       42773.000    90
Mean corrected    15140.456    89

       Raw  R-square (1-Residual/Total)        =        0.963
Mean corrected R-square (1-Residual/Corrected) =        0.897
         R(observed vs predicted) square       =        0.897

                                                    Wald Confidence Interval
Parameter         Estimate       A.S.E.    Param/ASE       Lower < 95%> Upper
 G                  32.942        1.031       31.952       30.893       34.992
 X                   0.703        0.031       22.898        0.642        0.764
 B                   2.129        0.229        9.299        1.674        2.585

          JUVENILES   JUVENILES
  Case     Observed    Predicted     Residual
      1       36.000       32.942        3.058
      2       31.000       32.942       -1.942
      3       22.000       32.942      -10.942
      4       25.000       32.942       -7.942
      5       39.000       32.942        6.058
      6       42.000       32.942        9.058
      7       39.000       32.942        6.058
      .       ......       ......       ......
      .       ......       ......       ......
      .       ......       ......       ......
      .       ......       ......       ......
     86        2.000        0.337        1.663
     87        0.000        0.337       -0.337
     88        0.000        0.337       -0.337
     89        1.000        0.337        0.663
     90        0.000        0.337       -0.337

Asymptotic Correlation Matrix of Parameters

     G           X           B
G        1.000
X       -0.696       1.000
B       -0.611       0.566       1
```

ICp, asymptotic standard error, and lower and upper 95% confidence limits

Figure J.3 **Example of the Initial Output Derived using the Gompertz Model in SYSTAT Version 11.** The initial output provides the residual mean square error used to select the model of best choice, as well as the ICps, the standard error for the estimate, and the upper and lower 95% confidence limits. The number of cases displayed has been shortened for the purpose of this diagram; however, the output within SYSTAT displays all cases including the actual variable measurement and the corresponding predicted estimate and residual.

additional models, or the data could be re-analyzed using the less desirable linear interpolation method of analysis. The comparison between weighted and unweighted regression is completed for each of the selected models while proceeding through the process of final model selection (i.e., model and regression of best choice). Alternatively, if Levene's test demonstrates that the data are not homogeneous, and the graphs of the residuals demonstrate a non-divergent pattern (e.g., Figure J.2C), an inappropriate or incorrect model might have been selected. The user is then advised to consult a statistician for further guidance on the use and application of alternate models.

1) Select **File**, **Open**, and then **Data**. Select the file containing the data set to be weighted. Insert the two new variable names into the column heading by double-clicking on a variable name, which opens the '**Variable Properties**' window. In this window, insert an appropriate name for the variable of interest, select the variable type, and specify comments if desired. The two new column headings should indicate the variance of a particular variable (e.g., varjdrywt), and the inverse of the variance for that variable (e.g., varinvsjdrywt). Save the data file by selecting **File**, and then **Save**.

2) Select **Data**, followed by **By Groups...**. Select the independent variable (i.e., logconc), followed by **Add**, to insert this variable into the '**Selected variable(s):**' box; this will enable the determination of the variance of the variable of interest by concentration or treatment level (i.e., "group"). Select **OK**.

3) Select **Analysis**, **Descriptive Statistics**, and then **Basic Statistics...**. Select the variable of interest to be weighted (e.g., jdrywt), followed by **Add** to insert this variable into the '**Selected variable(s):**' box. Select **Variance** within the '**Options**' box, followed by **OK**. This function will display the variance for the variable of interest, grouped by concentration or treatment level within the '**Output Pane**' tab of the main screen.

4) Select **Data**, **By Groups...**, and then click on the box beside **Turn off**, and select **OK** so that any analysis that follow will not be analyzed according to each individual concentration or treatment level; the analysis should consider the entire data set as a whole.

5) Return to the data file by selecting the '**Data Editor**' tab within the main screen. Transfer the variances for each concentration or treatment level to the corresponding concentration within the variance column (e.g., varjdrywt). Note that the variance is the same among replicates within a treatment.

6) Select **Data**, **Transform**, and then **Let...**, and select the column heading containing the inverse of the variance (e.g., varinvsjdrywt) for the variable of interest, followed by **Variable** within the **'Add to**' box to insert the variable into the '**Variable:**' box. Select the '**Expression:**' box and type in '1/', and then select the column heading containing the variances (e.g., varjdrywt) of the variable of interest for each replicate and concentration, followed by **Expression** within the '**Add to**' box to insert the variable into the '**Expression:**' box. Select **OK**. The inverse of the variance for each replicate and concentration will be displayed in the appropriate column. Save the data by selecting **File**, and then **Save**.

7) Select **File**, **Open**, and then **Command**; open the file containing the command codes for estimating the equation parameters (e.g., Section J.2.3, step 2) for the same model selected for the *unweighted* analysis.

8) Insert an additional row after the third line by typing 'weight=varinvsy', where 'y' is the dependent variable to be weighted (e.g., weight=varinvsjdrywt), as per the shaded area below:

```
nonlin
print=long
model jdrywt = t/(1+(logconc/x)^b)
weight=varinvsjdrywt
save resid2/ resid
```

```
estimate/ start = 0.5, 0.6, 2 iter=200
use resid2
pplot residual
plot residual*logconc
plot residual*estimate
```

9) Assign a new number for the residuals within the line entitled 'save resida' (where 'a' represents the assigned number).

10) Substitute the mean of the controls and the estimated ICp within the line entitled 'estimate/ start. . .' (refer to Table J.2 for details on the substitution for each model). These estimates will be the same as those used for the unweighted analysis.

11) Select **File**, and then **Submit Window** to run the commands. This will generate output of the iterations, the estimated parameters, and a list of the data points with the corresponding predicted data points and residuals within the '**Output Pane**' tab of the main screen. A preliminary graph of the estimated regression line will also be presented; this should be deleted. A normal probability plot and graphs of the residuals will also be presented.

12) Proceed with the analysis as described in Section J.2.4 to ensure that all model assumptions have been met.

13) Compare the weighted regression analysis with the unweighted regression analysis. Select the weighted regression if weighting reduced the standard error for the ICp by 10%, relative to the unweighted regression analysis.

J.2.6 The Presence of Outlier(s) and Unusual Observations

Outliers are indicative of a measurement that does not seem to fit the other values derived from the test. Outliers and unusual observations can be identified by examining the fit of the concentration-response curve relative to all data points, and by examining the graphs of the residuals. If an outlier has been observed, the test records (e.g., hand-recorded and electronic data sheets and experimental conditions) should be scrutinized for human error. If the outlier is a data point that has been obtained through a transcription error than cannot be corrected, or through a faulty procedure (e.g., adult *E. andrei* within one test unit were accidentally not removed, and left for 8 weeks of reproduction), then the data point should be removed from the analysis. If an outlier has been identified, the analysis should be completed with and without the presence of the outlier. The decision on whether or not to remove the outlier should also take into consideration natural biological variation, and biological reasons that might have caused the apparent anomaly. Regardless of whether or not the outlier is removed, a description of the data, outliers, analyses with and without the outlier, and interpretive conclusions, must accompany the final analysis. If it appears as if there is more than one outlier present, the selected model should be re-assessed for appropriateness and alternative models considered. Additional guidance on the presence of outliers and unusual observations is provided in EC (2004b) and should be consulted for further details.

The Analysis of Variance (ANOVA) function within SYSTAT can be performed to determine whether or not the data contain outliers. However, ANOVA assumes that the residuals are normally distributed, and therefore, assumptions of normality must be met before using the ANOVA to detect outliers. The presence of outliers can also be determined from the graphs of residuals.

1) Perform an Analysis of Variance (ANOVA) as described in Section J.4 of this appendix, to determine whether any outliers exist. Any outlier(s) will be identified as a case number that corresponds with the row number in the SYSTAT data file. The program uses the studentized residuals as an indication of outliers; values >3 indicate the possibility of an outlier. This should be confirmed with the graphs of the residuals.

2) If a decision is made to remove the outlier(s), delete the value from the original data table (file), and re-save the file under a *new* name (i.e., select **File**, and then **Save As...**). For example, the new file name might contain the letter 'o' (for outlier(s) removed) at the end of the file's original name

3) Repeat the regression analysis with the outlier(s) removed, using the same model and estimated parameters that were used before the outlier(s) were removed. Alternatively, additional models may be used for analysis if the alternative model results in a better fit and smaller residual mean square error. If the removal of the outlier(s) does not result in a significant change to both the residual mean square error and the ICp (including its corresponding confidence intervals), then the individual performing the analysis must make a subjective decision (i.e., professional judgement) as to whether or not to include the outlier(s). Justification for the removal or inclusion of the outlier(s) must be recorded along with the final analysis.

J.2.7 Selection of the Most Appropriate Model

Once all of the contending models have been fit, each one should be assessed for normality, homogeneity of the residuals, and the residual mean square error. The model which meets all of the assumptions and has the smallest residual mean square error (refer to Figure J.3) should be selected as the most appropriate model. However, in the case where more than one model has the same residual mean square error, and all other factors are equivalent, the simplest model should be selected as the model of best choice. If a weighted regression was performed, the weighted and unweighted analyses should be compared and the weighted analysis selected if weighting reduced the standard error for the ICp by more than 10%. The residual mean square error is presented in the '**Output Pane**' tab just following the iterations, and preceding the parameter estimates. However, if none of the models adequately fit the data, then the user is advised to consult a statistician for the application of additional models, or the data should be re-analyzed using the less desirable linear interpolation method of analysis (see Section 6.4.2.2).

Note: Since the concentration or treatment levels were logarithms in the calculations, the ICps and their confidence limits should be transformed to arithmetic values for the purpose of reporting them.

J.2.8 Creating the Concentration-Response Curve

Once an appropriate model has been selected, the concentration-response curve for that particular model must be generated.

1) Within the command editor window at the bottom of the screen, copy the model equation (i.e., the equation after the '=' sign, third line of the command codes depicted in Table J.2) from the command codes used to derive the estimates for the selected model; the equation should consist of the original alphabetic characters (e.g., t, b, h, etc.). The equation can be copied by highlighting the equation and selecting **Edit**, followed by **Copy** (or right-clicking the mouse and selecting **Copy**).

2) Select **File**, **Open**, and then **Command** and open an existing graph command file (i.e., any file with '*.cmd') similar to the following example (or, if and as necessary, create a new one), using the logistic model. The first plot (i.e., 'plot') is a scatter plot of the dependent variable against the log concentration series. The second plot (i.e., 'fplot') is the regression equation, which is superimposed upon the scatter plot.

```
graph
begin
plot jdrywt*logconc/ title = 'Dry Mass of Juvenile E. andrei', xlab = 'Log(ug boric acid/kg soil d.wt)',
ylab='Mass (mg)',
xmax = 2, xmin = 0, ymax = 6, ymin = 0
fplot y=4.01241/(1+(logconc/0.68263)^3.72182); xmin = 0,
xmax =2, xlab ='' ymin = 0, ylab ='', ymax = 6
end
```

3) Paste the previously copied equation in place of the pre-existing equation (as seen in the shaded area above) by highlighting the previous equation, and then selecting **Edit**, followed by **Paste** (or right-clicking the mouse and selecting **Paste**). Replace all of the alphabetical characters (e.g., t, b, h, x, a, etc.), together with the respective estimates, provided in the '**Output Pane**' tab generated by the application of the selected model.

4) Type in the correct information within the line entitled 'plot y*logconc . . .', where 'y' is the dependent variable under study (e.g., jdrywt). Adjust the 'xmax' (i.e., the maximum log-concentration used) and 'ymax' (refer to Section J.2.1, Step 7) numerical values accordingly. Ensure that all 'xlab' and 'ylab' (i.e., axis labels) entries are correct, and if not, adjust accordingly. Ensure that all quotation marks and commas are placed within the command program as depicted in the previous example; SYSTAT is case- and space- insensitive.

Note:		
	'title'	refers to the title of the graph
	'xlab'	refers to the x-axis label
	'xmin'	refers to the minimum value requested for the x-axis
	'xmax'	refers to the maximum value requested for the x-axis
	'ylab'	refers the y-axis label
	'ymax'	refers to the maximum value requested for the y-axis
	'ymin'	refers to the minimum value requested for the y-axis

The 'xmin', 'xmax', 'ymin', and 'ymax' must be the same for both plots to superimpose the regression line accurately on the scatter plot of the data. An example of the final regression graph is provided in Figure J.1 for each of the five proposed models.

5) Select **File**, then **Save As** to save the graph command codes in an appropriate working folder using the same coding used to generate the data file, with indication as to which model the regression corresponds to. Select **Save** to save the file.

6) Select **File**, then **Submit Window** to process the command codes. A graph of the regression, using the model estimate parameters for the selected model, will appear.

J.3 Determining Additional ICps

In some cases, it might be desirable to estimate another value for 'p' (besides or instead of an IC50). The models proposed by Stephenson *et al*. (2000b) enable the selection and determination of any ICp. The following section, as well as Figure J.1, provide guidance on determining an IC25, however, the models can be changed to suit any 'p' value (e.g., IC20).

1) Select **File**, **Open**, and then **Command** and open the file corresponding to the command codes used to generate the estimate parameters (refer to Table J.2 for the command codes for each model). Change the model equation such that it will calculate the desired ICp (e.g., IC25); Figure J.1 provides guidance on adjusting the models to calculate the IC25. However, any ICp can be determined by modifying the fractions used in each model. For example, to calculate an IC20 using the logistic model, the equation would change from 't/[1+(logconc/x)^b]' (for calculating an IC50) to 't/[1(0.20/0.80)*(logconc/x)^b]' (for calculating an IC20).

2) Once the equation has been adjusted for the ICp of interest, follow each step outlined in Section J.2.3 of this appendix. However, substitute the estimated ICp (e.g., IC25) within the fifth line entitled 'estimate/ start=' (refer to Figure J.1 for details on the substitution for each model). These values were initially derived from an examination of the scatter plot or line graph. The model, once it converges, will provide a set of parameters from which the ICp, and its corresponding 95% confidence limits, are reported (i.e., parameter 'x').

3) Proceed with the analysis as described in Sections J.2.4 to J.2.8 herein.

J.4 Analysis of Variance (ANOVA)

1) Select **File**, **Open**, and then **Data** to open the data file containing all of the observations for the data set under examination.

2) Select **Analysis**, **Analysis of Variance (ANOVA)**, and then **Estimate Model...**.

3) Select the variable within which the data are to be grouped (e.g., logconc), and place this variable into the '**Factor(s):**' box by selecting **Add**.

4) Select the variable of interest (e.g., jdrywt), followed by **Add**, to insert the variable into the '**Dependent(s):**' box.

5) Select the box beside '**Save**' (bottom left-hand corner of the 'Analysis of Variance: Estimate Model' window) and scroll down the accompanying selections to choose **Residuals/Data**. Type in an appropriate file name within the adjacent empty box to save the residuals (e.g., anova1). Select **OK**. A graph of the data and the generate output will appear within the '**Output Pane**' tab. At this point, any outlier(s), based on the studentized residuals, will also be identified (refer to Section J.2.6 of this appendix for guidance on assessing outlier(s)).

6) Assess the assumptions of normality and homogeneity of the residuals as per Section J.2.4 using the data file that was created to save the Residuals/Data prior to conducting the ANOVA (i.e., anova1). After assessing normality and homogeneity of the residuals using Shapiro-Wilk's and Levene's tests, respectively, the following coding is used to examine the graphs of the residuals:

```
graph
use anova1
plot residual*logconc
plot residual*estimate
```