Environment and Climate Change Canada
Environnement et Changement climatique Canada

BIOLOGICAL TEST METHOD:

Test for Measuring Reproduction of Oribatid Mites Exposed to Contaminants in Soil

Canada

PDF:
Cat. No.: En83-10/1-61-2020E-PDF
ISBN: 978-0-660-35725-6

Environment and Climate Change Canada
Public Inquiries Centre
12th Floor, Fontaine Building
200 Sacré-Coeur Boulevard
Gatineau QC K1A 0H3
Telephone: 819-938-3860
Toll Free: 1-800-668-6767 (in Canada only)
Email: ec.enviroinfo.ec@canada.ca

Photo on Cover: *Oppia nitens,* Valerie Behan-Pelletier, Acarology Unit, Research Branch, Agriculture and Agri-Food Canada © Environment and Climate Change Canada

Aussi disponible en français

Biological Test Method:
Test for Measuring Reproduction of Oribatid Mites Exposed to Contaminants in Soil

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
Ottawa, Ontario

STB 1/RM/61
September 2020

Readers' Comments

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

Richard Scroggins, Chief
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario
K1A 0H3

General inquiries regarding this method can be addressed to:
ec.methodes-methods.ec@canada.ca

Review Notice

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

Abstract

This document provides detailed procedures, conditions, and guidance on preparing for and conducting a biological test for measuring soil toxicity using an oribatid mite, *Oppia nitens*. This is a 28-day test for effects on mite reproduction. The method is conducted as a static 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. Water and food (granulated dry yeast) are added to the test vessels during the test.

The test is conducted at a mean temperature of 20 ± 2°C in 30-mL glass shell vials (~2.6 cm inner diameter), or other suitable vessels, containing a measured volume of approximately 20 mL of soil or that which results in a ≥ 3 cm soil depth, at optimal moisture content. This test is initiated by placing 15 age-synchronized (adults, aged 8–10 days post-ecdysis to adult stage) test organisms into each replicate vessel containing test or clean (negative control or reference) soil. A minimum of five replicates are prepared for each treatment. At the end of the test, the live mites (adults and progeny) are extracted from the soil using heat-extraction, and the numbers in each replicate and treatment are determined. The mean of the replicates for each treatment is calculated and the percentage effect concentration estimated for inhibition of reproduction (e.g. ICp).

General or universal conditions and procedures are outlined for test preparation and performance. Additional conditions and procedures specific to the intended use of each test are stipulated. The biological test method described herein 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 commercial chemical(s) or test substance(s). Instructions and requirements are included for test facilities, sample collection, handling and storing samples, culturing 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 positive control replicates or a reference toxicity test.

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 and Climate Change Canada (previously Environment Canada) and are favoured:

- for use in Environment and Climate Change Canada environmental toxicity laboratories;
- for testing that is contracted out by Environment and Climate Change 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 environmental protection and management programs carried out by Environment and Climate Change 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 samples 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 and Climate Change Canada as part of this series.

Words defined in the Terminology section of this document are italicized when first used in the body of the report according to the definition.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Chemical Formulae

AES	atomic emission spectrophotometry
Al	aluminum
ANOVA	analysis of variance
$CaCl_2$	calcium chloride
$CaCO_3$	calcium carbonate
CCME	Canadian Council of Ministers of the Environment
C	carbon
Ca	calcium
CEC	cation exchange capacity
Cl	chlorine
cm	centimetre(s)
CV	coefficient of variation
°C	degree(s) Celsius
d	day(s)
DQO	data quality objective
g	gram(s)
h	hour(s)
H_3BO_3	boric acid
HPLC	high performance liquid chromatography
H_2O	water
ICAP	inductively coupled argon plasma
ICp	inhibiting concentration (for a specified percent effect; e.g., IC50)
K	potassium
kg	kilogram(s)
L	litre(s)
LC50	median lethal concentration
LED	light-emitting diode
LOEC	lowest-observed-effect concentration
m	metre(s)
M	mole(s) (concentration)
Mg	magnesium
mg	milligram(s)
mL	millilitre(s)
mm	millimetre(s)
mS	millisiemens
MW	molecular weight
n	sample size
N	nitrogen
Na	sodium
nm	nanometre(s)
NOEC	no-observed-effect concentration
OM	organic matter
P	probability
PAH	polycyclic aromatic hydrocarbon
QA/QC	quality assurance / quality control
®	Registered Trade Mark
SD	standard deviation
S	sulphur
s	second
sp.	species (singular)
spp.	species (plural)
TOC	total organic carbon
TM (^{TM})	Trade Mark
v:m	volume-to-mass
v:v	volume-to-volume
WHC	water-holding capacity
wt	weight
α	alpha, denotes Type I error
β	beta, denotes Type II error
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., survival/reproduction)
≈	approximately equal to
~	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 (mite) is a mite that is sexually mature. (See also *juvenile*.)

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

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, age-synchronized 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.

Ecdysis refers to the process of moulting or shedding an outer cuticular layer (i.e., exoskeleton). For oribatid mites, the shedding of exoskeleton stops with the final developmental stage (i.e., *adult*), as the exoskeleton in the *adult* form becomes hardened and sclerotized. For the purpose of this method, *ecdysis* refers to the emergence of an *adult* mite from the tritonymph stage, marked by the shedding of its exoskeleton and distinguishable by its body shape and size, as well as the colour of its integument (see Appendix E).

Ecological risk assessment (ERA) is the process of *risk* analyses and evaluation of the adverse *effects* of *contaminated* environmental media (e.g., air, *soil*, water) on non-human organisms with respect to the nature, extent and probability of the occurrence of these *effects* (ISO, 2005).

Electrical conductivity is a numerical expression of the ability of a 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. For this method, *electrical 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.

Hormesis is an observed stimulation of performance (e.g., reproduction) among test organisms, compared with the *control* organisms, at low *concentrations* in a *toxicity test*.

Instar refers to a stage of an insect or other arthropod between molts.

Juvenile (mite) is a mite that is sexually immature (i.e., larvae, protonymphs, deutonymphs, and tritonymphs). (See also *adult*.)

L, F, and H layers refer to the combined LFH layer of a *soil*. This is an organic layer that occurs on the surface of the mineral *soil*, and is usually composed of the accumulation of leaves, twigs, and woody *materials*. The components of the L (leaf) layer, which is at the top, are usually identifiable. The next layer down (F) is distinguished by the original *materials* being difficult to identify as a result of the initiation of decomposition, while the H layer is composed of decomposed organic *materials* that are indiscernible. The H layer may be intermixed with mineral particles from the mineral *soil* below.

Light-emitting diode (LED) is a type of light source. It is a semi-conductor diode which glows when a voltage is applied. LED differ from fluorescent and incandescent light sources in the mechanism used to generate light.

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 [µmol/(m^2 · 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, 2014). Approximate conversions between quantal flux and *lux*, however, are:

- for cool-white fluorescent light: 1 *lux* ≈ 0.014 µmol/(m^2 · s);
- for full-spectrum fluorescent light (e.g., Vita-Lux® by Duro-Test®): 1 *lux* ≈ 0.016 µmol/(m^2 · s); and
- for incandescent light: 1 *lux* ≈ 0.019 µmol/(m^2 · 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 *soil* samples for *toxicity*.

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 0−14, with 7 representing neutrality, numbers < 7 indicating increasingly greater acidic reactions, and numbers > 7 indicating increasingly basic or alkaline reactions.

Photoperiod is the duration of illumination and darkness within a 24-hour period.

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 sexually mature (*adult*) mites.

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.

Redox potential (also known as the oxidation-reduction potential) is a measure (in volts) of the affinity of a *substance* for electrons relative to hydrogen.

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 and Climate Change 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*.

Setae are slender, usually rigid, bristles, hairs, or spines distributed in characteristic patterns on the exoskeleton that function as sensory receptors or in locomotion.

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*. A *batch* might also refer to a single group of *O. nitens* received from a supplier or source outside the laboratory at a discrete time.

Bulk soil samples are *unconsolidated*, typically large (> 1 L) *point samples* that consist of more than one individual block of *soil* removed from one sample location by a sampling device, and therefore are *point samples*, not *composite samples* (see *point* and *composite samples*). *Bulk soil* samples are often collected to satisfy the large volume requirements for biological testing.

Cation exchange capacity is the sum total of exchangeable cations that a *soil* can adsorb. It is sometimes called total-exchange capacity, base-exchange capacity, or cation-adsorption capacity. It is expressed in milliequivalents per 100 grams of *soil* (or other adsorbing *material* such as clay) (AAFC, 1998).

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, or that might enter the environment through spillage, application, or discharge.

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* to the test organisms.

Composite sample(s) are *soil* samples consisting of *point* or *bulk* samples combined from two or more sample locations at a *site* (Crépin and Johnson, 1993).

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

Consolidated sample (see also *unconsolidated soil sample*) is synonymous with undisturbed sample and is a sample obtained from *soil* using a method designed to preserve the soil structure (ISO, 2005).

Contaminant is a *substance* or *material* that is present in a natural system, or present at an increased *concentration*, often because of some direct or indirect human activity. The term is frequently applied to *substances* or *materials* 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* or *substance*. The *control* is used as a check for the absence of measurable *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*.

Core sample is a sample of *soil* that has been collected using a corer.

Data quality objectives (DQOs) are pre-defined criteria for the quality of data generated or used in a particular study so as to ensure that the data are of acceptable quality to meet the needs for which they were collected.

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

Deionized 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 essential food and fibre; Hausenbuiller, 1985).

Horizon – see *soil horizon*.

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 *deionized* or *distilled water*, reverse-osmosis water, or dechlorinated tap water. Depending on study design and intent, a surface water or groundwater from the *site* might be used instead of *deionized* or *distilled water* for the hydration of each *test soil* (including *negative control soil*). (See also *test water, deionized 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* that could affect the survival or reproduction 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 reproduction 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*) and gives information about the state of health (i.e., quality) of the test individuals coming from a *culture*.

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 many types of *soil*, the following equation (from AESA, 2001) is suitable for estimating the total OM content of *soil* from *total organic carbon* (TOC) measurements: % OM = % TOC × 1.78; however, the relationship between TOC and OM is slightly different among *soils*, and therefore the *total organic carbon* content should also be determined by laboratory analysis. (See also *total organic carbon.*)

Point sample(s) are individual blocks of *soil* removed from one sample location by a sampling device (e.g., a soil *core*).

Positive control soil is *contaminated soil* that contains *concentrations* of one or more *contaminants* that adversely affect the reproduction of the test organisms using the biological test method 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*. *Reference soil* used in a test frequently exhibits physicochemical properties (e.g., *texture*, *organic matter* content, *total organic carbon* content, *pH*, and *electrical conductivity*) 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 multi-*concentration toxicity test* with a *reference toxicant* or a positive control *concentration* prepared using a *reference toxicant* is used 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 multi-concentration 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 and should be investigated as to the cause. A *reference toxicity test* with mites is performed as a *spiked-soil* test using a standard *chemical*.

Sampling location means a specific location, within a *site*, where the sample(s) of field-collected *soil* are obtained for *toxicity tests* and associated physicochemical analyses (and is considered the same as a sampling station).

Site means a delineated tract of land that is being used or considered as a study area, usually from the perspective of its being *contaminated* or potentially *contaminated* by human activity. A *reference site* is a *site* uninfluenced by the source(s) of contamination but within the general vicinity of the *sites* where samples of *test soil* are collected.

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 mites. 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 biological activities (e.g., microbial, invertebrate [including mite], and plant) and abiotic factors therein, and by anthropogenic activities.

Soil horizon is a layer of mineral or organic *soil material* approximately parallel to the land surface that has characteristics altered by processes of *soil* formation. It differs from adjacent *horizons* in properties such as colour, structure, *texture*, and consistency and in chemical, biological, or mineralogical composition.

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 mites 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 (are) 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 and Climate Change Canada documents typically do not call this *spiking*, but instead refer to the manipulation as "dilution," "amendment," or simply "addition." (See also *chemical-spiked soil* and *spiked soil*.)

Stock solution means a concentrated solution of the *substance(s)* to be tested, following 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* (*deionized*, *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* (e.g., *site soil*) that is *contaminated* or potentially so, or a *chemical-spiked soil* that is to be evaluated for *toxicity* to mites. Boreal and taiga *test soils* are collected as separate *soil horizons*. 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 mites and for other purposes associated with the biological test method (e.g., to hydrate samples of *test soil*). *Test water* must be *deionized* 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*. *Texture* in the context of this guidance document is described according to the Canadian System of Soil Classification (AAFC, 1998), not the Unified Soil Classification, the United States Soil Conservation Service Classification, or any other soil classification system used for soil science, engineering, or geology. 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, sandy clays).

Further distinction as to *texture* (e.g., "sandy clay," "silt loam," "loam") can be made based on the Canadian classification scheme using the relative amounts of percent sand, percent silt, and percent clay in the *soil* (AAFC, 1998).

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

Unconsolidated sample (see also *consolidated soil sample*) is synonymous with disturbed sample and is a sample obtained from *soil* without any attempt to preserve the soil structure (ISO, 2005).

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; water weight:dry soil weight) retained in a sample of *soil* that has been saturated with water.

Statistical and Toxicological Terms

A priori literally refers to something that is independent of experience. In the context of test design and statistics, *a priori* tests are ones that have been planned before the data were collected. Test objectives and test design would influence the decisions on which *a priori* tests to select.

Acute means within a short period (seconds, minutes, hours, or a few days) in relation to the lifespan of the test organism and is generally used to describe the length of a test or exposure duration.

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 7 or 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 springtails, plants or earthworms), 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 with the known potency of a standard preparation. *Toxicity test* is a more specific and preferred term for environmental studies.

Chronic toxicity refers to discernable adverse *effects* observed during or after relatively long-term exposure to one or more *contaminants*, which 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).

Effect, in *toxicology*, means a measurable biological change. The change could be structural, physiological, behavioural, etc. In a *toxicity test*, the biological change should be assessed against a background of measurements on organisms in *control* conditions. The statistical analysis generally considers the degrees of *effect* that are beyond the *control* measurements, and are therefore presumed to result from exposure to *toxic*

components of the *material* being tested.

Endpoint means the response(s) of the test organism that is measured (e.g., *adult* death, number of *progeny*), or the value(s) that characterize the results of a test (e.g., *LC50*, IC25).

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 logarithmically. 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 EC, 2005b). 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 EC, 2005b). 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 with the *control*, in a *quantitative* (continuous) biological measurement such as number of *progeny* produced by individuals at the end of the test (e.g., IC25 or IC50).

LC50 is the median *lethal* concentration, i.e., the *concentration* (e.g., % or mg/kg) of *substance(s)* or *material(s)* 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., 28-day *LC50*). Depending on the study objectives, an LCp other than *LC50* (e.g., an LC25) 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*.

Normality (or *normal distribution*) refers to a symmetric, bell-shaped array of observations. The array relates frequency of occurrence to the magnitude of the item being measured. In a *normal distribution*, most observations will cluster near the mean value, with progressively fewer observations toward the extremes of

the range of values. The normal distribution plays a central role in statistical theory because of its mathematical properties. It is also central in biological sciences because many biological phenomena follow the same pattern. Many statistical tests assume that data are normally distributed, and therefore it can be necessary to test whether that is true for a given set of data.

Power is, loosely, the probability of correctly concluding that there is a difference between the variables being tested. By definition, it is "the probability of rejecting the null hypothesis when it is in fact false and should be rejected." In effect, it is the opposite of making a *Type II error*, in which an investigator accepts the null hypothesis when there is actually a difference. The probability of making that *Type II error* is called *ß*, and *power* is represented by (1 - *ß*). *Power* cannot be directly and precisely set by the investigator, before doing a *toxicity test*. *Power* can be increased, however, by strengthening the *toxicity test* (more organisms, more *replicates*, etc.). Calculating *power* at the end of a test is rather complex, but *power* is related to *Minimum Significant Difference*, which can be estimated by standard procedures in many statistical tests that operate on *quantitative* data.

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. An example would be number of *progeny* produced at test end. Generally, *quantitative effects* are determined and expressed as measurements. (See also *quantal*.)

Replicate (*treatment*, *test vessel,* or *test unit*) refers to a single test vessel 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 vessel; therefore, any transfer of organisms or test *material* from one test vessel to another would invalidate a statistical analysis based on the replication (see Sections 5.1 and 5.6.1 herein, and Section 2.5 of EC, 2005b).

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

Residual, in the context of Section 4.8.1.1, 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* (or 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 resulting from exposure to the *concentration* or level of

contamination below that which 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 amounts at the right location (i.e., receptor/organ). *Toxic* is an adjective and, in some situations, a noun (usually found in the plural). In this context, *toxicant* is the better choice for the 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 result from exposure to either *lethal* or *sublethal concentrations* of *contaminants* in *soil*.

Toxicity test is a determination of the adverse *effect(s)* of a *substance* or *material* that results from exposure of a group of selected organisms of a particular species (e.g., *Oppia nitens*), 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* observed (*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*, *material*s, 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 safety limits 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 location*, 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*.)

Type I error, commonly designated as *α (alpha)*, occurs when an investigator rejects a null hypothesis that is true. In other words, the investigator concludes that there is a significant difference, when there is in fact none.

Type II error, commonly designated as *ß (beta),* occurs when an investigator fails to reject the null hypothesis when it is false (concludes that there is no significant difference, when there is in fact one).

Warning chart is a graph used to follow changes over time, in the *endpoints* for a *reference toxicant*. The date of the test or test number is on the horizontal axis. For multi-concentration tests, the effect-concentration is plotted on the vertical logarithmic scale, whereas for positive controls, the percent *effect* relative to the *control* is plotted on the vertical arithmetic scale.

Warning limit is plus or minus two standard deviations of the mean from tests with a *reference toxicant*. For multi-concentration tests, a *warning limit* is calculated logarithmically from a historical *geometric mean* of the *endpoints* (i.e., IC50), whereas for positive controls, a *warning limit* is calculated arithmetically from a historical mean of *endpoints* (% *effect* relative to *control*).

Acknowledgements

This document was prepared by Ryan Hennessy (Environment and Climate Change Canada [ECCC], Ottawa, ON) and Jennifer Miller (Miller Environmental Sciences Inc.), with introductory text largely written and supplied by Juliska Princz (ECCC, Ottawa, ON). Strategic direction and detailed review comments were supplied by Rick Scroggins and Juliska Princz (ECCC, Ottawa, ON). Thanks also to Leana Van der Vliet (ECCC, Ottawa, ON), who provided review comments on the test design and statistical analyses. This method was based on research conducted by ECCC's Soil Toxicology Laboratory and would not have been possible without the tremendous efforts of the Laboratory's staff: Juliska Princz, Heather Lemieux, Muriel Jatar, Myriam Malette, Ellyn Ritchie, Christopher Fraser, and Patrick Boyd. Work at the Soil Toxicology Laboratory was, in turn, based on guidance provided by Dr. Valerie Behan-Pelletier of Agriculture and Agri-Food Canada and Dr. Steven Siciliano of the University of Saskatchewan.

The interlaboratory studies undertaken to validate the test method were coordinated by Heather Lemieux and performed by the following participating laboratories: Eurofins Agroscience Services EcoTox GmbH (Niefern-Öschelbronn, Germany), ECT Oekotoxikologie GmbH (Flörsheim am Main, Germany), Nautilus Environmental (Calgary, Canada), Exova (Edmonton, Canada), University of Coimbra (Coimbra, Portugal), CREAF - Universitat Autònoma de Barcelona (Barcelona, Spain), Vrije Universiteit Amsterdam (Amsterdam, Netherlands), University of Saskatchewan (Saskatoon, Canada), AGAT Laboratories (Quebec City, Canada), and ECCC (Ottawa, Canada). We gratefully acknowledge the contributions of all of the inter-laboratory participants: Marlene Fuchs, Stephanie Kirschner, and Thomas Moser from Eurofins Agroscience Services EcoTox GmbH; Cornelia Bandow, Jörg Römbke, Adam Sceffczk, and Lilli Senn from ECT Oekotoxikologie GmbH; Holly Lopez from Nautilus Environmental; Darlene Lintott from Exova; Ana Daniela Alves, Tiago Natal da Luz, and Carla Pereira from the University of Coimbra; Xavier Domene and Alba Llovet from CREAF – Universitat Autònoma de Barcelona; Kees van Gestel and Rudo Verweij from Vrije Universiteit Amsterdam; Richard Nhan, Kayode Jegede, and Steven Siciliano from the University of Saskatchewan; Virginie Bérubé, Mélanie Picard, and Pierre Yves Robidoux from AGAT Laboratories; and Heather Lemieux, Juliska Princz, and Rick Scroggins from ECCC (Ottawa, Canada).

We gratefully acknowledge the many useful review comments provided by our review panel, including Thomas Moser from Eurofins Agroscience Services EcoTox GmbH (Niefern-Öschelbronn, Germany); Lilli Senn and Jörg Römbke from ECT Oekotoxikologie GmbH (Flörsheim am Main, Germany); Carla Sofia Pereira, Mathieu Renaud, and Tiago Natal da Luz from University of Coimbra (Coimbra, Portugal); Xavier Domene from CREAF – Universitat Autònoma de Barcelona (Barcelona, Spain); Kees van Gestel from Vrije Universiteit Amsterdam (Amsterdam, Netherlands); and Kayode Jegede, John Owojori, and Steven Siciliano from the University of Saskatchewan (Saskatoon, Canada).

The ongoing support of members of ECCC's Regional Environmental Testing Laboratories (Appendix B) and the Inter-Governmental Ecotoxicological Testing Group (Appendix C) is also acknowledged. Special thanks to members of ECCC's Soil Toxicology Laboratory (past and present) for providing photographs for inclusion in this test method document.

Section 1

Introduction

1.1 Background

The Method Development and Applications Unit (MDAU) of Environment and Climate Change Canada (ECCC; previously 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, MDAU, the Canadian Association of Petroleum Producers (CAPP), and the federal Program for Energy Research and Development (PERD) initiated 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. The goal was to develop biological test methods applicable to diverse types of Canadian *soils* using terrestrial species that were representative of Canadian soil ecosystems. There have been two comprehensive reviews of existing biological test methods used internationally to evaluate the toxicity of *contaminants* to soil invertebrates (Bonnell Environmental Consulting, 1994; Römbke *et al.*, 2006).

Four standardized soil toxicity test methods have been published by Environment Canada: i) Tests for Toxicity of Contaminated Soil to Earthworms (*Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris*), EPS 1/RM/43 (EC, 2004a); ii) Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil, EPS/1/RM/45 (EC, 2005a, amended 2007); iii) Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil, EPS 1/RM/47 – 2nd edition (EC, 2014a); and iv) Test for Growth in Contaminated Soil using Terrestrial Plants Native to the Boreal Region, EPS 1/RM/56 (EC, 2013a).

Universal procedures for preparing and conducting soil *toxicity tests* using the oribatid mite *Oppia nitens* are described herein. Guidance is also provided for specific sets of conditions and procedures that are required or recommended when using this biological test method 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 biological *endpoint* for this method is reproductive success measured at the end of the test.[1]

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).

[1] ECCC soil methods have traditionally been designed to capture both lethal (i.e., survival) and sublethal (e.g., reproduction, growth) endpoints (EC, 2004a, 2014a). For this oribatid mite method, however, reproductive success is the preferred endpoint as it has been found to be significantly more sensitive than adult survival (e.g., EC, 2010, 2013b; Princz *et al*., 2010, 2012, 2018; Li. *et al*., 2018; Gainer *et al*., 2018). The difference in magnitude between the two endpoints is often large enough to make it very challenging to determine an appropriate concentration series that would capture both endpoints simultaneously. Recent research sheds light on species' traits, such as the long lifespan and small clutch size exhibited by oribatid mites, which may be linked to this observed sensitivity for the reproduction endpoint (Gainer *et al*., 2018). Although the focus of this test is on reproduction as an endpoint, the test does not preclude the derivation of an *LC50* if warranted (see Section 4.7). Adult survival data are collected and recorded at the end of the test for all soils to help with the assessment of reproduction (e.g., confirm adult survival, identify outliers) and for the assessment of test validity (see Sections 4.4 and 4.7).

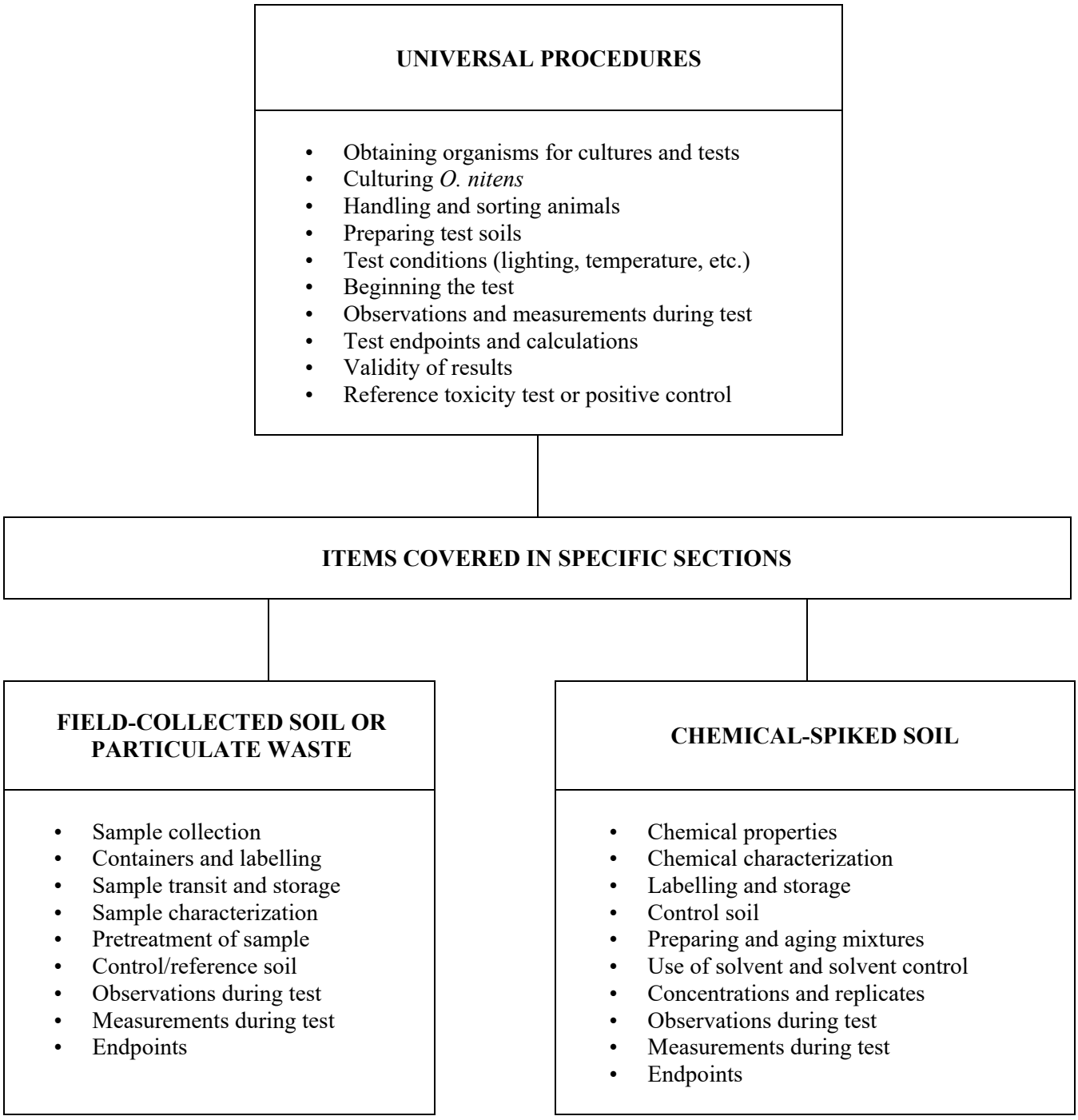


Figure 1 **Considerations for preparing and performing soil toxicity tests using mites and various types of test materials or substances**

This biological test method is intended for use in evaluating the *sublethal* toxicity of samples of material such as the following:

- field-collected soil from agricultural or non-agronomic regions that are contaminated or potentially contaminated;
- soils under consideration for removal and disposal or *remediation* treatment;
- soils that have undergone remediation treatment;
- dredged material destined or under consideration for land disposal after dewatering;
- industrial or municipal sludge and similar particulate wastes that might be deposited on land; and
- *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 this biological test method, an attempt has been made to balance scientific, practical, and cost considerations, and to ensure that the results will be sufficiently precise for the majority of 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 this and other biological test methods, and on the interpretation and application of endpoint data for soil toxicity, the reader should consult Sections 4.12, 5.5, and 5.6.4 of Environment Canada's *Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology* (EC, 1999). In-depth direction on the use of statistics in determining effect endpoints in ecotoxicology testing is available in Environment Canada's Guidance Document on Statistical Methods for Environmental Toxicity Tests (EC, 2005b).

1.2 Identification, Distribution, and Life History of Oppia nitens *(C.L. Koch)*

The test species to be used for the biological test method described herein (*Oppia nitens*) belongs to the suborder Oribatida. Oribatid (Oribatida = Cryptostigmata) mites are members of the class Arachnida (phylum Arthropoda, subphylum Chelicerata), commonly known for spiders, scorpions, and mites. Mites belong to the infraclass Acari, which is the most diverse and abundant of all arachnid infraclasses (Walter *et al.*, 1996). Fossilized mites have been found dating back nearly 400 million years, making them some of the most ancient terrestrial animals (Norton *et al.*, 1988). The Acari are divided into three superorders: the Parasitiformes (which includes ticks), the Opilioacariformes, and the Acariformes. Oribatids are members of the Acariformes or "mite-like" mites. Acariformes are separated from the other two superorders primarily by their modified sensory *setae* and the presence of optically active chitin (actinochitin) (Proctor, 1998).

Oribatid mites are free-living and typically the most abundant and diverse microarthropod present within organic *horizons* (Crossley and Bohnsack, 1960; Walter, 1985; Behan-Pelletier, 1999; Heneghan *et al*., 1999), with densities ranging from 200,000 (Maraun and Scheu, 2000) to approximately 500,000 individuals per m^2 (Behan *et al*., 1978). Of particular importance to Canada, these mites are often the most numerically dominant arthropods in the cold ecosystems of the Northern Hemisphere (Behan, 1978). Oribatid mites significantly contribute to nutrient (e.g., nitrogen) immobilization and mineralization (Singh *et al*., 1996; Hansen, 2000; Johnston and Crossley, 2002), and soil formation (Coleman *et al*., 2004).

Oribatid mites are typically medium to dark brown, with some exceptions, and are on average 300 to 700 µm long, but can range in size from 150 to 1500 µm (Behan-Pelletier, 1999). Growth and maturity are characterized by six post-embryonic developmental stages (an inactive prelarva, an active larva, protonymph, deutonymph, tritonymph, and *adult*) whereby growth is accomplished with the shedding of exoskeleton. Some aspects of development are unique in that the legs of

subsequent *instars* are formed within the body, rather than within the hull of previous instars' legs (Proctor, 1998). The moult (*ecdysis*) involves a pre-ecdysial (development of integument) resting stage, characterized by immobility for a prolonged period of time and, depending on the species, can occupy up to one third of a mite's total lifespan (Luxton, 1981). The shedding of exoskeleton stops with the final developmental stage, as the exoskeleton within the adult form becomes hardened and sclerotized. Melanization is also typical, manifesting in a medium to dark brown colouration.

The life-history traits of oribatid mites are generally described as "*k*-selected," characterized by low metabolism, slow development, and low fecundity. However, given the abundance and diversity of oribatid species, they exhibit broad and opportunistic feeding habits (Norton, 1985), are able to disperse (albeit slowly, and primarily as adults) and colonize different soil habitats and horizons, and occupy varied trophic levels (Behan-Pelletier, 1999). The life-span of oribatid mites (i.e., egg to adult) varies between species, ranging from a few weeks to several months, to a year in temperate soils (Norton, 1985), and ranging from three to as long as seven years in cooler climates (Cannon and Block, 1988; Webb, 1989). However, time to maturity has mainly been documented under laboratory conditions, of which temperature can affect the duration of maturation (Norton, 1994). Also of note is that some species have the ability for super-cooling during freezing temperatures, thus also possibly extending longevity, but also allowing for winter dormancy (Cannon and Block, 1988).

Oppia nitens C.L. Koch 1836 is a member of the largest oribatid family, Brachypylina:

- superorder Acariformes;
- order Sarcoptiformes;
- suborder Oribatida;
- infraorder Brachypylina;
- superfamily Oppioidea;
- family Oppiidae;
- subfamily Oppiinae.

Adult *O. nitens* are approximately 510 µm in length and 290 µm in breadth (Michael, 1884), while larvae are ~200 µm in length and ~105 µm in breadth, and tritonymphs are ~372 µm in length and ~195 µm in breadth (Seniczak, 1975). During *juvenile* stages, the mites are white/translucent in colour but transition through golden hues to a rich chestnut brown within a week of reaching adulthood (Appendix E). Development time appears to be dependent on environmental conditions, with Princz (2014) finding that the species matures within four weeks of hatching at 20°C to 23°C, while Sengbusch and Sengbusch (1970) observed a developmental period of 45 to 46 days at 20°C, with females laying eggs only three months after hatching; Stefaniak and Seniczak (1981) observed a developmental period of 28 days at 23°C, and Yu *et al*. (1997) observed a developmental period of 2–3 weeks at 23°C with a fairly high humidity of 85 ± 5%. There is conflicting information regarding reproductive mode (i.e., sexual or parthenogenetic); however, observations of spermatophores under laboratory conditions (Stefaniak and Seniczak, 1981) are suggestive of sexual reproduction. The literature contains no documented sexual dimorphism exhibited within the species (i.e., males do not differ in size from females), and sexual dimorphism has not been observed to date in present laboratory cultures. *Oppia nitens* eggs are oval and whitish, with a smooth surface, and sizes ranging from 90 to 150 µm (Seniczak, 1975); eggs hatch within about one week of oviposition (Princz, 2014).

Diagnostic features for identifying *O. nitens* have historically included the number and type of setae but more recently, emphasis is being placed on DNA-based taxonomic identification (i.e., barcoding). The International Organization for Standardization (ISO) has published a standardized procedure for the identification of ecotoxicological test species using DNA barcoding (ISO, 2019a). For *O. nitens*, sequencing of the 5' region of mitochrondrial Cytochrome Oxidase Subunit 1 from several specimens is available for comparison on the International Barcode of Life data portal: http://www.boldsystems.org. The Barcode of Life Data Systems (BOLD) is one of several international databases that allows access to, and provides a platform for, analysis of DNA barcode sequences.

Oppia nitens has been documented as a polyphagous fungivore but does show some selective feeding preferences (Seniczak and Stefaniak, 1978). Singh *et al*. (1996) observed a strong preference for ground leaf litter mixed with dried mushrooms, with a

moderate preference for leaf litter or mushrooms alone, and very little preference for granulated yeast; however, *O. nitens* has been successfully reared on granulated yeast alone (Princz, 2014). Other feeding substrates in nature include lichens, raw humus, or carrion (dead and rotting tissue), and the species may be predatory and cannibalistic (Seniczak, 1975; Seniczak and Stefaniak, 1978).

1.3 Historical Use of Mites in Toxicity Tests

The development of biological test methods for soil toxicity testing has lagged behind that for other media (e.g., water and sediment) (Bonnell Environmental Consulting, 1994). That delay was partially due to the fact that researchers and regulators had focused on the aquatic environment. Soil systems are more complex than aquatic systems, with many problems inherent in their lack of homogeneity. The variety of exposure routes available to investigators (e.g., via pore water, soil vapours, or direct contact with soil particles), coupled with the high cost of running soil toxicity tests, in the past have led investigators to rely on extrapolations from aquatic test methods to soil-based exposures (Bonnell Environmental Consulting, 1994).

Assessment of soil quality during the 1970s and earlier primarily involved the evaluation of the physicochemical properties of soil. It was not until the 1980s that the initial use of standardized biological test methods for measuring soil toxicity emerged from agencies responsible for pesticide registration and application (e.g., the United States Environmental Protection Agency [USEPA], and the Office of Pesticides Programs [Holst and Ellanger, 1982]; the Organization for Economic Cooperation and Development [OECD], 1984a, b).

The toxicity of *site soils* became a "new" concern in the mid-1980s, and regulatory programs such as SUPERFUND in the United States, and the National Contaminated Sites Remediation Program (NCSRP) in Canada, were established to address the urgent need for guidance on the assessment and remediation of high-priority contaminated *sites*. Under the NCSRP, a review of existing whole-organism *bioassays* for soil, freshwater sediment, and fresh water (Keddy *et al.,* 1995) was conducted to lead to the establishment of a suite of tests that could be used immediately for contaminated-site assessment in Canada (Bonnell Environmental Consulting, 1994). Keddy *et al.* (1995) concluded that most of the existing methods or procedures for measuring the toxicity of samples of soil from contaminated sites were inadequate for proper ecotoxicological assessment, and recommended that attempts be made to develop a suite of standardized biological test methods for soil that used test species and conditions applicable to Canadian soil ecosystems. The Canadian Council of Ministers of the Environment (CCME) published a framework for *ecological risk assessment* (ERA) in 1994 (CCME, 1994), which had a subsequent impact on the management of contaminated sites (CCME, 1996, 1997). The ERA approach, which relied on the results of single-species toxicity tests, led to the need to develop reliable, reproducible, and realistic soil toxicity tests with ecologically relevant terrestrial test species for the assessment of contaminated site soils (Bonnell Environmental Consulting, 1994). In the late 1990s, biological assessments in the form of toxicity testing were becoming a useful complement to chemical analyses, especially when applied to site-specific risk assessments.

Oribatid mites have served as bioindicators of environmental disturbances (Lebrun and van Straalen, 1995; Behan-Pelletier, 1999; Gergócs and Hufnagel, 2009), responding to changes in soil quality, and demonstrating susceptibility to contaminants (e.g., metals, pesticides, petroleum hydrocarbons) in soil (as reviewed in Princz, 2014). Relative to other Acari and soil arthropods, oribatid mite populations have been shown to be more sensitive to soil disturbances such as pollutant impacts, soil amendments, and forestry practices (Al-Assiuty *et al*., 2000; Battigelli *et al*., 2004; Minor and Norton, 2004). Their life-history characteristics (e.g., low metabolism, slow development, low fecundity) and slow dispersal capabilities limit their ability to adapt to short-term disturbances, leading to declined populations, which can in turn, be detected as a sign of environmental degradation (Lebrun and van Straalen, 1995). The use of oribatid mites for ecotoxicological testing has been reviewed (Lebrun and van Straalen, 1995; Hugier *et al*., 2015), demonstrating a recent rise in their applicability in the assessment of both contaminated site soils and chemically spiked soils,

evaluating endpoints such as lethality, avoidance, and reproduction.

Initial attempts at method standardization began with van Gestel and Doornekamp (1998) developing a mortality and reproduction soil assay using the parthenogenetic species *Platynothrus peltifer* (C.L. Koch, 1839). The test methods included both dietary and soil exposure studies, whereby the soil exposure studies demonstrated greater sensitivity to selected *toxicants* (copper and sodium salt of linear alkyl benzene sulphonates [LAS]). The results of this research demonstrated the need for a soil exposure system to account for multiple exposure pathways. However, standardization of a test method using *P. peltifer* was limited by the need to use field-collected specimens because of difficulties in the establishment and maintenance of laboratory cultures. Although the tests were effective at discerning effects on reproduction after exposure to contaminated soil, the tests were compromised by high adult mortality, and were lengthy (e.g., 6 to 12 weeks for reproduction endpoint) to accommodate the species' long development cycle (e.g., > 150 d to maturity) (van Gestel and Doornekamp, 1998). Although prevalent in boreal and arctic ecosystems, difficulties associated with culturing and testing also led to the recommendation for exclusion of this species from soil toxicity test method development (Römbke *et al*., 2006).

Additional studies have incorporated the use of *Archegozetes longisetosus* (Aoki 1965) as a laboratory test species (Seniczak and Seniczak, 2002; Köhler *et al*., 2005; Seniczak, 2006; Seniczak *et al*., 2006, 2009; Heethoff *et al*., 2007) as this species is parthenogenetic, easily cultured under laboratory conditions, and characterized by a short generation time with relatively high fecundity (Heethoff *et al*., 2007). However, this species is limited to a pan-tropical distribution; therefore, the relevance of this species to non-tropical habitats is questionable, particularly when assessing soils from boreal and northern regions. For other mite species, much research has been conducted to standardize *Hypoaspis* (*Geolaelaps*) *aculeifer* as a standard test species (OECD, 2008; Smit *et al*., 2012; ISO 2019b); however, this species occupies a higher trophic level as a result of its predatory habits, and does not represent the ecological niche that oribatid mites occupy.

Although research on metals dominates toxicological studies associated with oribatids, researchers have also evaluated the effect of pesticides in laboratory and field settings. In general, pesticide toxicity is substance- and species-specific (e.g., Al-Assiuty and Khalil, 1995; Cortet *et al*., 2002; Prinzing *et al*., 2002), with some species negatively affected [e.g., azadirachtin (Stark, 1992); 2,4,6-trinitrotoluene and p-nitrophenol (Parmelee *et al*., 1993)], and others remaining unperturbed [e.g., chlorpyrifos (Stark, 1992; Michereff-Filho *et al*., 2004); endosulfan (Osler *et al*., 2001); zinc-manganese ethylene-bis-dithiocarbamate (Adamski *et al*., 2007); and neonicotinoids (de Lima e Silva *et al*., 2017)].

With regard to hydrocarbon contamination, oribatid mites demonstrate some degree of sensitivity to petroleum hydrocarbon contaminated soil, as well as smaller-ring polycyclic aromatic hydrocarbons (PAHs), with no effect associated with five-ring PAH compounds, such as benzo[a]pyrene and creosote (Blakely *et al*., 2002; Owojori and Siciliano, 2012; Princz *et al.*, 2012). Similar results were found when *H. aculeifer* were exposed to benzo[a]pyrene for three weeks, in that no effects on adult survival or reproduction were observed at *concentrations* up to their highest test concentration of 947 mg/kg dry soil (Sverdrup *et al*., 2007). However, other researchers have found PAH contamination associated with a decline in the abundance of Acari in soil (Erstfeld and Snow-Ashbrook, 1999).

The potential of *O. nitens* as a test species has only been explored relatively recently. Yu *et al*. (1997) explored the toxicity of *Bacillus thuringiensis* toxins present in transgenic cotton and potatoes to *O. nitens*. The effect of dietary exposure only was evaluated on adult survival and reproduction, and no adverse effects were detected. These authors used cadmium as a positive control, but did not report the effects of the metal on either adult survival or reproduction. Princz *et al*. (2010) report success using age-synchronized cultures of *O. nitens* for mortality and reproduction testing in field-collected soils, and provided guidance on their use in soil toxicity testing. However, the authors cautioned that soil *organic matter* content might serve as a limiting factor for their use in highly mineral soils. Since this time, additional research has demonstrated the

applicability of using *O. nitens* in the assessment of contaminated site soils (e.g., Princz *et al*., 2012) and chemically spiked soils (e.g., Owojori and Siciliano, 2012; de Lima e Silva *et al*., 2017; Gainer *et al*., 2018, 2019; Li *et al*., 2018; Princz *et al*., 2018), warranting their inclusion as part of a test battery approach to the risk assessment of contaminated soils. A recent review by Huguier *et al*. (2015) cautions that *O. nitens* are in general less or as sensitive to anthropogenic contamination as other soil invertebrates, but echoes the sentiment that mites represent communities that cannot be omitted from environmental hazard assessment. More recently, the ISO experts of technical Committee TC-190: Soil Quality have agreed to proceed with the development of a standardized procedure for measuring the inhibition of reproduction in oribatid mites *(Oppia nitens*) exposed to contaminants in soil, led by Canadian experts (ISO 2019c).

Section 2

Test Organisms

2.1 Species and Life Stage

The biological method described herein must be performed using laboratory-cultured *Oppia nitens* C.L. Koch 1836. The identification, distribution, and life history of *O. nitens* is summarized in Section 1.2. Species identification must be confirmed and documented[2] upon establishment of a new *culture*, and/or with each new *batch* of *O. nitens* introduced to the laboratory culture (Römbke *et al*., 2016). Cultures of *O. nitens* held for a prolonged period at a testing laboratory should be identified to species at least once every two years. Species identification may be made using the distinguishing taxonomic features described and illustrated in taxonomic keys by qualified personnel experienced with identifying oribatid mites, or using DNA-based taxonomic identification (i.e., barcoding) (ISO, 2019a). The soil toxicity test described herein must be started using age-synchronized adult *O. nitens* that are collected over a ≤ 3-day period and aged 8 to 10 days post-ecdysis to adult form (see Section 2.3.8).

2.2 Source

Laboratory-cultured mites (see Section 2.3) must be used as the source of the test organisms. Sources of *Oppia nitens* for establishing laboratory cultures may be government or private laboratories that are culturing this species of mite for soil toxicity tests.[3]

Breeding stock of *O. nitens* can be obtained by contacting the following Canadian source:
Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario K1A 0H3
Email: ec.méthodes-methods.ec@canada.ca,

and the following International source:
Dr. Cornelius (Kees) A.M. van Gestel
Department of Ecological Science
Vrije Universiteit Amsterdam
De Boelelaan 1085
1081 HV Amsterdam
The Netherlands
Email: kees.van.gestel@vu.nl

All mites used in a soil toxicity test must be derived from the same population. Mites to be used as a source of breeding stock should be transported to the laboratory using a portion of the soil or other substrate to which they are adapted. Breeding stocks are best transported as a mixed-age culture in small containers with the plaster of Paris substrate described in Section 2.3.5[4] or in a small container of soil. Additional quantities of this substrate might be obtained for *acclimation* or culturing purposes, depending on culturing conditions and requirements (Section 2.3). Shipping and transport containers should be insulated to minimize changes in temperature during transit, and the temperature

[2] Acceptable forms of documentation include identification of laboratory specimens by a qualified taxonomist, and identification of laboratory specimens by molecular analysis (such as DNA barcoding).

[3] Investigators might wish to use progeny produced from organisms that occupied a particular locale. Accordingly, cultures may be established using wild populations or may be genetically enhanced by introducing breeding stock from different sources. If animals are obtained from a wild population, their taxonomy must be confirmed and they or their progeny should be evaluated for sensitivity to reference toxicant(s) before being used in toxicity tests. 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.

[4] The plaster of Paris substrate might loosen from the bottom of the container during transportation; therefore, steps should be taken to prevent the mites from being crushed between the loosened substrate and the sides of the container. The container should be sealed with Parafilm® to prevent moisture loss until receipt, after which the Parafilm® may be removed.

should be maintained at ~20°C. 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 in transit.

Upon arrival at the laboratory, organisms may be held in the substrate (i.e., soil or plaster of Paris) used in transit while temperature adjustments are made, or they may be transferred to new culturing substrate (Section 2.3.5). If the nature (including the *texture* and *moisture content*) of the substrate in which mites were initially held (e.g., by a supplier) or transported differs markedly from that in which they are to be cultured (Section 2.3.5), it is prudent to adapt the mites to a new substrate over several days. The impact of the transition can be lessened by adding some of the transport substrate to the culture substrate.

Soil temperature should be adjusted gradually (e.g., ≤ 3°C per day) to the temperature to be used during culturing (Section 2.3.4). Guidance on handling mites given in Section 2.3.7 should be followed when transferring organisms from an outside source to culture vessels (Section 2.3.2). Other conditions during this interim holding period for acclimation of breeding stock or for acclimation of test organisms to laboratory conditions should be as similar as possible to those used for maintaining cultures (Section 2.3).

2.3 Culturing of Oppia nitens

2.3.1 General

General guidance and recommendations for culturing *Oppia nitens* 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,"* explicit directions regarding many aspects of culturing, including the choice of culture vessel, number of organisms per vessel, 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[5] 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[6] and behave and feed normally (see Section 2.3.9). Additionally, those used as *controls* in the test must meet all criteria for a valid toxicity test (see Section 4.4). The acceptability of the culture is also demonstrated by ongoing *reference toxicity tests* or positive controls using a *reference toxicant* (see Section 4.9). If a culture of organisms fails to meet these criteria, its cause should be investigated. Care must be taken to ensure that each culture is not contaminated with other similar species (i.e., mixed with different mite or invertebrate species). Periodic (e.g., every two years) taxonomic checks of the laboratory's cultures are recommended (see Section 2.1).

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 *O. nitens*. For this purpose, intra-laboratory *precision*, expressed as a *coefficient of variation* for the respective IC50 data, should be determined by performing five or more full length tests (i.e., 28-d duration) 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 4.9).

When performing soil toxicity tests with *O. nitens*, consistency must be demonstrated either through the inclusion of a positive control concentration with each *definitive* test (Section 4.9) or through reference toxicity tests, which must be conducted a minimum

---

[5] Performance-based indices include those related to the survival and condition of cultured organisms 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 (Section 4.4), and those related to the performance of groups of animals in a positive control concentration run concurrently with each definitive test or in reference toxicity tests (Section 4.9).

[6] ECCC labs have observed that adult mites which lack the dark pigmentation typical of *O. nitens* (i.e., opaque with a "milky" appearance) cannot reproduce and are therefore not suitable for testing (see Section 2.3.9). Refer to Appendix E for a photographic example.

of twice per year with the laboratory's cultures, using the conditions and procedures outlined in Section 4.9. Additionally, the performance of any cultures that have been established recently using new breeding stock (Section 2.2) should be checked with a reference toxicity test or positive control, and the results determined to be acceptable (see Sections 2.3.9 and 4.9) before these cultures are used to provide test organisms.

Cultures of *O. nitens* should be observed frequently (e.g., once or twice per week). Ideally, records should be maintained documenting:

- the date a culture is started with adults;
- the dates of substrate renewal;
- the feeding and watering regime (including type and quantity added on each occasion);
- facility and substrate quality (e.g., air temperature, *photoperiod* and light quality, *pH* of substrate); and
- observations of culture health (e.g., behaviour and appearance of mites in culture, reproductive rates, presence of organisms of difference ages, odour of substrate, location of mites in the vessel, amount of uneaten food in vessel, presence of fungi or organisms other than *O. nitens*).

A checklist of required and recommended conditions and procedures for culturing *O. nitens* to generate organisms for use in soil toxicity tests is given in Table 1.

2.3.2 Facilities and Apparatus

Mites should be cultured in a controlled-temperature laboratory facility. Equipment for temperature control (i.e., an incubator or a room with constant temperature) should be adequate to maintain temperature within the recommended 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 condensation).

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

A variety of culture vessels, such as plastic trays or breeding boxes, are suitable for culturing *O. nitens*. The sides and/or lid may be translucent or transparent, to enable light to come into contact with the surface of the culturing substrate (see Section 2.3.3), but this is not a requirement. Each vessel should have a lid, which may be solid, to minimize drying of the surface substrate and the *risk* of contamination, or perforated (e.g., with holes covered with fibreglass mesh or Nitex™ screening) to allow air exchange. The use of culture vessels 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 and juvabiones).

The choice of size and numbers of culture vessels required might be influenced by the number of adult mites required by the testing facility for one or more series of soil toxicity tests. Each culture vessel should accommodate a minimum depth of 1 cm of soil or plaster of Paris substrate, or a combination of both (i.e., ≥ 1 cm layer of plaster of Paris substrate covered with a ≥ 1 cm layer of soil).

2.3.3 Lighting

Cultures of *O. nitens* can be cultured with incandescent, fluorescent, or *light-emitting diode* (LED) light and a regulated photoperiod (e.g., 16 h light:8 h dark or 12 h light:12 h dark). Light intensity adjacent to the top of the culture vessels should range within 400−800 *lux*. This range is equivalent to a quantal flux of 5.6−11.2 µmol/($m^2 \cdot s$) for cool-white fluorescent, 6.4−12.8 µmol/($m^2 \cdot s$) for full-spectrum fluorescent, or 7.6−15.2 µmol/($m^2 \cdot s$) for incandescent. The lights should be positioned sufficiently far from the culture vessels to prevent evaporation caused by heat buildup.

Table 1 Checklist of required and recommended conditions and procedures for culturing *Oppia nitens* to provide test organisms for use in soil toxicity tests

Source of breeding stock for culture	− mixed-age culture obtained from a government or private culture; identification to species confirmed
Acclimation	− gradually, for temperature (recommend ≤ 3°C/day) and substrate differences upon arrival
Culture vessels	− plastic trays or breeding boxes, covered with solid or perforated lids; sides and/or lid may be transparent or translucent to enable light to come into contact with the surface of the culturing substrate
Air temperature	− daily average, 20 ± 2°C; instantaneous, 20 ± 3°C
Lighting	− incandescent, fluorescent, or LED; intensity of 400−800 lux at the surface of the culture vessel; fixed photoperiod (e.g., 16 h L:8 h D or 12 h L:12 h D)
Type of substrate	− ≥ 1 cm of soil rich in organic matter; or ≥ 1 cm 8:1 mixture of plaster of Paris and activated charcoal with plaster of Paris caps or pie-shaped pieces of filter paper coated with plaster of Paris placed on the surface of the substrate; alternatively, ≥ 1 cm plaster of Paris substrate covered with ≥ 1 cm soil rich in organic matter
Hydration of substrate	− hydrated with test water; moisture content sufficient to keep surface of substrate moist but with no standing water on the surface of the plaster of Paris culture substrate or in the bottom of the soil culture vessels
pH of substrate	− 6.0−7.5
Renewal of substrate	− as required, and at least once every 4–6 months; transfer mites to fresh culture substrate; mix organisms between culture vessels
Monitoring culture	− air temperature of culture facility monitored weekly; pH measured when new batch of soil or new plaster of Paris substrate is prepared
Maintaining culture	− aerate vessels at least once/week (minimum twice/week recommended); moisture level of substrate observed for each culture vessel at time of aeration; spray surface or add several drops of test water to maintain humidity; record condition of culture; maintain loading density at ~5 to 15 adult mites/cm^2 for plaster of Paris substrate; maintain soil cultures with organisms of differing ages moving actively over the substrate surface or clustered around piles of yeast following a feeding event; addition of organic matter to substrate might increase reproduction
Feeding	− granulated dry yeast (e.g., Fleischmann's™), divided into several piles or sprinkled onto substrate surface; feed twice/week

Maintenance of age-synchronized cultures	–	newly emerged adult mites are selected from plaster of Paris substrate over a period of ≤ 3 days based on the colour of their integument; selected organisms are separated from the main culture; organisms held in soil are extracted from the soil cultures with heat and then transferred to a plaster of Paris substrate for selection; organisms selected for age-synchronization are maintained on plaster of Paris as per main culture (e.g., fed and aerated at least twice weekly)
Age for test	–	collected as newly emerged (ecdysis) adults (lightly coloured integument) over a period of ≤ 3 days and then aged for 8 to 10 days before test start
Indices of culture health	–	considered healthy if (1) organisms of differing ages are moving actively over the surface of the substrate, and there is a low incidence (i.e., ≤ 10%) of non-pigmented ("milky") mites (2) results for reference toxicity tests or positive controls using mites from the culture fall within historical warning limits; mites that appear to be injured or are otherwise unhealthy (e.g., milky) should be discarded and must not be used for testing; reproduction data from negative control soils are monitored

[*] **The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.**

2.3.4 Temperature

The organisms should be cultured in a facility with an air temperature of 20 ± 2°C as a daily average. Additionally, the instantaneous temperature of the facility should be 20 ± 3°C.

2.3.5 Culturing Substrate

The choice of substrate for culturing *O. nitens* is left to the discretion and experience of laboratory personnel; however, the following culture substrates are proven and recommended.

In Environment and Climate Change Canada's experience, *O. nitens* cultures demonstrate higher survival and reproductive rates when maintained in a soil substrate relative to those maintained on plaster of Paris (ECCC, 2018). The substrate ECCC recommends for culturing earthworms (Section 2.3.5 of EC, 2004a) has been used successfully for culturing mites; however, other soil types might be suitable as well. Field soils rich in organic matter (which have been demonstrated to be uncontaminated – i.e., no previous application of pesticides, etc.) have proven to be suitable for culturing substrates for *O. nitens* as well and can also be used for maintaining mass or back-up cultures in the laboratory. ECCC laboratories have had the most success raising mites in sandy soils with high leaf litter / organic matter content (ECCC, 2018). A minimum soil depth of 1 cm is recommended for culturing. When organisms are needed for testing, they are extracted from the soil cultures with heat and then transferred to a plaster of Paris substrate for easier observation, handling, and age-synchronization.

Alternatively, a substrate consisting of 8 parts plaster of Paris (Stucco)[7] and 1 part charcoal (e.g., analytical-grade activated charcoal 375 µm mesh; e.g., Fisher cat #35-474) has been recommended for the culturing of soil invertebrates such as springtails by Wiles and Krogh (1998), ISO (1999), Greenslade and Vaughan, (2003), EC (2014a), and OECD (2009) and has been found by ECCC to be appropriate for the culturing of *O. nitens* as well. Working in a chemical fume hood, the culture substrate is prepared in a 1-L wide-mouth plastic bottle with screw-cap lid. First, 120 g of plaster of Paris and 15 g of charcoal are added into the 1-L bottle. Then, 130 mL of *test water* is added, and the bottle is closed and shaken for 30 seconds. The amount of water needed might vary depending on the type of plaster used. Once prepared, the plaster of Paris mixture is then poured into the culture vessel(s) to a depth of 1 cm (Becker-van Slooten *et*

[7] The quality of the plaster of Paris might vary. If the plaster of Paris has a strong odour and reproduction is low, a new batch of plaster of Paris should be used.

al., 2003).[8] This should be done fairly quickly to prevent the substrate from hardening before it is poured into the breeding vessels. The vessels are gently tapped on the sides and on the laboratory bench top to release any air bubbles that might have formed during mixing, as well as to evenly distribute the culture substrate and to create a flat substrate surface.[9] The culture vessels should be placed on a level surface and allowed to air-dry flat for at least three hours. Once hardened, test water is added to the culture vessels to almost saturation (i.e., there should be no standing water on the substrate surface). If the prepared culture vessels are not being used immediately, they can be stored at room temperature. Before storage, the substrate should be saturated with test water (i.e., slowly add ~1 cm of test water on top of the set surface) to prevent it from drying out during storage. Over-drying will result in the substrate shrinking away from the edges of the vessel, thereby creating a gap. If a gap is created between the sides of the vessels and the substrate due to over-drying, the substrate should be discarded since the mites will reside and lay eggs down the sides and at the bottom of the vessel (i.e., where they are inaccessible). The substrate should be rinsed with test water before any organisms are added. Approximately 1 cm of test water should be added to the substrate and the edges and surface gently rubbed with a gloved fingertip to remove any sharp or uneven edges. The substrate should then be rinsed three times. Excess water can be poured off, the surface lightly blotted with paper towel, and the vessels sealed with lids, after which the vessels are ready to use.

A third option for culturing substrate is a combination of the two previously described (i.e., a ≥ 1 cm plaster of Paris substrate covered with a ≥ 1 cm layer of soil).[10]

The pH of a new batch of soil or plaster of Paris substrate should be verified prior to use. The pH of soil can be measured using the $CaCl_2$ method described herein. For plaster of Paris substrate, the pH can be verified by placing pH paper on the wet substrate surface. The pH of both culturing substrates should be between 6.0 and 7.5.

Culture vessels should be rehydrated with test water once or twice/week to maintain the humidity (e.g., optimum humidity is provided by keeping the soil or plaster of Paris moist). This can be accomplished by using a fine mist spray bottle to gently spray the surface of the soil as needed, or by adding several drops of test water with a pipette or squeeze bottle until the water just begins to remain on the surface for the plaster of Paris substrate. Care should be taken to avoid overhydrating (i.e., no standing water on the plaster of Paris substrate or in the bottom of any soil culture bins) and not to damage the mites or to blow organisms out of the culture vessel during the rehydration process.

The culture vessels must be aerated a minimum of once/week; however, twice/week is recommended, and more frequently if there is a history of fungal problems in the cultures. Aeration can be achieved during the weekly rehydration or twice weekly feeding processes by simply removing the lids for ≥ 1 minute.

2.3.6 Food and Feeding

Success in culturing *O. nitens* has been achieved using granulated dry yeast or baker's yeast. Yeast to be used as food for cultures can be purchased from the grocery store. Fleischmann's™ is a brand that has been used successfully during the development of this method (M. Jatar, Environment and Climate Change Canada, personal communication, 2015); however,

---

[8] 120 g of plaster of Paris, 15 g of charcoal, and 130 mL of water make enough substrate for a 16 × 11 × 5.5 cm culture vessel (Becker-van Slooten *et al*., 2003).

[9] Air bubbles leave crevasses on the surface of the culture substrate, which might encourage the mites to lay eggs. Pie-shaped pieces of filter paper coated with plaster of Paris or plaster of Paris caps (made by filling slightly bent aluminum weigh boats with plaster of Paris and allowing them to harden) can be placed on the surface of the substrate to promote egg production. Alternatively, a thin layer of organic matter (e.g., soil or composted leaf litter) placed on the substrate surface might enhance egg production.

[10] In addition to potentially enhancing egg production, the addition of soil on top of the plaster of Paris substrate might extend the duration of healthy culture conditions. It might also facilitate the extraction of organisms from the culture, as mites will cluster around piles of yeast following a feeding event, and the soil can then be easily scooped out of the culture for heat-extraction of the mites.

other brands may be used. The quantity of food added to each culture vessel depends on mite density and developmental stage and, therefore, should be based on observations and records of food consumed, or not consumed, during preceding feedings.

The food may be divided into several piles or sprinkled over the surface of the culture substrate (i.e., soil or plaster of Paris). Piles facilitate easier cleanup of uneaten yeast but should be well spaced so the mites do not have to walk far to find food. The yeast should be placed onto the surface of the substrate of each vessel twice/week at the time of aeration and rehydration. The old, unconsumed yeast (if remaining) may be removed before the new yeast is added.[11] Care should be taken to avoid excessive fungal and bacteria growth in the culture vessels.[12] To activate the yeast, it should be added after the substrate has been hydrated. Alternatively, the yeast can be activated by hydrating it with a few drops of test water, but care should be taken to not over-hydrate as mites could become stuck if the yeast is too wet.

2.3.7 Handling Organisms and Maintaining Cultures

Mites 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 fine-tipped paintbrush is suitable for moving mites to and from the culture or test vessels; however, care must be taken to avoid damaging the organisms. Mites can also be transferred by gently tapping one vessel over another. When handled, any animals that are injured or appear stressed should be discarded and must not be used for testing. Note that grooming behaviour is often exhibited for several minutes after handling but is not reflective of injury.

It is recommended that the contents of each culture vessel be inspected just before each feeding to determine the apparent condition of the mites and the culture substrate. Records should be kept of the apparent condition of the culture (organisms and substrate) noted during each observation period (Section 2.3.1).

The loading density of mites in each culture vessel should be restricted to prevent overcrowding and the resulting adverse effects on mite growth, reproduction, and culture health. For plaster of Paris substrate, a loading density of ~ 5 to 15 adult mites per cm^2 is suggested as it has been observed that the mites do well when slightly crowded (M. Jatar, Environment and Climate Change Canada, personal communication, 2015). For soil, loading density is difficult to assess; however, if the number of juveniles present in the culture begins to diminish (e.g., there are only adults clustered around piles of yeast 24–48 hours following a feeding event), it might be indicative that the culture bins have reached their loading density capacity, and that the culture substrate should be renewed (see Section 2.3.9).

The substrate in each culture vessel should be renewed as required and every four to six months, regardless of organism density. For mites maintained on plaster of Paris substrate, this can be achieved by preparing new culture vessels and by transferring the mites into the new vessels via plaster of Paris caps or pieces which have been baited with food and left in the old vessel.[13] For mites maintained in a soil substrate, cultures can be renewed by heat extracting

---

[11] Old, uneaten yeast might lead to excess populations of bacteria or fungi, which could be harmful to the cultures, but care must be taken as juveniles might be frequently found on or under old food.

[12] Excessive fungal and bacterial growth in the culture vessels might be avoided with the following procedures: use ultra-pure (e.g., Milli-Q®) water for culture substrate preparation and hydration, aerate the culture vessels more frequently (e.g., more than twice/week), and remove any unconsumed yeast (Stämpfli *et al*., 2005). If fungal and/or bacterial growth is excessive in any culture vessel, mites can be baited onto new pieces of plaster of Paris with a few grains of fresh food and transferred to a new culture substrate, or the culture vessel can be discarded.

[13] Plaster of Paris caps baited with yeast are placed on the surface of the culture vessel. Once the mites have congregated on the caps (i.e., after 24–48 hours), they can be moved into a new culture container that contains fresh substrate baited with food. Alternatively, the old plaster of Paris substrate containing mites can be broken into pieces and placed onto a new culture substrate baited with food. Once the mites have moved to the new substrate (i.e., after 24–48 h), the old plaster of Paris substrate pieces can be removed.

the soil in the existing culture bins, and transferring the heat-extracted mites to a fresh culture substrate (see Section 2.3.5).[14] Regardless of substrate type, the population of mites can be reduced in crowded culture vessels by transferring only a portion of the total culture (e.g., 75% of individuals). It is important that new cultures contain a mixture of organisms from different culturing vessels to avoid inbreeding. Alternating cultures between plaster of Paris and soil substrates can help maintain culture health and stimulate culture growth. It is recommended that mite cultures be maintained on at least two different types of substrate to protect against loss of the entire culture, should the organisms fail to thrive in one of the substrate types. The change of substrate and/or the addition of organic matter (e.g., organic manure) on the surface of the culture substrate might enhance the health of the culture and stimulate oviposition (ECCC, 2018).

The air temperature of the culture facility should be monitored weekly and the moisture level of the culture substrate should be observed at the time of weekly aeration. Adjustments should be made as and if necessary (see Sections 2.3.4 and 2.3.5).

2.3.8 Age-synchronized Cultures for Toxicity Tests

To be successful, the culturing procedures used must produce the required number of healthy test organisms of a known developmental stage, and of similar age and size. Additionally, the cultured organisms must meet specific health- and performance-related indices (Section 2.3.9). The following paragraphs describe procedures that should be followed to obtain age-synchronized test organisms (i.e., adults, aged 8 to 10 days after ecdysis to adult stage) for use in the toxicity test described in this method document.

Age-synchronized cultures are established by selecting newly emerged adults from existing plaster of Paris or soil culture extractions and transferring them to fresh vessels containing plaster of Paris substrate. As clear visualization of mites chosen for age-synchronization is necessary, they must be selected from organisms being held on a plaster of Paris substrate (e.g., from cultures maintained on a plaster of Paris substrate, or from soil cultures that have been heat-extracted to the plaster of Paris substrate at the bottom of the heat-extraction unit; see Appendix G). Newly emerged adults are identified based on the appearance of their integument, which ranges in colour from pink to light orange/amber and is almost translucent (see Appendix E for images). Mites must be selected and added to the age-synchronized culture for a period of ≤ 3 days to reach desired numbers. Once the required numbers of mites have been collected, the culture must be fed and hydrated as described for the main culture (see Sections 2.3.5 and 2.3.6) while being allowed to mature for 8 to 10 days before being used in testing.

Any laboratory-cultured *O. nitens* used to start a toxicity test (including that with a reference toxicant) for effects on reproduction should be acclimated in the laboratory as much as possible to conditions representing those in this toxicity test (Section 4.3). During the age-synchronizing period, temperature conditions should be the same as those to be used in the toxicity test, and mites must be fed dry yeast (see Sections 2.3.4, 2.3.6, and 4.3).

2.3.9 Health and Performance Indices

Each culture vessel 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. If the culture appears unhealthy or atypical during any weekly (or more frequent) check, it should be checked daily to make sure that "cascade mortality" (i.e., rate of death increasing exponentially over time) is not occurring. Cultures are considered healthy if *O. nitens* of differing ages are moving actively over

[14] During heat-extraction, the mites are collected on a plaster of Paris substrate in the lower half of the heat-extraction unit (see Appendix G). The mites can then be moved from the heat-extraction vessel into the new soil by tapping the heat-extraction vessel over a culture bin containing fresh substrate or by removing the plaster of Paris base from the heat-extraction vessel and laying it directly on top of a fresh culture substrate baited with food. Once the mites have moved to the new substrate (i.e., after 24–48 h), the plaster of Paris base can be removed.

the substrate surface or are clustered around piles of yeast following a feeding event[15] and there is a low incidence (e.g., ≤ 10%) of non-pigmented (i.e., opaque with a "milky" appearance) organisms (see footnote 6 in Section 2.3.1 and Appendix E).

Technicians at ECCC and some participants of the inter-laboratory investigation have reported occasionally finding mites that are opaque and have a "milky" appearance (ECCC, 2018, 2019). The magnitude and frequency of this occurrence varies, but seems to increase if the culture is over-extended for testing. It has been hypothesized that this colour deviation among oribatid mites might be a genetic defect, a result of stress, or a symptom of malnutrition (Woodring and Cook, 1962; Taberly, 1987; ECCC, 2018). If observed in the cultures, these mites should be discarded and must not be used as test organisms as they do not reproduce during testing (see Appendix E for a photographic example; ECCC, 2018).

There are two possibilities for meeting minimum QA requirements using a known reference substance (e.g., boric acid). The first option is to conduct two multi-concentration reference toxicity tests annually (i.e., once every six months) using age-synchronized organisms derived from the same culture(s) from which the test organisms for the definitive soil toxicity test(s) are obtained. The second option is to include a positive control concentration with each definitive mite toxicity test using the same age-synchronized test organisms as those used in the definitive test (see Section 4.9 for details). All tests with the reference toxicant(s) should be performed using the conditions and procedures outlined in Section 4.9. 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 Section 4.4.

A laboratory that routinely performs toxicity tests with mites might find it useful to monitor the data on the number of *progeny* produced in negative control soil as a measure of culture health and performance. A plot of such data over time might show problems with respect to reproductive success that are attributable to diet or other conditions to which cultures are exposed (G. Stephenson, AquaTerra Environmental, personal communication, 2016).

[15] If the number of juveniles observed one to two days following a feeding event starts to diminish (i.e., only adults are clustered around piles of yeast), this might be indicative of declining culture health and the need for culture renewal (J. Princz, Environment and Climate Change Canada, personal communication, 2019).

Section 3

Test System

3.1 Facilities and Apparatus

Tests must be performed in an environmental chamber or equivalent facility having acceptable temperature and lighting control (see Section 4.3). The test facility should be well ventilated to prevent personnel from being exposed to harmful fumes, and it should be isolated from physical disturbances or any contaminants that might affect the test organisms. The area used to prepare *test soils* should contain a fume hood and be properly ventilated.

The test facility must be isolated from the area where the mites are cultured (Section 2.3) 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 vessels and their contents from these sources. The ventilation system should be designed, inspected, and operated to prevent air within the testing facility from contaminating the culturing 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, soil, water, or test vessels within this facility must be non-toxic (see Section 2.3.2) and should minimize sorption of chemicals. Borosilicate glass, nylon, high-density polyethylene, high-density polystyrene, polycarbonate, fluorocarbon plastics, Teflon™, Nalgene™, porcelain, fibreglass, and type 316 stainless steel should be used whenever possible to minimize chemical sorption and leaching. The use of toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must be avoided.

The test facility must have the basic instruments required to monitor the quality (e.g., temperature, 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 that can be set at 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 vessels, equipment, and supplies that might contact site soils, test soils, control soils, test (hydration) water, *stock solutions,* or test solutions must be clean and rinsed with *deionized* or *distilled water* (i.e., test water) before use. All non-disposable materials should be washed after use. The following cleaning procedure is recommended (EC, 2004a, 2005a, 2013a, 2014a):[16]

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[17]) nitric (HNO_3) or hydrochloric acid (HCl) (metal-free grade) to remove scale, metals, and bases;

4. rinse twice with deionized 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%

[16] Steps 1−4 of the cleaning procedure should be used if metal contamination is of concern; steps 1, 2, 5, 6, and 7 should be used if contamination with organics is of concern; and all steps should be followed if both metal and organics contamination is suspected.

[17] To prepare a 10% solution of acid, carefully add 10 mL of concentrated acid to 90 mL of deionized water.

purity) hexane for oily residues (use a fume hood);[18]

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

7. rinse three times with deionized water (or other test water).

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

3.2 Initial and Definitive Tests

3.2.1 Initial Tests

Before definitive soil toxicity tests 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 3.3) be undertaken by laboratory personnel. Additionally, a minimum of five reference toxicity tests should be performed 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 Section 4.9). These initial tests are recommended to confirm that acceptable performance of the test species can be achieved in a candidate natural or artificial negative control soil (see Section 3.3) in a specific laboratory and under the culturing conditions and procedures specified in this report (see Section 2.3).

The conditions and procedures used to perform these initial tests with negative control soil should be identical and according to Section 4, whereas the conditions and procedures used to perform the initial reference toxicity tests should be identical and according to Section 4.9. Each test with negative control soil or reference toxicant(s) should be performed using a different lot of test organisms from the same source.

Data from the control performance tests ($n \geq 5$) must show that the criteria for test validity (see Section 4.4) can be met using a natural or artificial soil intended for use as a negative control soil in a definitive soil toxicity test. Data from the initial reference toxicity tests ($n \geq 5$) should be compared by calculating and appraising the magnitude of the coefficient of variation (CV) for the respective series of tests and endpoint values (see Section 4.9).

3.2.2 Definitive Tests

Test vessels to be used in definitive tests must be inert to test and reference substances or contaminant mixtures (i.e., the test or reference substances, or mixtures thereof, should not adhere to or react in any way with the test vessel). The volume of the vessel should be sufficiently large to accommodate mite survival and reproduction for the duration of the test. Glass shell vials with a capacity of ~30 mL (~2.6 cm internal diameter) have been successfully used as test vessels in the development of this test method; however, vessels with other volumes and dimensions may be used. Each vessel must be cleaned thoroughly before and after use, and rinsed well with deionized or other test water before use. Each test vessel should be covered with a plastic or metal lid with a small hole to allow for gas exchange.

3.3 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 mites 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

[18] Rinsing Plexiglas™ or any plastic equipment or vessels with acetone or hexane is **not** recommended, since plastic could become pitted and etched by these solvents and could turn from transparent to opaque.

natural soil with acceptable properties.[19] For definitive tests with field-collected boreal forest and taiga soils, it is recommended that uncontaminated natural soil be used as the negative control soil. Regardless of soil type, there must 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 (Section 4.4).

The biological test method described herein was developed and tested using 11 negative control soils with diverse physicochemical characteristics (EC, 2010, 2013b, 2014b; Hennessy, 2010; Princz *et al*., 2010, 2012, 2018; Princz, 2014; Ritchie *et al*., 2017; ECCC, 2018, 2019). These clean soils included two artificial soils, four natural agricultural soils (i.e., two sandy loam soils from Alberta and Quebec, a clay loam soil from Ontario, and loamy sand LUFA soil obtained from Germany), and five soils collected from boreal forest and taiga ecozones within Canada (i.e., two Gleyed Humo-ferric Podzols from Newfoundland and Ontario, a Dark Grey Luvisol and an Orthic Eutric Brunisol from Saskatchewan, and a Rego Humic Gleysol from Alberta). 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 therefore were considered to be "clean." The origin and physicochemical characteristics of these natural soils are further described in Appendix D. The test validity criteria for *O. nitens* described in Section 4.4 are based on the performance data for these organisms in negative control soil that were generated for each of these diverse soils (ECCC, 2018). During the development of the test method, there was an observed trend of lower numbers of juveniles produced in some of the sub-surface soil horizons (often, but not limited to those soils with lower % OM).

3.3.1 Natural Soil

Negative control soil may be natural soil collected from a clean (uncontaminated) site that is 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 mites were collected to establish a culture (Section 2.2).

All samples of natural soil selected for possible use as negative control soil in soil toxicity tests (as well as samples of candidate *reference soil*) must be analyzed for the following physicochemical characteristics:

- particle size distribution (% sand, % silt, and % clay)
- *total organic carbon* content (%)[20]
- organic matter content (%)[20]
- pH
- *electrical conductivity*
- moisture content (%)
- *water-holding capacity* (WHC)
- *cation exchange capacity* (CEC)

Additionally, the following analyses should be performed:

- major cations and anions (Na^+, K^+, Mg^{2+}, Ca^{2+}, Al^{3+}, S^{2-}, Cl^-)
- nitrogen as total N, nitrate (NO_3^-), nitrite (NO_2^-), and ammonium (NH_4^+)
- phosphorus as total and/or bioavailable
- potassium as total and/or bioavailable

---

[19] The Canadian Council of Ministers of the Environment (CCME) provides a comprehensive website 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 website 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 (1-204-948-2090) or email (info@ccme.ca).

[20] Organic matter content can be used to calculate total organic carbon (TOC) by multiplying the organic matter content (OM) of a soil by a soil constant (AESA, 2001). However, the relationship between TOC and OM is slightly different among soils, and the total organic carbon content should also be determined by laboratory analysis.

- C:N ratio

To confirm that the negative control and/or reference soils are not contaminated, the following screening analyses are recommended:

- organophosphorus insecticide suite
- organochlorine insecticide suite
- herbicide suite
- metal suite
- petroleum hydrocarbons (including PAHs)
- other site- or area-specific contaminants of concern

Pesticide and metal concentrations should not exceed CCME soil quality criteria, if available (see footnote 19). If indigenous organisms are present and/or problematic in the sample(s) of natural soil at any time (i.e., during storage or testing), their presence (e.g., physical description and estimated numbers) should be recorded, and they should be removed manually (e.g., by sieving), if possible. Alternatively, most indigenous organisms (predatory mites being of particular concern here) can be killed by at least one or more freeze/thaw cycles if it is suspected that they are too small to remove manually (see footnote 75 herein, and Section 5.6.6 of EC, 2012). If the results of both the initial 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 (4−10 mm, depending on the type of soil),[21] transferred to clean, thoroughly rinsed plastic pails, and stored in darkness at 4 ± 2°C until required. Plastic pails should not be used for collection and storage of soils if there are concerns about chemical constituents of the plastic leaching into the soil.

3.3.2 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).

In keeping with the formulation of artificial soil used in four other Environment Canada soil toxicity test methods (EC, 2004a, 2005a, 2013a, 2014a), the following three ingredients should be used to prepare artificial soil to be used in the biological test method described herein (based on dry mass):

i) 10% *Sphagnum* sp. peat, air-dried and sieved (e.g., through a 2-mm mesh screen)

ii) 20% kaolin clay with particles < 40 μm

iii) 70% "grade 70" silica sand

The ingredients (above percentages expressed as dry mass fraction) should be mixed thoroughly in their dry form using a mechanical stirrer and/or gloved hands.[22] Reagent-grade calcium carbonate should be added to the dry mixture in a quantity sufficient to attain a pH (measured using a calcium chloride slurry method; see Section 4.6) for the artificial soil ranging within 6.0−7.5 once it is hydrated.[23]

---

[21] The more porous sieve sizes (e.g., 6–10 mm) might be needed for soils with a higher organic content. Further guidance on the requirement for sieving, including appropriate sieve size selection, is provided in EC (2012).

[22] 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.

[23] 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 might occur when the soil is first formulated without the addition of $CaCO_3$. The initial pH adjustment should attempt to raise pH to a range within 7.0 to 7.5, since the pH of artificial soil typically drops slightly (to 6.5 to 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, 1998; G. Stephenson, AquaTerra Environmental, personal communication, 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 (or until it has the optimal texture for testing) when

Thereafter, the mixture should be hydrated gradually using test water (i.e., deionized or distilled water) until its moisture content is ~20% (which is ~28% of the soil's water-holding capacity),[24] while mixing further until the soil is visibly uniform in colour and texture. As necessary, additional 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. 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 or until it has the optimal texture for testing (i.e., a homogeneous crumbly consistency with clumps ~1 to 3 mm in diameter; see Section 5.3).

Samples of artificial soil selected for possible use as negative control soil in soil toxicity tests must be analyzed for the following physicochemical characteristics:

- particle size distribution (% sand, % silt, and % clay)
- total organic carbon content (%)[20]
- organic matter content (%)[20]
- pH
- electrical conductivity
- moisture content (%)
- water-holding capacity (WHC)
- cation exchange capacity (CEC)

Additional analyses, such as those described for natural soils (Section 3.3.1) may also be carried out, as necessary.

3.4 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 mites 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 may be a sample of negative control soil that is spiked with a reference toxicant for which historical data are available on its toxicity to mites using the specified test conditions and procedures. For the biological test method described herein, one or more reference toxicants must be used in a multi-concentration test or as *replicates* of a positive control soil (i.e., at a specific concentration) when appraising the sensitivity of the test organisms and the precision and reliability of results obtained by the laboratory for that material (see Section 4.9). A test might also include a sample of negative control soil (natural or artificial; see Section 3.3) 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 *O. nitens*, and according to the biological test method described herein. In some instances, a test might include a positive control soil that consists of a highly contaminated sample of field-collected soil or sludge shown previously to be consistently toxic to the mites according to the biological test method described herein.[25]

---

required 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 those 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 the 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 the time for pH stabilization associated with dry storage of artificial soil.

[24] The % hydration might need to be adjusted higher or lower depending on the type of peat used in preparing artificial soil.

[25] If the positive control soil consists of a highly contaminated sample of field-collected soil, it is important that its toxic potential be stable over time (i.e., the sample

3.5 Reference Soil

One or more samples of reference soil might be included in a soil toxicity test using mites. 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, and electrical conductivity) represent the sample(s) of test (contaminated) soil as much as possible. Ideally, the reference soil is collected from the general vicinity of 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 near the test site(s) might also be chosen due to their known lack of toxicity in previous tests with mites, and their possession of physicochemical characteristics similar to the test soil samples. Boreal forest and taiga reference soils must be collected as separate *soil horizons*, where possible. Each soil horizon must then be stored and tested individually (i.e., each horizon is treated as a separate soil sample) (see Section 5.1 and EC, 2012). The sample(s) of field-collected reference soil used in a study could be tested for toxic effects as undiluted soil only, or this soil could be mixed with the sample(s) of test and reference soils to prepare a range of concentrations to be included in a multi-concentration test[26] (see Sections 3.6, 4.1, and 5.3). 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 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 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 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 3.3). 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 5.6); 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.

3.6 Test Soil

This biological test method is intended to measure the toxicity of one or more samples or mixtures of contaminated or potentially contaminated soil (test soil), using mites 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-

---

is old enough that the bioavailability has been stabilized).

[26] 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 soil*s* are likely known to 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 those of the test soil.

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).

Field-collected soils collected by horizon take into account contamination stratified due, in part, to the different speciation and resultant mobility of contaminants (EC, 2012). Therefore, for soils collected from the boreal or taiga ecozones, both reference and contaminated soils must be collected in separate horizons. Soils collected in horizons must be treated as individual soil samples and tested separately (see Section 4.1). Soils without distinct soil horizons (e.g., where the surface soil horizons have been mixed or disturbed due to human activity) should be collected according to depth (see Section 5.1). 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). Guidance on the collection, handling, analyses, and testing of field-collected soils is provided in Section 5.

Section 4

Universal Test Procedures

General procedures and conditions described in this section for toxicity tests with mites apply when testing the toxicity of samples of soil, particulate waste, or chemicals, 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. Specific guidance on conducting tests with boreal and taiga soils has been incorporated throughout this test method document.

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 *O. nitens* in preparation for soil toxicity tests also apply.

Summary checklists in Table 2 describe required and recommended conditions and procedures to be universally applied to each 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 (including boreal and taiga soils), biosolids mixed into soil (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 measures the effects of exposure to contaminated soil on the reproductive success of mites. Test organisms must be age-synchronized laboratory cultured *O. nitens*. Test duration is 28 days[27] and the test soils are hydrated during the test but not renewed.

[27] Studies were conducted to assess the reproduction from exposure of 15 and 20 age-synchronized adults to soil for either 28 d or 35 d (EC, 2013b). Reproduction was greater after 35 d of exposure to the soil, but it became more difficult to distinguish between the original adults and the F1 generation (i.e., progeny), as newly emerged adults matured within the 35-d exposure scenario. As a result, a decision was made to remain with the 28-d test duration (ECCC, 2018).

Table 2 Checklist of required and recommended conditions and procedures for conducting tests for effects of exposure to contaminated soil on the reproduction of *Oppia nitens*

Universal		
Test type	−	whole soil toxicity test; no renewal (static test)
Test duration	−	28 days
Test organisms	−	*O. nitens*: age-synchronized laboratory cultures; collected over a ≤ 3-day period, aged 8 to 10 days after ecdysis to adult stage
Number of replicates	−	≥ 5 replicates per treatment: 15 age-synchronized test organisms per replicate (test vessel)
Negative control soil	−	depends on study design and objectives; clean field-collected soil or artificial soil if testing site soils; artificial soil recommended for tests with chemical(s) or chemical product(s) spiked in soil
Test vessel	−	30-mL glass shell vial (~2.6 cm inner diameter) or other appropriate vessels, covered; plastic or metal lid with pinhole for gas exchange recommended
Amount of soil/ test vessel	−	a volume of ~ 20 mL at optimal soil moisture content for 30-mL glass shell vials or equivalent; ≥ 3 cm soil depth; 3–4 cm soil depth recommended
Moisture content, test soils	−	for soil preparation, hydrate to the optimal percentage of its WHC if field-collected soil (see Section 5.3), ~70% of WHC if artificial soil (see Sections 3.3.2 and 6.2); during test, hydrate as necessary
Air temperature	−	daily average, 20 ± 2°C; instantaneous, 20 ± 3°C
Lighting	−	incandescent, fluorescent, or LED; intensity, 400−800 lux adjacent to surface test vessels (must be ≥ 400 lux); fixed photoperiod (e.g., 16 h L:8 h D or 12 h L:12 h D)
Feeding	−	granulated dry yeast (e.g., Fleischmann's™); every 7 days from Day 0 to Day 21; ~0.5–1 mg per test vessel each feeding; sprinkled onto soil surface
Aeration and hydration	−	open test vessels briefly, minimum once/week to aerate and to assess moisture level; determine moisture loss by weighing test vessels (e.g., weigh vessels at test start and weekly thereafter), and hydrate if loss is > 2% of the initial water content
Measurements during test	−	air temperature in test facility, daily or continuously; moisture content, pH, and electrical conductivity (if necessary) of soil in each treatment/concentration, at test start and end; any excessive growth of fungi, presence and estimated quantity of any uneaten food, and apparent "wetness" of soil, at least weekly
Endpoints	−	total number of surviving adult mites and total number of surviving progeny in each test vessel at the end of the test (Day 28); mean (± SD) percent survival of adults in each treatment, at test end (Day 28); mean (± SD) number of surviving progeny in each treatment, at test end (Day 28); if multi-concentration test: 28-day ICp for mean number of surviving progeny produced per treatment at test end

Test validity	−	invalid if mean survival of adults (first generation) in negative control soil at test end is < 70%; invalid if mean reproduction for adults in negative control soil is < 30 progeny/vessel
Test with reference toxicant	−	choose between a positive control concentration or a multi-concentration reference toxicity test: − if the positive control is chosen, it must be performed with every definitive test; use boric acid (H_3BO_3) or similar; prepare and test ≥ 5 replicates of a predetermined concentration, using artificial soil as a substrate; 15 mites/replicate; follow procedures and conditions described in Section 4.9 and Appendix H; determine % reduction in progeny production (as a percent of the control response) at test end (i.e., Day 28) − if the multi-concentration reference toxicity test is chosen, it must be performed twice per year; use boric acid (H_3BO_3) or similar; prepare and test ≥ 5 concentrations plus a negative control, using artificial soil as the substrate; ≥ 5 replicates for negative control and test concentrations; 15 mites/replicate; follow procedures and conditions described in Section 4.9; determine 28-day IC50 for inhibition of number of progeny (including 95% confidence limits); express as mg boric acid/kg soil dry weight; validity criteria are the same as those for definitive test

Field-collected Soil

Transport and storage	−	seal in plastic or other appropriate material, and minimize air space; labelled or coded; 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 28-day tests have shown that all criteria for test validity could be regularly met; analyzed for at least the following: particle sizes (% sand, % silt, % clay), TOC (%), OM (%), pH, electrical conductivity, moisture content (%), WHC, and CEC.
Reference soil	−	one or more samples for tests with field-collected soil; taken from site(s) presumed to be clean but near sites of test soil collection; characteristics (e.g., organic carbon content, OM [%], particle size distribution, texture, pH, and electrical conductivity) similar to test soil(s); analyzed as described for natural negative control soil
Characterization of test soils	−	must include at least moisture content (%), WHC, pH, electrical conductivity, TOC (%), OM (%), particle sizes (% sand, % silt, % clay), and CEC; should include at least nitrogen, phosphorus, potassium, C:N ratio, major cations and anions; and, optionally, bulk density, total inorganic carbon, total volatile solids, biochemical oxygen demand, chemical oxygen demand, redox potential, soluble salts, sodium adsorption ratio, contaminants of concern (e.g., metals, polycyclic aromatic hydrocarbons, pesticides), and characteristics of the contamination (e.g., odour, staining, debris, presence of fuel or solvent)
Preparation of test soils	−	if necessary, remove debris and indigenous macro-organisms using forceps; if necessary, gently pass through a sieve of suitable mesh size (e.g., 4−10 mm); at least one or more freeze/thaw cycle for soil horizons with high organic content; homogenize; determine percent moisture content, and WHC; hydrate with test water (or, if and as necessary, dehydrate) to the optimal percentage of its WHC; 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	−	artificial soil or a clean field-collected soil for which previous 28-day tests have shown that all criteria for test validity could be regularly met; analyzed for at least the following: particle sizes (% sand, % silt, % clay), TOC (%), OM (%), pH, electrical conductivity, moisture content (%), WHC, and CEC.
Characterization of chemical(s) or chemical products	−	information on concentration of active ingredients and impurities, water solubility, vapour pressure, stability, dissociation constants, adsorption coefficients, toxicity to humans and terrestrial organisms, and biodegradability of chemical(s) or chemical product(s) spiked into negative control soil should be known beforehand
Solvent	−	deionized water is the preferred solvent; if an organic solvent is used, the test must include a solvent control soil in addition to a negative control soil
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 and/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 consisting partly or completely of the test substance(s); mixing conditions are standardized for each treatment; ensure homogeneity
Concentration within soil mixture of chemical(s) or chemical product(s) added	−	normally measured at beginning and end of test, in high, medium, and low concentrations as a minimum

---

[*] **The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.**

This definitive test method was applied and validated by several participating laboratories in three rounds of concurrent tests using *O. nitens* in artificial and field-collected soils spiked with boric acid (ECCC, 2019).[28]

4.1 Preparing Test Soils

Each test vessel (see Section 3.2.2) 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 vessels should be positioned such that observations and measurements can be made easily. Treatments should be positioned randomly within the test facility, and the position of test vessels within the test facility should be changed regularly during the test (i.e., once per week, randomly) (EC, 2004a, 2005a, 2013a, 2014a).

The day that mites are first 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[29] (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 or soil horizon mixed as a batch should be enough to establish the *replicates* of that treatment (see Table 2), plus an additional amount for the physicochemical analyses to be performed (Section 4.6) and a surplus to account for the unused soil that adheres to the sides of the mixing container. 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 can 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, a volume of test soil that provides a 3–4 cm soil depth

---

[28] In the first round of the interlaboratory validation tests, eight laboratories participated in a 28-day test with *O. nitens* using three different clean control soils: a natural forest soil, a natural sandy loam soil, and artificial soil. This initial round was intended to allow the laboratories to establish familiarity with the test species and test methodology. Eight laboratories participated in the second round of the interlaboratory validation tests, which consisted of reproduction tests with *O. nitens* exposed to boric acid in a field-collected sandy loam soil for 28 days. All laboratories met the proposed minimum acceptable control performance criteria for adult survival of ≥ 70% and the proposed control performance criteria for reproduction of ≥ 30 mean juveniles (i.e., progeny) per replicate; one laboratory failed the proposed test validity criteria for artificial soil (29 ± 12 juveniles) only. The coefficients of variation for adult survival in the control soils were 5.1% and 8.2% for the artificial and field-collected sandy loam soils, respectively; the corresponding coefficients of variation for progeny production were 41 and 42%, respectively. The mean IC50 for progeny production was 105 ± 7.0 mg H_3BO_3/kg soil dry weight, with values ranging from 98 to 118 mg/kg. The interlaboratory variability, expressed as the coefficient of variation, was 6.7% (ECCC, 2019). Eight laboratories participated in the third round of the interlaboratory validation tests. These were reproduction tests with *O. nitens* exposed to boric acid in LUFA 2.2 soil for 28 days. All laboratories met the proposed minimum acceptable control performance criteria for adult survival of ≥ 70% and the proposed control performance criteria for reproduction of ≥ 30 mean juveniles per replicate. Coefficients of variation for adult survival in the control soils were 8.2% and 5.9% for the artificial and LUFA 2.2 soils, respectively; the corresponding coefficients of variation for progeny production were 37% and 52%, respectively. The mean IC50 for progeny production was 86 ± 24 mg H_3BO_3/kg soil dry weight, with values ranging from 42 to 114 mg/kg. The interlaboratory variability, expressed as the coefficient of variation, was 28%, demonstrating acceptable agreement among laboratories (ECCC, 2019).

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

(i.e., ~20 mL for the 30-mL glass shell vials or equivalent) should be transferred to each replicate test vessel. The volume of soil in each replicate test vessel must be the same and must provide a minimum soil depth of 3 cm in the test vessel. The soil added to each test vessel should be smoothed (but not compressed) using a spatula or by gently tapping the glass vial on the benchtop or with a hand.

For soils collected as distinct horizons (e.g., boreal or taiga soils), each horizon must be prepared and tested separately in independent definitive tests. For soils to be assessed in multi-concentration tests, each horizon of the test soil should be mixed with the same horizon of negative control or reference soil (see Section 5) at the various test concentrations (0%, 6.25%, 12.5%, 25%, etc.). In some cases, it might not be possible to collect the same horizons of negative control soil and test soil. For example, negative control soils might be collected in horizons, but this might not be possible at the site of contamination, i.e. more than one horizon of test soil might not be present or horizons might be mixed. In this case, test concentrations should be prepared by mixing suitable weights of test soil into the available horizon(s) of negative control soils at the appropriate test concentrations.

For a single-concentration test (e.g., site soil tested at 100% concentration only; a particular concentration of test soil; or a chemical-spiked soil tested at one concentration [e.g., Maximum Label Rate]), a minimum of five replicate test vessels as well as five replicate negative control test vessels must be set up by adding soil from the same batch to each replicate vessel. For site soils, replicates should represent *replicate samples* (i.e., field replicates) collected individually from a given sample location (see Section 5.1). For a multi-concentration test, a minimum of five replicate test vessels per negative control soil and a minimum of five replicate test vessels per treatment must be set up. *Power* analysis carried out on reproduction data generated using this method (Section 5.6.2) indicated that for five laboratory replicates, an effect size of 40% or greater can be reliably detected (power ≥ 80%). More replicates (i.e., ≥ 8) are recommended to reliably detect a smaller (i.e., 30%) effect size (Section 5.6.2). In the case of appreciable uncertainty about sample toxicity, a *range-finding* test might prove worthwhile for selecting, more closely, the concentrations to be used for the definitive test. For a range-finding test, the number of replicates used may be reduced (e.g., three replicates). For any test that is intended to estimate the inhibiting concentration for a specified percent effect (*ICp*) in a definitive multi-concentration test, at least seven concentrations plus the *control* treatment(s) must be set up, and more (i.e., ≥ 10 plus controls) may be used to improve the likelihood of bracketing the endpoint sought.[30] If a range-finding test is conducted prior to definitive testing, fewer concentrations may be used in the definitive test, since more information on the effect concentration / dilution range will be available. In this case, a minimum of five test concentrations must be used.

It is recommended that a minimum of 25 g of soil for each treatment (including any control or reference soils used) be included in the test for the purposes of conducting physicochemical analyses on Day 0 and at the end of the test (see Section 4.6).[31]

Concentrations should be chosen to span a wide range, including a low concentration that evokes no adverse effects (e.g., similar to that for the negative control treatment), and a high concentration that results in "complete" or severe effects. If the anticipated endpoint is bracketed with a closely spaced series of concentrations, all might 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 in order to split the selected range more finely. In any case, a consistent

---

[30] The use of 10 or more concentrations (plus the controls) can be used to better show the shape of the concentration-response relationship and to choose the appropriate linear or non-linear regression model (see Section 4.8.1.1).

[31] The heat-extraction at the end of the test often modifies the physicochemical properties of the test soil. Therefore, extra replicates (with or without test organisms, depending on the objectives) should be prepared for the sole purpose of conducting physicochemical measurements at test end (see Section 4.6).

geometric series should be used (see Appendix F). See EC (2005b) for additional guidance on selecting test concentrations that apply here.

Following the addition of a measured aliquot of test soil to each test vessel, lids (Section 3.2.2) should be placed onto the test vessels and closed tightly to minimize moisture loss. The test vessels should be held overnight under specified test temperature and lighting conditions (Section 4.3), for chemical equilibration (e.g., of chemical-spiked soil or site soil diluted with control soil) of the test soils. If there is concern over the volatilization, degradation, or metabolism of contaminants or chemicals in test soils, the test can be initiated immediately after the preparation of the test soil (see Section 6.2). The dates test and control soils are prepared and organisms are added to the test vessels must be recorded and reported.

4.2 Beginning the Test

Test organisms are transferred to each test vessel the day after the soil is prepared (i.e., Day 0). A number of test organisms in excess of those required for the test should be available from a group of age-synchronized culture vessels established to yield the appropriate number of organisms required for a test (Section 2.3.8).

Adult mites, aged 8 to 10 days after ecdysis to adult stage, from age-synchronized cultures (see Section 2.3.8 for information on the age-synchronization) must be used. Fifteen organisms must be transferred into each test vessel.[32]

Organisms may be gently transferred from the age-synchronized culture to a piece of folded stiff cardboard (e.g., 8.5 × 11 in. paper folded in half), a small glass container, or a weigh boat (previously washed and dried to remove the waxy film that coats the weigh boats) using a fine paintbrush. Final observation of mites should be made to confirm the correct number and that their appearance is normal (i.e., organisms chosen should appear healthy and active, demonstrating movement, lack of visible defects or damaged bodies, and should be similar in colouration[33]). Any atypical mites should be discarded. Thereafter, the organisms should be carefully transferred to the surface of the soil in a test vessel by gently tapping the cardboard or the weigh boat over the test vessel. The group of mites transferred to each test vessel should be random across the replicates and treatments.

4.3 Test Conditions

- This is a 28-day soil toxicity test during which the soil in each test vessel is not renewed.
- The test vessel (e.g., 30-mL glass shell vial) and its contents (i.e., a volume providing a ≥ 3 cm depth of test soil) are covered (Sections 3.2.2 and 4.1).
- For a single-concentration test, at least five replicate test vessels must be set up for each test soil (i.e., each treatment) and for each control soil. For a multi-concentration test, a minimum of five replicate test vessels per test concentration and five replicate test vessels per control soil must be set up.
- For multi-concentration tests, at least seven concentrations plus the appropriate control treatment(s) must be used. If a range-finding test is conducted prior to the definitive test, the number of concentrations may be reduced but a minimum of five concentrations must be used.
- The test must be conducted at a daily mean temperature of 20 ± 2°C. Additionally, the

[32] Initial studies were carried out using 15 and 20 age-synchronized adults. The use of 15 adults was preferable, as it decreased the total number of organisms required for a test by 25%, yet allowed for sufficient juvenile production (ECCC, 2018). Further studies were also conducted to evaluate gender-specific biases in selecting the individual test organisms, given that gender could not be selected *a priori*. The gender of adult mites extracted from five tests using artificial soil was examined under a microscope once the mites were cleared with lactic acid; the resulting ratio of males to females was 49 ± 11% to 51 ± 11%, respectively, through random selection of test organisms from age-synchronized cultures (ECCC, 2018).

[33] Any individuals that appear damaged, undersized (relative to the others chosen), or coloured differently (e.g., opaque or "milky" – see Appendix E) must not be used in the test.

instantaneous temperature must always be 20 ± 3°C.

- Test vessels 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), and should use incandescent, fluorescent, or LED lights. Light intensity adjacent to the surface of the soil in each test vessel should be 400−800 lux, and must be at least 400 lux as a minimum (Section 2.3.3).

4.4 Criteria for a Valid Test

For the results of the test described in this biological test method to be considered valid, each of the following two criteria must be achieved:[34]

i) the mean survival for adult mites held in negative control soil for 28 days must be ≥ 70% at the end of the test, and

ii) the reproduction for the adult mites in negative control soil for 28 days must average ≥ 30 progeny per control vessel at the end of the test.

4.5 Food and Feeding

During a toxicity test, *O. nitens* in each test vessel must be fed granulated dry yeast every seven days, starting at Day 0 and continuing until and including Day 21. Approximately 0.5–1 mg of granulated yeast should be added to each test vessel at the time of feeding. The type of yeast used must be a dried, granulated yeast (e.g., Fleischman's™) and should be prepared by distributing the yeast uniformly over the surface of the moist test soil (optimally moistened soil should be sufficient to hydrate and activate the yeast with time). It is important that the same amount of yeast is available to organisms in each test vessel. If, when adding yeast to a test vessel, it is noticed that the yeast from a previous feeding period has not been consumed, the unconsumed yeast should not be removed but a reduced amount of fresh yeast is added to the test vessel at that time.[35]

4.6 Observations and Measurements During the Test

The biological endpoint for this test is the number of progeny produced in each test vessel at the end of the test (Day 28). The condition, appearance, and number of live mites transferred to each test vessel on Day 0 must be observed and recorded. The lid must be removed from each test vessel for the purpose of aeration at least once/week or more frequently (i.e., ≥ 2 times per week) as necessary, or as the test progresses and the number of organisms per test vessel increases.[36] Observations and records should be made at this time regarding any excessive growth of bacteria or fungi, and the presence and estimated quantity of any uneaten food.

Air temperature in the test facility (Section 4.3) must be measured daily (e.g., using a maximum/minimum thermometer) or continuously (e.g., using a continuous chart recorder).

The contents of each replicate vessel must be examined at least once weekly for apparent "wetness." [37] Moisture loss should be determined by weighing test vessels. All test vessels can be

---

[34] The test validity criteria presented here are based on control data generated in many studies carried out during the development of the method (EC, 2010, 2013b, 2014b; Hennessy, 2010; Princz *et al*., 2010, 2012, 2018; Princz, 2014; Ritchie *et al*., 2017; ECCC, 2018, 2019). Clean soils included in the development of the test validity criteria included two artificial soils, four agricultural soils, and five boreal soils (including nine different horizons in total; see Appendix D). The validity criteria were based on a calculation of the 5th percentile (ECCC, 2018).

[35] If mycelium develops on the soil surface, simply disturb it by carefully breaking it up using forceps and a microscope while taking care to not disrupt the test organisms.

[36] Gently tapping on the test vessel lid to dislodge any mites and slowly removing the lid allows any individuals hiding under the lid to fall back into the test vessel.

[37] The apparent "wetness" of a soil is affected by the nature of the soil and the amount of water lost from test vessels due to evaporation. Soils might appear too dry when the WHC has been underestimated (see Section 5.3).

weighed at the beginning of the test. The weight of each test vessel can then be checked at least once per week and test water added to compensate for weight loss (i.e., due to water loss), if the loss is > 2% of the initial water content (ISO, 1999). For a large number of test vessels, the average amount of water lost can be calculated by weighing a random sample of 10% to 20% of the test vessels at the beginning of the test and once per week thereafter. This amount of test water can then be added to all of the test vessels. Test water can be added using a gentle mist, or with a dropper.

The pH and moisture content of the test soil or soil horizon 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 electrical 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 (Day 0) measurements should be made using a *composite sample* made up of subsamples of each batch of test soil or soil horizon used to set up replicates of a particular treatment (see Section 4.3).[38] The final (i.e., Day 28) measurements should be made using additional replicates set up for each treatment (see Section 4.1) that are analyzed at the end of the test.

Soil pH should be measured using a calcium chloride ($CaCl_2$) slurry method (modified from Hendershot *et al*., 1993; as recommended by Becker-van Slooten *et al*., 2004).[39] For these analyses, 4 g of hydrated soil[40] are placed into a 30-mL glass beaker (~3 cm in diameter and ~7 cm high) with 20 mL of 0.01 *M* $CaCl_2$.[41] 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 when the meter reading is constant, the pH is recorded.

The moisture content of each test soil or soil horizon should be measured by placing a 3−5 g subsample of each test soil or horizon 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 the calculation of moisture content (%) be based on dry weight (**not** on wet

[38] On the day before the start of the test (Day -1), one or more additional replicates of each test soil should be placed into a test vessel within the test facility. These replicates should be reserved for physicochemical analyses of Day 0 conditions to which the mites are exposed. A separate set of replicates should also be set up on Day -1, for physicochemical analyses of test end (Day 28) conditions. These additional replicates might or might not have organisms added on Day 0.

[39] The method by Hendershot *et al.* (1993) includes a step that involves air-drying the sample for 48 h before its analysis for pH. The experience of Environment Canada investigators is that this step is needlessly time consuming (K. Doe, Environment Canada, personal communication, 2004; J. Princz, Environment Canada, personal communication, 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, Environment Canada, personal communication, 2004).

[40] It might be necessary to use a lower soil:$CaCl_2$ solution ratio (e.g., 2 g of soil: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).

[41] To prepare 0.01 *M* $CaCl_2$, dissolve 2.940 g of calcium chloride dihydrate ($CaCl_2 \cong 2H_2O$) with distilled water in a 2000-mL volumetric flask. The electrical 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 range using a hydrogen chloride (HCl) or calcium hydroxide [$Ca(OH)_2$] solution. If the electrical conductivity is not within the acceptable range, a new solution must be prepared.

weight), since the results of these calculations are used with calculations of water-holding capacity (also calculated based on dry weight) 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 5.5). For a multi-concentration test with chemical-spiked soil, such measurements should be made for the high, medium, and low concentrations tested, as a minimum (see Section 6.3). Aliquots for these analyses should be taken for each soil or soil horizon as described previously for pH and moisture content; analyses should be according to proven and recognized analytical techniques (e.g., SPAC, 1992; Carter, 1993; Carter and Gregorich, 2008).

4.7 Ending the Test

The test must be terminated after 28 days of exposure (i.e., Day 28). At that time, the number of surviving adult mites and the number of live progeny produced in each test vessel must be observed and recorded. Before opening a test vessel, the lid should be tapped (e.g., three times) to dislodge any individuals from the underside. Live mites (adults and progeny) must be extracted from the test soil via heat-extraction (see Appendix G).

The heat-extraction method described by Wiles and Krogh (1998) and OECD (2009) is based on the principles of MacFayden and of Petersen, and involves a controlled temperature gradient extractor, where the organisms are collected over a 48-h period.[42] Becker-van Slooten *et al*. (2005) developed a simpler and more cost-effective heat-extraction technique. This method, which was then further refined by Environment Canada (EC, 2014a; ECCC, 2018) using equipment available in Canada, is recommended for the extraction of mites from test soil. The heat comes from a lamp fitted with a 60-watt lightbulb, and is regulated by the distance of the lightbulb from the surface of the soil in the heat-extraction unit.[43] One heat-extraction unit should be prepared for each test vessel. At test termination, the soil from each test vessel is transferred into a heat-extraction unit. Soil from a test vessel can be transferred to a heat-extraction unit by pouring the test soil from the test vessel into the open end of the heat-extraction unit containing the cheesecloth and/or mesh. The bottom of the test vessel should be tapped several times to dislodge any soil stuck to the sides and bottom. Any live mites attached to the empty test vessel must be recorded and combined with the final count data. Water can be used to float and count any mites still remaining in the test vessel (see Appendix G). The soil surface is gently smoothed out evenly over the mesh using a spoon or a scoopula. The heat-extraction units are placed

---

[42] In the heat-extraction method described by Wiles and Krogh (1998), the heat comes from a heating element at the top of an extraction box (regulated through a thermistor placed on the surface of the soil sample). The temperature in the cooled liquid surrounding the collecting vessel is regulated through a thermistor situated at the surface of the collection box (placed below the soil). The thermistors are connected to a programmable controlling unit that raises the temperature according to a pre-programmed schedule (i.e., the soil is gradually heated from 25°C to 40°C at a rate of 5°C every 12 h). The organisms are collected in a cooled collecting vessel (2°C) with a plaster of Paris / charcoal layer at the bottom.

[43] The heat-extraction unit consists of two plastic cups (e.g., Fisher cat #11-838-17), one of which has ~1 cm cut off of the bottom, and the other ,which has ~1 cm of plaster of Paris substrate (see Section 2.3.5) on the bottom. A piece of plastic canvas (used for needlework; 7 mesh) is cut to size and glued (with a hot glue gun and non-toxic glue sticks) into place ~1 cm below the top edge (not the cut edge) of the cup that has had the bottom removed. The heat-extraction unit is assembled by placing the cut cup (i.e., with the mesh insert) upside down on top of the whole cup (i.e., with the plaster of Paris substrate on the bottom) so that the two widest parts (i.e., the original top of each cup) of both cups meet (i.e., the cup with the mesh insert is inverted on top of the cup containing the plaster of Paris). A piece of Parafilm® should be wrapped around the seam between the two cups and secured with a piece of tape, if necessary. A small square of cheesecloth can be placed onto the mesh so that soil does not drop through to the bottom portion of the extraction unit (Appendix G).

underneath the lamps, limiting the number of units per lamp to no more than five or six, so that the heat and light are kept consistent for each unit. The bottom of the lightbulb is adjusted to ~25 cm above the top of the soil and a thermometer (e.g., electronic thermometer) is set up within one of the units (i.e., one thermometer per lamp) to monitor temperature changes throughout the extraction. The temperature should be checked twice per day, or more frequently if necessary, along with the hydration of the soil (with lamps being raised during periods when technicians are not present to monitor the units, e.g., overnight, to prevent the soil from drying completely). The lamp height does not need any adjustment, and the temperature should reach ~32°C after 48 hours. Test water should be added (e.g., as a gentle mist over the surface of the transferred test soil) as required to ensure the test soil does not completely desiccate, but no further wetting of the soil should occur in the 16 h prior to the end of the extraction. At the end of the extraction period (i.e., 48 h), the lamp is turned off and the Parafilm® connecting the two halves of the heat-extraction unit is removed. The organisms that have dropped down through the mesh to the plaster of Paris substrate should be counted using flotation (see Appendix G). They can be counted immediately, either manually or through image analysis. If the organisms cannot be counted right away, the substrate can be moistened with deionized water, and placed in cool storage (e.g., 4°C–10°C) until processing can continue, or they can be preserved (e.g., in 70% alcohol) for enumeration at a later date. Details on the construction of the heat-extraction units and performing the heat-extraction are provided in Appendix G. Other heat-extraction methods may be used, and two alternative procedures are also described in Appendix G.

Laboratories that are not experienced with the heat-extraction procedure described must initially validate and document the efficiency of their heat-extraction system (i.e., demonstrate and record data that show that a significant number of test organisms are not being left in the soil following heat-extraction). This can be accomplished by further processing the heat-extracted soil for test organisms using a flotation method[44] to check on the efficiency of the heat-extraction technique. The heat-extraction process is considered acceptable if there are < 5% of the total number of test organisms remaining in the soil (i.e., extracted from the soil using flotation, following heat-extraction). If the heat-extraction efficiency is not acceptable, all treatments must be processed in a similar matter (i.e., using flotation following heat-extraction). Once laboratory personnel are experienced with heat-extraction and have demonstrated the efficiency of their system, they should continue *monitoring* the efficiency periodically.

All mites (adults and progeny) extracted from the soil are counted and recorded, regardless of whether they are dead or alive. Juvenile mites (i.e., progeny) must be distinguished from the adults and counted separately. In general, adults can be easily distinguished from their progeny by their significantly greater size and darker colouration (see Appendix E). Test organisms should also be observed for short period of time to detect movement or a lack thereof. A mite is considered dead if there is complete cessation of movement of any type of body part, including legs, abdomen, head, and antennae. Dead *O. nitens* tend to have their legs curl underneath them and might, but generally do not, change colour. Care should be taken to distinguish dead adults from molted carapaces; the latter are translucent and collapsed. If dead adults or progeny are observed following heat-extraction, they should be noted and must be included in the "survival" counts and endpoint calculations,[45] whereas dead mites observed prior to

[44] The flotation method described in EC 2014a can be used to check the efficiency of the heat-extraction technique. For this method, the heat-extracted soil is transferred into a glass petri dish. Deionized water is added to the petri dish and the slurry is swirled or stirred lightly with a glass stir rod. The mites, both adult and progeny, float to the surface, and the dish is placed under a microscope for enumeration.

[45] Following heat-extraction from the soil, all adult mites and their progeny observed in the collection vessel (i.e., lower portion) of the heat-extraction unit, regardless of whether they are alive or dead, are included in the count for adult survival and the endpoint calculation for surviving progeny production. The rationale for this is that during heat-extraction, the mites need to be alive in order to move through the soil and into the lower chamber of the heat-extraction unit. Any organisms that survived

heat-extraction must not be included in the final survival counts. Any missing adult mites are considered to be dead and must not be included in the survival count, assuming that such mites have died and decomposed prior to the extraction.

Test vessels, irrespective of concentration levels, should be processed in a random manner since counting might become more or less accurate. Extra replicates of each test soil (including the negative control soil and, if included in the test, reference soil) set up for the purpose of physicochemical analyses must be analyzed to determine the pH and moisture content at the end of the test (Section 4.6). Analyses for other chemical constituents (i.e., concentrations of contaminants) should also be made at this time using additional replicates prepared for each test soil (Section 4.6).

4.8 Test Endpoints and Calculations

For each test, the percent survival of adult mites in each test vessel at the end of the test must be calculated. The mean (± SD) percent adult survival for all mites exposed to each concentration (including the negative control soil and, if used, reference soil) must be calculated and reported at the end of the test using the survival data determined from all treatment replicates (e.g., the mean of the replicates within each treatment).

The reproductive endpoint for this test is based on the number of surviving progeny produced in each replicate and each treatment during the 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 mites. The mean (± SD) number of surviving progeny in the test soil on Day 28 must be determined and reported for each treatment, including reference and negative control soils.

The two most common possibilities for a typical test design involve:

i) Multiple *sampling locations*, in which responses at one or more test site sampling locations are compared with those at a *reference site* sampling location,[46] with other test sampling locations, or with the control soil (i.e., single-concentration test). Hypothesis testing is frequently used in the statistical assessment, and the common outcome is that a response at a sampling location is either "different" or "not different" from another sampling location (Section 5.6.1).

ii) Multiple concentrations of a test soil, achieved by mixing a test soil with reference or control soil (Section 5.3), or by *spiking* a soil with various concentrations of a chemical or chemical product (Section 6.2). For a multi-concentration test, the 28-day ICp for reproductive inhibition must be calculated and reported (data permitting).[47]

---

the soil exposure will therefore move through the soil to the collection vessel below. Thereafter, the test organisms in the heat-extraction collection vessel might become desiccated and die; or if test organisms are collected and preserved (e.g. in alcohol) prior to enumeration, distinguishing between live and dead organisms would not be possible.

[46] Throughout this document, reference site is used to describe an area in which there is clean soil uninfluenced by the contaminant under study (i.e., reference soil). A reference soil should be collected for these comparisons, as described in Section 5. However, in the absence of a reference soil, a negative control soil may be substituted.

[47] Historically, investigators have 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, 2005b). Given these disadvantages, ICp is the required statistical endpoint for reproduction data derived in a multi-concentration test using mites. Contrary to recent criticism blaming the continued generation and publication of NOEC/LOEC data on the failure of governments and international organizations to formally discredit and cease recommending these approaches (van Dam *et al*., 2012), it is evident that Environment and Climate Change Canada has fully adopted regression-based methods in aquatic-, sediment-, and soil-based environmental toxicity testing (EC, 2004a; 2005a, b; 2007a, b; 2011a, b; 2014a; Van der Vliet *et al*., 2012).

In a scenario where there are multiple sampling locations, an understanding of the strengths of various study designs is critical for the successful application of statistical tests. The study objectives should be clearly defined before data are collected, with an appreciation both for the power (ability to detect an effect) of the test design and the ease of interpretation of the results. In general, it is advantageous to limit the number of comparisons made, and this is typically done by choosing a test design and statistical tests that compare test sampling locations with a reference sampling location. Further gains in power can be made if a gradient can be assumed (i.e., samples collected in sequential order away from the point source; see Section P.4 in EC, 2005b). In some cases, study objectives and test design might not have been given adequate attention before the collection of the data, and to compensate, investigators will perform a comparison among all possible sampling locations, maximizing the number of comparisons made. This is strongly discouraged, particularly when large numbers of sampling locations are involved, because undesirable effects on *Type I* and *Type II error* rates might occur; interpretation of results is often more difficult; and unwarranted focus might be given to particular comparisons after data have been collected.[48] Detailed statistical guidance on hypothesis testing for the number of progeny at test end is provided in Section 5.6 and EC 2005b.

Environment Canada (2005b) provides direction and advice for calculating the ICp endpoint, which should be followed; Section 4.8.1 gives further guidance in this regard. Initially, regression techniques (see Section 4.8.1.1) must be applied to multi-concentration data intended for calculation of an ICp.[49] In the event that the data do not lend themselves to calculating the 28-day ICps for the reproductive inhibition using the appropriate regression analysis, linear interpolation of these data using the program ICPIN should be applied in an attempt to derive an ICp (see Section 4.8.1.2). Although the reproductive endpoint is the biological endpoint of interest for this method, situations might arise where effects on adult survival warrant the calculation of a *lethal* endpoint (e.g., LCp). If such an occurrence arises with the test data, EC (2005b) may be consulted for the appropriate analyses of lethality data.

An initial plot of the raw data (number of progeny) against the logarithm of concentration is highly recommended, both for a visual representation of the data and to check for reasonable results by comparison with later statistical computations. Any major disparity between the approximate graphic ICp and the subsequent computer-derived ICp must be resolved. The graph would also show whether a logical relationship was obtained between log concentrations (or, in certain instances, concentration) and effect, a desirable feature of a valid test (EC, 2005b).

4.8.1 ICp

When a multi-concentration test for effects of exposure of mites to field-collected or spiked-soil mixtures is conducted, the *quantitative* data representing reproductive inhibition must be used to calculate the ICp (see introductory paragraphs of

[48] Zajdlik & Associates Inc. (2010) made this last point in the defence of the application of an overall test for significance: "All too often an observed difference catches the eye of the data analyst and a search begins to apply a statistical test to 'validate' the observed difference. This is an example of data snooping; conclusions made using this data analytic approach are suspect." This same flaw is apparent in poorly defined study designs, as described here.

[49] Regression is the method of choice for estimating an 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. Non-linear regression techniques were originally recommended by Stephenson *et al.* (2000) for several reasons, such as the relationship that exists between exposure concentration and mite reproduction responses is typically non-linear, the *heteroscedasticity* of the data is rarely reduced by transformation, the more standard bootstrap simulation technique has several limitations for these types of data, and non-linear 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, 2005b).

Section 4.8, and Section 6.2). The ICp is a quantitative estimate of the concentration causing a fixed percent reduction in the mean number of progeny produced by the adult mites during the test.

The ICp is calculated as a specified percent reduction (e.g., the IC25 and/or IC20, which represent 25% and 20% inhibition, respectively). 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 surviving progeny produced in each replicate must be used to calculate the average number of surviving progeny produced per treatment (concentration) in relation to the average number produced in the *negative control* replicates. A value of zero is assigned for the number of juveniles in a replicate, if all of the adult mites in that replicate died before producing progeny. If any of the adult mites died during the test, after producing young, the number of surviving progeny produced is still to be used in the analyses. If there are no surviving progeny in a replicate (test vessel), 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 progeny in all replicates at a given concentration, that concentration is still included in the analysis, using an average value of zero juveniles.

As previously indicated, an ICp for mean number of surviving progeny produced in each treatment must be calculated and reported (data permitting) upon completion of a 28-day multi-concentration test with *O. nitens*. These calculations must be made using the appropriate linear or non-linear regression analyses (see the following Section 4.8.1.1). If, however, regression analyses fail to provide a meaningful ICp for the mean number of progeny produced, the ICPIN analyses described in Section 4.8.1.2 should be applied to the corresponding data. Any procedures applied to the data, details regarding any transformation of the data, and the statistical method used for the calculation of ICp must be reported.

4.8.1.1 Use of regression analysis

Upon completion of a definitive 28-day multi-concentration test, an ICp (including its 95% confidence limits) for the mean number of surviving progeny produced in each treatment must be calculated using regression analysis, provided that the assumptions below are met. A number of models are available to assess reproduction data (using quantitative statistical tests) via regression analysis. The proposed models for application consist of one linear model, and the following four non-linear regression models: exponential, Gompertz, logistic, and logistic adjusted to accommodate *hormesis*[50] (see Section 6.5.8 in EC, 2005b). Use of regression techniques requires that the data meet assumptions of *normality* and *homoscedasticity*. The reader is strongly advised to consult EC (2005b) for additional guidance on the general application of linear and non-linear regression for the analysis of quantitative toxicity data.[51]

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 2. 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-

[50] A hormetic response (i.e., hormesis) might be observed at one or more of the lowest, sublethal concentration(s), i.e., performance at such concentration(s) is enhanced relative to that in the negative control (see Section 10.3 in EC, 2005b). For instance, there might be more progeny produced in soil with low concentrations than in the control treatment. This is not a flaw in the testing. Rather, 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.

[51] Some of the specific guidance provided in EC (2005b) refers to the use of a general purpose statistical package (i.e., SYSTAT); however, CETIS (a software package designed for environmental toxicology) contains the models described herein for regression analysis. The latest version of SYSTAT is available for purchase by contacting SYSTAT Software, Inc.; see website www.systatsoftware.com/products/Systat. The latest version of CETIS is available for purchase by contacting Tidepool Scientific Software; see website www.tidepool-scientific.com/Cetis/Cetis.

response curve. The shape of the curve is then compared with available models so that one or more appropriate models that best suit the data are selected for further examination (refer to Figure O.1, Appendix O, in EC, 2005b 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 Section 10.2 of EC (2005b). 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.[52] 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 (2005b). 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 is dependent upon 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 perform the less-desirable linear interpolation (using ICPIN, see Section 4.8.1.2) method of analysis.

Homoscedasticity of the residuals should be assessed using Levene's test as described in EC (2005b), 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 O.2A of Appendix O in EC, 2005b) 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 O.2B, Appendix O in EC, 2005b for an example), then the data analysis should be repeated using weighted regression. Traditionally, the data have been weighted by dividing by the inverse of the variance; however, other options are available. Before choosing the weighted regression, the standard error of the ICp is compared with that derived from the unweighted regression.

If there is a difference of greater than 10% between the two standard errors,[53] 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 reanalyzed using the less-desirable linear interpolation (using ICPIN, see Section 4.8.1.2) 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 O.2C, Appendix O in EC, 2005b, for an example), and the user is again urged to consult a

---

[52] The Akaike Information Criterion (or an equivalent, such as the Bayesian Information Criterion) is another option for determining best model fit.

[53] 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 might introduce additional complications to the modelling procedure. A statistician should be consulted when weighting is necessary.

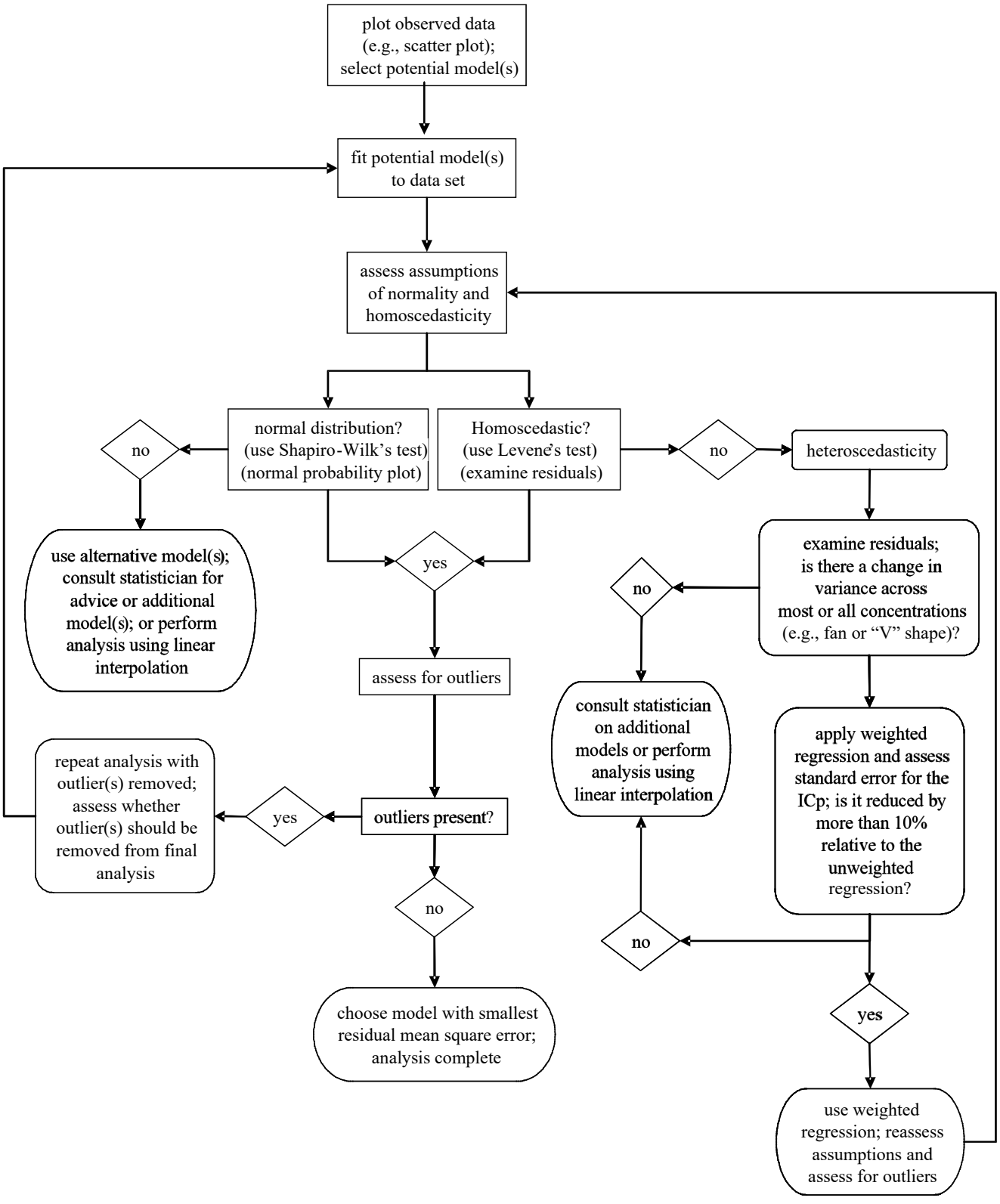


Figure 2 **The general process for the statistical analysis and selection of the most appropriate model for quantitative toxicity data** (adapted and modified from Stephenson *et al*., 2000)

statistician for further guidance on the application of additional models.

Endpoints generated by regression analysis must be bracketed by test concentrations; extrapolation of endpoints beyond the highest test concentration is not an acceptable practice (EC, 2005b).

4.8.1.2 Linear interpolation using ICPIN

If regression analyses of the endpoint data (see preceding Section 4.8.1.1) fail to provide an acceptable ICp for reproductive inhibition (i.e., assumptions of normality and homoscedasticity cannot be met), linear interpolation using the computer program called *ICPIN* should be applied. This program (Norberg-King, 1993; USEPA, 1995, 2002) is not proprietary and is included in most computer software for *environmental toxicology,* including TOXSTAT (1996) and CETIS. The original instructions for ICPIN from the USEPA are clearly written and make the program easy to use (Norberg-King, 1993).[54] 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 the data as necessary, then using the two data points adjacent to the selected ICp (USEPA, 1995, Appendix L; USEPA, 2002, Appendix M). 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, 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 may range from 80 to 1000. At least 400 is recommended here, and 1000 would be beneficial.[55]

If there are several adjacent high concentrations with no surviving progeny, 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).

Besides determining and reporting the computer-derived ICps for mite reproduction at test end, a graph of percent reduction in number of progeny produced 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, 2005b).

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 (see Option 4, Section 10.3.3 in EC 2005b). The correction is applied for any test concentration in which the average effect (i.e., the *geometric mean* of the replicate means) is higher ("better") than the average for the control. To apply this correction, replace the observed mean numbers of progeny of the replicates in the hormetic

---

[54] 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 of individual organisms within the same vessel. 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.

[55] ICPIN has some deficiencies, which is why it is recommended herein only in cases where the 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 the 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 more accurate methods of estimating the ICp and its 95% confidence limits (EC, 2005b) (see Section 4.8.1.1).

concentration(s) with the means of replicates in the control. The geometric mean for that/those concentration(s) will then be the same as that for the control.

4.9 Tests with a Reference Toxicant

The routine use of a reference toxicant is used to assess, under standardized test conditions, the relative sensitivity of a portion of the population of adult mites within a particular culture (Section 2.3.9) 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 for that reference toxicant, under standardized test conditions, as well as the technical proficiency of the laboratory staff conducting the test (EC, 1995). Testing with a reference toxicant, conducted according to the procedures and conditions described herein, must be performed according to one of the following two regimes:

i) multi-concentration reference toxicity test at least twice per year[56] using organisms taken from the population of mites that is being cultured for use in the definitive test(s) (Section 2.3); or

ii) a positive control concentration run concurrently with each test (Section 2.3.9 and Appendix H).

A laboratory that chooses to monitor the sensitivity of its culture(s) to a reference toxicant in a multi-concentration reference toxicity test should conduct these tests at least once every six months. Reference toxicity tests may be run concurrently with a definitive soil toxicity test using organisms from the same age-synchronized culture, if the number of age-synchronized organisms allows.

Described herein are the procedures and conditions to be followed when performing multi-concentration reference toxicity tests in conjunction with a 28-day test of soil toxicity using *O. nitens*. These procedures also apply to tests for assessing the acceptability and suitability of cultures to be used in soil toxicity tests. They should be applied to assess intralaboratory precision when a laboratory is inexperienced with the biological test method defined in this document and during initial test setup (see Sections 2.3.1, 2.3.9).

For the first option of testing with a reference toxicant, a reference toxicity test must be conducted as a *static* multi-concentration test using the reproductive endpoint. The test conditions and procedures described herein for performing a 28-day reproductive test must be applied to each reference toxicity test. Additional conditions and procedures described in Section 4 for performing a multi-concentration test with samples of test soil apply equally to each reference toxicity test. 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 multi-concentration reference toxicity test must be performed using the same test vessels as those used for definitive tests (Section 3.2.2), with the same volume of soil (i.e., ≥ 3 cm soil; Section 4.1) at optimal moisture content. The number of replicate test vessels per reference toxicant concentration and negative control soil must be ≥ 5. The number of mites per test vessel must be 15 as described in Section 4.2.

Procedures for starting and ending a reference toxicity test must be consistent with those described in Sections 4.2 and 4.7. Test conditions described in Section 4.3 must be applied. Test organisms must be fed as described in Section 4.5. Test observations and measurements given in Section 4.6 must be followed.

The validity criteria for reference toxicity tests are the same as those described for definitive tests (see

[56] Environment and Climate Change Canada typically includes monthly reference toxicity tests as the option for routine testing (EC, 2004a); however, due to the age-synchronization process required in this test method, the number of organisms for testing that are available each month is limited.

Section 4.4). Results for a reference toxicity test should be expressed as mg reference chemical/kg soil, dry weight.

Appropriate criteria for selecting the reference toxicant to be used in conjunction with a definitive test for soil toxicity using mites 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.

The reference toxicity test requires a minimum of six treatments (i.e., negative control soil and five concentrations of reference toxicant). Reagent-grade boric acid (H_3BO_3) is recommended for use as the reference toxicant when performing soil toxicity tests with mites, although other chemicals may be used if they prove suitable. Each test concentration should be made up according to the guidance in Sections 4.1 and 6.2, using artificial soil (Section 3.3.2) as the substrate.

Routine reference toxicity tests (e.g., those performed twice per year) using boric acid (or another suitable reference chemical) spiked in negative control soil should consistently apply the same test conditions and procedures described herein. A series of test concentrations should be chosen based on preliminary tests, to enable calculation of a 28-day IC50 (see Section 6.4).

Environment and Climate Change Canada's Biological Assessment and Standardization Section is introducing the use of positive control replicates, included with each definitive toxicity test, as an alternative to routine multi-concentration reference toxicity testing. As such, the second option for testing with a reference toxicant offered herein is to include replicates of a single-concentration of a known toxicant, which elicits a consistent partial response, with each definitive test to serve as a positive control. Positive controls are defined as an exposure of test organisms to conditions similar to a negative control (i.e., same number of replicates, number of organisms per replicate, vessels, test conditions, etc.) except exposed to a single concentration of a known toxicant. This option could be more feasible and practical for longer term sublethal- and lifecycle-type toxicity tests, such as the 28-day reproduction test with *O. nitens*, described in this test method document.

If chosen, the traditional multi-concentration reference toxicity test is required to be conducted twice per year. The alternative, however, is to run replicates of a positive control concurrently with every definitive toxicity test conducted. This approach could have several advantages: it is economical (reduced effort and resources); it reflects a response by organisms sub-sampled from the lot (group) used for testing; and it can measure the same endpoint(s) in the same matrix and duration as the definitive test, especially for longer, sublethal soil toxicity tests).

The choice of toxicant for the positive control concentration should be made using the same selection criteria as those used for a multi-concentration reference toxicity test and reagent-grade boric acid (H_3BO_3) is recommended herein. A single concentration known to elicit a consistent partial response must be used (as compared with traditional reference toxicity tests conducted using multiple concentrations to capture a range of effects, e.g., complete lack of reproduction to no effect on reproduction). The positive control replicates must be prepared using the same test vessels as those used for definitive tests (Section 3.2.2), with the same volume of soil (i.e., ≥ 3 cm soil; Section 4.1) at optimal moisture content. The number of replicate test vessels per positive control sample must be ≥ 5. The number of mites per test vessel must be 15, as described in Section 4.2. The positive control concentration should be made up according to the guidance in Sections 4.1 and 6.2 using artificial soil (Section 3.3.2), and the procedures and conditions for testing must be consistent with those used in the definitive test, as described in Sections 4.2 to 4.7. For the positive control option, the required endpoint is the mean response (i.e., number of progeny produced) in the positive control concentration subtracted from the mean in the negative control, divided by the mean negative control response and multiplied by 100 to provide a percent inhibition

(see Appendix H).

If selecting this option, the positive control response (i.e., target effect level) must be defined and include acceptability limits for each endpoint. Acceptability limits for the purposes of this method are synonymous with *warning limits* and must be operationally defined at each laboratory with variability limits that are fit for purpose. For example (see Appendix H), a laboratory might define for its positive control that boric acid (e.g., 95 mg H_3BO_3/kg dry soil) must produce a 41% inhibition of progeny production (i.e., target effect level) that falls in between calculated warning limits (i.e., ≥ 27% and ≤ 56%), with a coefficient of variation (CV) of response over time of ≤ 30%. Keeping in line with currently required multi-concentration reference toxicity test results, the results of an individual positive control test are not to be used to determine the acceptability of the corresponding test result (i.e., as test validity criteria), but rather can be used to monitor consistency over time (i.e., similar means among positive control tests) and precision over time (i.e., overlapping ranges among positive control tests). Identifying outliers in test organism response or extreme variability in response for individual tests must be used to trigger investigations into potential causes such as culture sensitivity, culture health, environmental/facility conditions, and technician performance. Data obtained from negative controls, positive controls, and culture health data should be monitored over time (i.e., by trend analysis) to proactively indicate changes in the organism response. Appendix H provides an example of how to choose a positive control concentration for this test and how to derive warning limits.

For both multi-concentration reference toxicity tests and positive controls, once sufficient data (e.g., minimum of five data points) are available (EC, 1995, 2005b), all comparable endpoints (i.e., IC50s for a particular reference toxicant derived from multi-concentration reference toxicity tests, or percent reduction of progeny production relative to control for a single concentration of reference toxicant tested as positive controls) must be plotted successively on a *warning chart*. For multi-concentration reference toxicity tests, the warning chart should plot logarithm of concentration on the vertical axis against date of the test or test number on the horizontal axis. For positive control concentrations, the warning chart should plot the percent reduction in response on the vertical axis against the test date or test number on the horizontal axis (Appendix H). Each new data point for the reference toxicant should be examined to determine whether it falls within the warning limits (± 2 SD of values obtained in previous comparable tests using the same reference toxicant and test procedure) (EC, 2004a, 2005a, 2013a, 2014a, Appendix H). A separate warning chart must be prepared and updated for each dissimilar procedure (e.g., differing reference toxicant) and endpoint. Each new data point for the reference toxicant should be compared with the established limits of the chart; the reference toxicant result is acceptable if it falls within the warning limits.

For multi-concentration reference toxicity tests, the logarithm of concentration (including IC50) must be used in all calculations of mean and standard deviation, and in all plotting procedures. This represents continued adherence to the assumption by which each IC50 was estimated based on the logarithms of concentrations. The warning chart can be constructed by plotting the mean and ± 2 SD as the logarithms, or by converting them to arithmetic values and plotting them on a logarithmic scale of concentration. Different approaches to creating a warning chart (e.g., Levey-Jennings, moving average) are acceptable. For positive control concentrations, the warning chart can be constructed by plotting the mean and ± 2 SD for percent reduction in reproduction relative to the control on an arithmetic scale.

The mean of the available endpoint values, together with the upper and lower warning limits (± 2 SD), should be recalculated with each successive endpoint for the reference toxicant until the statistics stabilize (EC, 1995, 2004a, 2005a, 2013a, 2014a, Appendix H). Warning charts can be used to detect trends over time. Examples of trends that might be observed include an increasing or decreasing trend, several successive points on one side of the mean, changes that are observed at different times of the year, and successive data points outside the ± 2 SD warning limits. If a particular data point fell outside the warning limits, the sensitivity of the test organisms, and the performance and precision of the test are suspect. Since this might occur 5% of the

time due to chance alone, an outlying data point would not necessarily indicate abnormal sensitivity of the mite culture, nor unsatisfactory precision of toxicity data. Rather, it provides a warning that this might be the case. A thorough check of all culture and test conditions and procedures, as well as technical proficiency, is required at this time. Depending on the findings, it might be necessary to repeat the reference toxicity test or positive control concentration, establish a new culture, select mites from an alternate culture, or obtain a new population of test organisms from an outside source before undertaking further soil toxicity tests.

Results that fall within the warning limits do not necessarily indicate that a laboratory is generating consistent results. A laboratory that produced extremely variable historical data for a reference toxicant would have wide warning limits; a new datum point could be within the warning limits but still represent an undesirable variation in results obtained in the test. A coefficient of variation (CV) of no more than 30%, and preferably 20% or less, has been suggested as a reasonable limit by Environment Canada (EC, 1995, 2005b) for the mean of the available values of log(IC50) (see preceding paragraph). For this biological test method, the CV for mean historical data derived for reference toxicity tests or positive controls performed using boric acid should not exceed 30%.

If an IC50 or positive control result fell outside the control limits (mean ± 3 SD), it would be highly probable that the test was unacceptable and should be repeated, with all aspects of the test being carefully scrutinized. If endpoints fell between the control and warning limits more than 5% of the time, a deterioration in precision would be indicated, and again the most recent test should be repeated with careful scrutiny of procedures, conditions, and calculations.

Concentrations of the reference toxicant (including single concentrations used as a positive control) in all stock solutions can be measured chemically using appropriate methods (e.g., analytical methods involving AES with ICAP scan, for the concentration of boron). Test concentrations of the reference toxicant in soil are prepared by adding a measured quantity of the stock solution to negative control soil,[57] and mixing thoroughly.[58] 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, or from the single concentration used for a positive control.[59] Each aliquot should either be analyzed directly, or stored for future analysis (i.e., at the end of the test) if the 28-day IC50 or positive control response 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 testing with the reference toxicant. The 28-day IC50 (for multi-concentration reference toxicity tests) or % reduction in response relative to

---

[57] Section 6.2, Preparing Test Mixtures, includes an example showing the amounts of test water and boric acid to be added to artificial soil to prepare a given treatment for a reference toxicity test or a positive control 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.

[58] 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 deionized water, and then adding a measured volume of this mixture to the soil. The flask is then rinsed three times with deionized 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.

[59] If the IC50 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 the IC50 for each test is based on nominal concentrations, however, sampling and analysis of aliquots from at least the low, middle, and high test concentrations is recommended.

the control (for positive control concentrations) 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 for a reference toxicity test or for a positive control, 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 a 0.01 *M* solution of $CaCl_2$ by boiling a slurry of soil in 50 mL of this extraction solution and then readjusting 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 boron using ICAP/AES. The boric acid concentration in the soil is then calculated using the following equation:

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

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

Section 5

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 on the collection, handling, transport, storage, and preparation of field-collected soil for biological testing is given in Environment Canada's *Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing* (EC, 2012). General procedures are outlined therein for the preparation of collecting soil samples, including: developing study objectives; identifying the study area; collecting background data; conducting site surveys, soil surveys, and ecological land classifications; selecting sampling strategies and locations; determining the size and number of samples to collect; establishing proper *quality assurance* and *quality control* (QA/QC) procedures; considerations for environment, health, and safety; and developing sampling plans. Guidance is also provided for soil collection, including: selecting sampling devices; collecting soil samples by horizon or by depth; handling soil samples on-site; selecting sample containers; and transporting samples. Procedures for personnel receiving, preparing (i.e., drying, wetting, sieving, grinding, homogenizing, reconstituting, and characterizing), and storing soil samples for biological testing at the laboratory are also described in EC (2012). Additional procedures and considerations are included that are specific to the nature of the contaminants (i.e., soils contaminated with volatile or unstable contaminants), biological testing requirements, and study objectives. Specific guidance is provided for sampling, handling, transporting, storing, and preparing soil from boreal forest, taiga, and tundra ecozones, as well as organic and wetland soils. Environment Canada's soil collection guidance document (EC, 2012) should be consulted, and the guidance therein followed (in addition to the guidance provided here), when collecting samples of field-collected soil and preparing them for toxicity tests with mites using the biological test method described herein.

5.1 Sample Collection

Environment Canada (2012) provides substantial guidance on field-sampling design and appropriate techniques for sample collection. The guidance provided therein assumes that some data on the characterization of the chemical and soil properties of the land under investigation are already available. Field surveys of soil toxicity using biological tests with suitable soil-associated test organisms (e.g., EC, 2004a, 2005a, 2013a, 2014a) are frequently part of more comprehensive land assessments and remediation (Stephenson *et al*., 2008; EC, 2012). Such assessments often include a *battery of toxicity tests* to evaluate the toxicity of soil using more than one test type and test species in conjunction 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. This integrated approach can provide more accurate information on the risk associated with soil contamination in ecological risk assessments and contaminated land management (EC, 2012). Statistical correlation in these assessments 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 the biological test method described herein (Section 4) might be collected quarterly, semi-annually, or annually from a number of contaminated or potentially contaminated sites for monitoring and *compliance* purposes. Soil samples 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. Increasingly, biological (toxicity) testing is being used at all levels (i.e., Tiers) of risk assessment. Depending on the specific objectives of the assessment and the conditions at a contaminated site, site-specific toxicity data can be used in a number of ways, including:

- to screen soil at a site to locate highly toxic or sublethally toxic areas;

- to identify site soil (determine concentration of contaminant in a site soil) that has a toxic impact;
- to evaluate contaminated soil for lethal or sublethal toxic effects;
- to identify soil characteristics that modify bioavailability;
- to derive (in part) site-specific standards and/or remedial objectives;
- to identify the efficacy of bioremediation technologies and/or site remediation; and
- for long-term monitoring of a remediated site (EC, 2012).

Further guidance on the application of biological testing in contaminated soil assessment is provided in EC (2012).

Environment Canada (2012) provides extensive guidance on defining study objectives and developing a study plan that incorporates biological testing into contaminated land assessments and management. A study plan provides specific guidance on the methods and strategies for sample collection and the procedures required to ensure that all *data quality objectives* (DQOs) are met. Information incorporated into a study plan includes: identification of DQOs; definition of the study area; collection of background data; selection and location of sampling; selection of sampling strategies; QA/QC; and considerations for environment, health, and safety. The sampling strategy (i.e., the process by which the type, location, and collection method of samples is determined) is driven primarily by the study objectives and secondarily by the site characteristics, and is discussed in detail in EC (2012).

The number of locations to be sampled at a study site and the number of replicate samples per location will be specific to each study. The number of samples to collect depends upon the study objectives, the data quality objectives, the desired level of certainty, and site-specific considerations. The number of sample replicates required further depends on the experimental design of biological tests and, in most cases, logistical and budgetary constraints (e.g., time and cost). Various types of samples (i.e., *point*, *composite,* and *bulk*) might be collected depending on the study objectives.

The majority of samples collected for biological testing are *unconsolidated samples* in which particles become loosened and separated in the sampling process. *Consolidated samples* are those collected such that the soil particles and pore structure remain unaltered (i.e., *cores*). Guidance on the collection of consolidated samples for biological testing is provided in EC (2012); however, this biological test method document and the guidance provided herein apply primarily to the use of *unconsolidated soil samples*.

Specific procedures for the collection, handling, and preparation of soils contaminated with volatile or unstable compounds are described in EC (2012), and include modifications to procedures for sample collection, transport, storage, preparation, and contaminant analyses. All of the procedures described therein should be applied in order to minimize the loss of contaminants when sampling and handling soils in the field, transporting soils to the toxicity laboratory, and any further loss of these contaminants in the laboratory prior to testing (i.e., during sample storage, handling, or preparation). Environment Canada's soil sampling guidance document (2012) also addresses issues related to QA/QC.

For certain monitoring and regulatory purposes, multiple replicate samples of soil (i.e., five field replicates or separate samples from different point or bulk samples taken at the same location) should be taken at each sampling location, including one or more reference location(s). These replicate samples[60]

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

provide information about the variability of the toxicity/bioavailability of the contaminants at the location and allow for statistical comparisons of soil toxicity among more than one location (EC, 2005b). Each of these "true replicate" samples of soil may be tested for its toxicity to mites as a single laboratory replicate (i.e., using only one test vessel per replicate sample) or as multiple laboratory replicates (i.e., using more than one test vessel per replicate sample; see Section 5.6.1). The use of power analysis (see Section 5.6.2) with endpoint data obtained in previous tests of the same type, performed with previous samples from the same or similar sites, will assist in determining the number of field and/or laboratory replicates that need to be tested. Also, some of the statistical tests have requirements for a minimum number of replicates. For certain other purposes (e.g., preliminary study or extensive surveys of the spatial distribution of toxicity), the survey design might include only one replicate sample (i.e., field replicate) from each location, in which case the sample (including reference and/or control soils) must be homogenized and split between five replicate test vessels (i.e., laboratory replicates).[61] The latter approach precludes any determination of mean toxicity at a given sampling location, and completely prevents any conclusion on whether a sampling location 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, using appropriate statistical tests (see Section 5.6.1). It is important to realize that any conclusion(s) about differences, which arise from testing single field samples lacking field replication, must not be extended to make any conclusion(s) about the sampling locations.

Regardless of the study objectives, one or more sites should be sampled for *reference* (presumably clean) *soil* during each field collection (see Section 3.5).[62] 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. Some of the most critical soil physicochemical properties that should be matched between the reference and contaminated soils include: particle size distribution, total organic carbon content (%), organic matter content (%), pH, and electrical conductivity. In addition, other properties to match might include CEC, total inorganic carbon, *redox potential*, and water-holding capacity (EC, 2012). Matching of total organic carbon content (%) or organic matter content (%) might not be warranted in cases where *pollution* (e.g., from or within sewage

---

[61] Power analysis carried out on reproduction data generated using this method (Section 5.6.2) indicated that, for five laboratory replicates, an effect size of 40% or greater can be reliably detected (power ≥ 80%). To detect a 30% effect size with the same power, eight replicates are recommended. More replicates may also be set up to meet specific study objectives, such as those defined for Phase I (i.e., site soil screening tests) in the recommended framework for toxicity assessments in support of the development of site-specific remediation objectives for petroleum hydrocarbons in soil (ECASG, 2006). This framework for toxicity assessment of contaminated lands is divided into two phases, the first of which includes site soil screening tests using undiluted soil samples representative of the study site. The purpose of the screening tests are to: 1) quickly determine if there is toxicity associated with short-term (*acute*) exposure of the test organisms to the site soil; and 2) if there is no *acute toxicity*, continue the test to assess for *chronic toxicity* associated with prolonged exposure to the site soil. An investigator, therefore, might choose to expand the test design for the single-concentration tests described in this test method document by setting up extra replicates to look for potential acute responses (i.e., adult mortality) early in the test. This approach serves only to judge the potential of an acute response, but is not suitable for defining remedial or cleanup objectives. Phase II of the proposed framework uses multi-concentration tests to determine the magnitude of the toxicity. As described in Section 4.1 of this test method document, a range-finding test can be useful, and is recommended in the framework, for determining the range of effect concentrations (i.e., narrow the range of concentrations to be used in a definitive sublethal test).

[62] 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, and 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 due to anthropogenic chemicals. 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.

or industrial sludge) is responsible for the high organic carbon content of 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. Further guidance on obtaining reference soils for biological testing and procedures to be followed when a site-specific reference soil cannot be located is provided in EC (2012).

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 mites, and for geochemical and contaminant analyses. Other particulate wastes being considered for disposal to land might also be collected for toxicity and physicochemical evaluation. Environment Canada (2012) provides guidance on additional considerations unique to waste pile sampling.

A sampling plan is an important component of the study plan. The sampling plan is a written description of the detailed procedures to follow when collecting samples, handling and preparing samples on-site (if required), packaging, labelling, storing (if necessary), and transporting samples. Prior to extracting soil samples, it is important to obtain a thorough field description of the soil to be sampled. In addition, soils should be described at a detailed site-specific level. In Canada, soils are classified using the Canadian System of Soil Classification (CSSC). Soils collected for biological testing should be classified to the subgroup level according to the CSSC, following the guidance provided in EC (2012). Appendix E in EC (2012) provides detailed information on the CSSC and the basic components of soil taxonomic identification.

Procedures used for sample collection (i.e., point, bulk, 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. Shovels, scoops, or trowels are among the most commonly used tools in soil sampling when large volumes of soil are needed; however, care must be exercised to ensure that a representative and unbiased sample is collected (e.g., a constant depth or soil horizon must be removed). More precise sampling devices include soil corers, ring samplers, cutting frames, or soil cylinders, but they are less convenient for extracting large soil sample volumes. If soil samples are collected at a depth, an auger can be a more efficient and less labour-intensive tool for soil collection. Descriptions of the more commonly used soil collection devices and the procedures that should be followed for collecting soils are provided in EC (2012).

Most Canadian forest or non-agronomic, ecozone soils are highly stratified into soil horizon*s*. The structure and chemistry of soil horizons are often very different, and this can result in different bioavailability and toxicity of contaminants to soil organisms. The top layer (A horizon) is the most commonly sampled horizon for biological testing. This horizon contains the most organic matter and most of the biological activity in mineral soils. Depending on the study objectives, the forest litter (*L layer*), fulvic/humic (*FH horizon*) (e.g., at a forested site), or surficial organic layer (O horizon) of mineral soils (e.g., at a tundra site) might also be collected when present. Subsurface B horizons and less commonly C horizons might also be sampled. Soils from the boreal or taiga ecozones sampled for the assessment of effect(s) on mites, described in this test method document, must be collected as separate soil horizons, where possible. Collection of soil samples according to depth is recommended for soils without distinct soil horizons (e.g., where the surface soil horizons have been mixed or disturbed due to human activity). To sample soil by horizon, the soil profile must first be classified, as described earlier and in EC (2012). Care should be taken when sampling soil horizons that dilution of the soil contamination does not occur. This is particularly important in cases where the vertical contamination extends only partially through a soil horizon. In this situation, the horizon may be sampled only to a certain depth, or collected as two different samples at two sampling depths (EC, 2012).

Guidance on the collection of soil samples for toxicity testing is provided in detail in EC (2012). The first step is to establish the boundaries of the sample location. The surface of the location where each sample is to be collected should then be cleared of debris such as twigs, leaves, stones, thatch, and litter (unless the L layer is being collected as part of

the study design). If the location is an area of grass or other herbaceous 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 release soil particles adhering to the roots. The soil sample to be collected for toxicity evaluation and chemistry should be collected 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 are concerns about historical deposition of contaminants). Soils exhibiting distinct horizons (e.g., undisturbed forest soils) must be sequentially collected in separate horizons as a soil pit is excavated (EC, 2012).

The minimum volume or mass of soil required for testing depends upon the study objectives, site conditions, and the test to be conducted. For a given test, the amount of soil required varies and depends on the experimental design of the toxicity test (e.g., single concentration test versus multi-concentration test), as well as the physical characteristics of the soil (e.g., bulk density, moisture content, amount of debris in the soil), the nature of the chemical analyses to be performed, and the distribution of the contaminants in the soil (e.g., vertical distribution). The required volume of soil per sample should be calculated before commencing a sampling program. 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. Soil collection volume recommendations for specific biological tests are provided in EC (2012). To obtain the required sample volume, it is frequently necessary to combine subsamples retrieved using the sampling device. Guidance provided in EC (2012) for compositing subsamples in the field should be followed. The same collection procedure should be used at all field sites sampled. For samples collected as distinct soil horizons, each horizon must be placed and stored in separate containers unless the soil profile has been disturbed through attempts to remediate the site.

The preparation of soil samples might begin in the field before the samples are shipped to a testing laboratory. This might include hand-sorting (to remove debris and/or organisms), air-drying, sieving, and homogenization of soil samples. All of these procedures are described in detail in EC (2012).

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 non-toxic, inert material. The choice of container for transporting and storing samples depends on the sample volume, the potential end uses of the sample, and the type and nature of the soil contamination. The containers must be clean and sealable and should be practical for handling and able to support the weight of the sample (EC, 2012). Thick (e.g., 4 mil) plastic bags are routinely used for sample transport and storage. 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 maintain darkened conditions during sample transport (ASTM, 2004). Plastic containers or liners should not be used if there is concern about the plastic affecting the characteristics of the soil (e.g., compounds from plastic leaching into the soil). Containers recommended for the transport and storage of soils are listed in Appendix H of EC (2012).

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 must include at least a code or description that identifies sample type (e.g., point, bulk, composite), sample date and time, sample site, precise location of sampling, sample condition, sample identification number (including replicate number, where applicable), and sample volume. The label information should also include the name and signature of sampler(s). Persons collecting soil samples should also keep field 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 bulk or *point samples* in the field;
- any sample preparation (e.g., sieving, drying) carried out in the field;
- the number of replicate samples taken at each sampling location;
- the sampling schedule;
- the types and numbers of containers used for transporting samples;
- any field measurements (e.g., temperature, pH, soil moisture content, bulk density) of the soil at the collection site;
- soil horizon characterization;
- any in-situ field testing (e.g., litterbag, earthworm exposure, bait lamina) performed;
- procedures and conditions for cooling and transporting the samples;
- observations of environmental conditions at the time of sampling (e.g., raining);
- observations and any field sampling of soil fauna and flora at the collection site;
- sample storage duration and conditions prior to arrival at the laboratory; and
- information on sample transportation.

Additional recommendations for site observations and field measurements are provided in Table 10 of EC (2012).

Soil samples should be kept cool during transport and storage and should not freeze or become overheated. 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. 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 photo-activated or otherwise altered due to exposure to sunlight. All samples must be shipped with appropriate documentation, including chain-of-custody forms, as well as any specific regulatory documentation for transport of contaminated material (see EC [2012] for further guidance on sample transport).

The date the sample(s) is (are) received at the laboratory must be recorded. Sample temperature and moisture content upon receipt at the laboratory must also be measured and recorded. In addition, each sample of field-collected test soil or each separately collected soil horizon should be inspected and the following qualitative descriptions made and recorded: colour, texture, informal description of moisture content, presence of standing water, presence of indigenous invertebrates, fungi or plant material, and any strong odours (EC, 2012). Samples to be stored for future use must be held under conditions that maintain the characteristics and quality of the soil for its intended use (EC, 2012). If volatile contaminants are in the soil or are of particular concern, any air "headspace" in the storage container should be purged with nitrogen gas before being capped tightly. Samples should 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.[63] 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 recommended that samples of soil or similar particulate material be tested as soon as possible after collection. The effects of storage time and temperature on soil properties and toxicity depend on the contaminants and soil characteristics. 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

[63] Air-drying soil is another practical option for preserving natural soils and/or soils containing non-volatile or light sensitive contaminants, as it allows a fast and more precise rehydration, and allows for the storage of samples at room temperature. Guidance on air-drying soils is provided in Section 3.10.3.1 of EC (2012).

the soil contaminants are aged and/or weathered, and therefore considered stable. Further considerations for the storage of contaminated soil are provided in EC (2012), and the guidance therein should be followed.

In the laboratory, each sample of field-collected soil or distinct soil horizon should be thoroughly mixed (Section 5.3), and representative subsamples collected for physicochemical characterization. Each sample or soil horizon to be tested (including all associated 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 (%)[64]
- organic matter content (%)[64]
- pH
- electrical conductivity
- moisture content (%)
- water-holding capacity (WHC)
- cation exchange capacity (CEC)

Additionally, the following analyses should be performed:

- major cations and anions (Na^+, K^+, Mg^{2+}, Ca^{2+}, Al^{3+}, S^{2-}, S^{2-}, Cl^-)
- nitrogen as total N, nitrate (NO_3^-), nitrite (NO_2^-) and ammonium (NH_4^+)
- phosphorus as total and/or bioavailable
- potassium as total and/or bioavailable
- C:N ratio

Other analyses could include:

- bulk density
- total inorganic carbon
- total volatile solids
- biochemical oxygen demand
- chemical oxygen demand
- redox potential
- soluble salts
- sodium adsorption ratio
- contaminants of concern
- characteristics of the contamination (e.g., odour, staining, debris, presence of fuel or solvent)

Unless indicated otherwise, identical chemical, physical, and toxicological analyses should be performed with subsamples representative of each replicate sample of field-collected soil or soil horizon (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 undesirable coarse debris (e.g., plant material, wood chips, glass, plastic, large gravel) or large macroinvertebrates, these may be removed by gently passing the soil through a coarse sieve (e.g., mesh size of 4 to 10 mm; EC, 2012). Dry sieving might also be desirable to ensure that the sample structure (i.e., aggregation, organic matter, or clay distribution) is amenable for testing. Soils should not be sieved in the laboratory if they were sieved in the field, or if they have the crumbly texture that is optimal for testing (i.e., 1- to 3–mm clumps). Soil samples consisting of moist clayey subsurface soils are very cohesive and often cannot be directly sieved or homogenized. These soils should first be broken up manually and then dried prior to sieving and homogenization, as described in EC (2012). In general, grinding of soil samples should be avoided when possible, but might be necessary with some soils (i.e., clayey soils) or if greater homogeneity of a sample is desired than can be achieved by sieving. As with soil sampling and storage procedures, any soil preparation procedures (i.e., *pretreatment*) should be documented and must be reported.

[64] Organic matter content can be used to calculate total organic carbon (TOC) by multiplying the organic matter (OM) content of a soil by a soil constant (AESA, 2001). However, the relationship between TOC and OM is slightly different among soils and the total organic carbon content should also be determined by laboratory analysis.

Reconstitution of soil sample constituents might be required prior to testing if the soil contained standing water that was decanted during preparation, or if portions of the sample were removed during preparation (e.g., thatch, plant root, or other organic material) but need testing along with the soil (EC, 2012). Soil horizons collected as separate components of a soil sample must be tested independently as separate soil samples. If the contaminants of concern have only been confirmed in one soil horizon (e.g., upper organic horizon) based on previous analyses and/or toxicity testing, then, depending on the study objectives, a decision must be made as to whether to conduct toxicity testing on this horizon alone or in the additional soil horizons collected from the sampling location.

Unless research or special study objectives dictate otherwise, each sample or horizon of field-collected unconsolidated test material should be homogenized in the laboratory before use (USEPA, 1989).[65] Any moisture that separates from a sample during its transport and/or storage must be remixed into it, if possible. Mixing can affect the concentration and bioavailability of contaminants in the soil, and sample homogenization might not be desirable for all purposes. To prepare a homogeneous sample, transfer the precalculated amounts of test and/or reference soil to a clean, rigid mixing container (e.g., a large stainless steel or plastic bowl) or, for larger volumes of soil, to clean plastic sheets spread out on the floor. The sample should be mixed manually (using a gloved hand or a non-toxic 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. A number of methods used to homogenize soil samples (e.g., folding, mixing, coning) are described in detail in EC (2012). While mixing, care should be taken to ensure that the impact of mixing on soil structure is minimal and that the structure is not destroyed entirely. As soon as the texture and colour of the sample appear to be homogeneous, mixing should be discontinued.

For each sample or soil horizon 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.

As indicated in Section 3.6, one or more samples or horizons of field-collected test soil may 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. Guidance on concentration series that might prove suitable is found in Section 6.2, along with that for preparing test mixtures that might apply equally when performing a multi-concentration test with one or more samples of field-collected soil. Refer to Section 4.1 for additional guidance when selecting test concentrations. In each instance, the test must include a treatment consisting solely of negative control soil (see Section 3.3).

As indicated in Section 4.1 for soils collected as distinct horizons, each horizon must be tested separately in independent definitive tests. For a multi-concentration test, the test soil horizon should be mixed with the same horizon of negative control or reference soil at the various test concentrations (0%, 6.25%, 12.5%, 25%, etc.). In some cases, it might not be possible to collect the same horizons of negative control soil and test soil. For example, preliminary remedial action might have already been taken at the test site, resulting in disturbed or mixed natural soil horizons. In these scenarios, the test soil may be tested as a mixed soil where test concentrations are prepared by mixing suitable weights of test soil into the available horizon(s) of negative control soils at the appropriate test concentrations. The study objectives must take into account the soil profile of the reference soil and the location and/or mobility of the contaminants in the test soil. The goal is to match equivalent horizons in reference and contaminated soil, if possible.

---

[65] One of the reasons for routinely homogenizing samples is to mix into the soil any pore water that 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.

Soil structure is an important factor that influences the reproduction of mites, and moisture content plays an important role in the determination of soil structure. A qualitative procedure, informally known as a "squeeze test," can be 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. 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 (OECD, 2016). As the test proceeds, test vessels should be weighed to determine water loss (see Section 4.6).

The moisture content of a given sample of field-collected test soil should be standardized during its preparation 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 prior to sample preparation and test initiation. 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 (Sections 4.1 and 4.6). Thereafter, the WHC of each sample must be determined using a recognized standard procedure (see following three paragraphs). A subsample of each soil sample should then be hydrated (or, if and as necessary, dehydrated) to a homogeneous, crumbly consistency with clumps approximately 1 to 3 mm in diameter. The moisture content, WHC, and optimal percentage of the WHC of each soil horizon must be determined separately. Soil horizons with higher organic matter content can be expected to have a higher WHC than mineral horizons, so will require greater amounts of water to hydrate to a moist, crumbly texture. 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.[66] 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., deionized or distilled water[67]) should be added to each sample with a moisture content that is less than the predetermined optimal percentage of its WHC, until this moisture content is achieved[68] (AquaTerra Environmental Ltd., 1998). 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 vapour-barrier plastic) or a clean, non-reactive (e.g., stainless steel or plastic) tray, and allowed to air-dry by evaporation at ambient (~20°C) room temperature;[69] rehydration to the predetermined

---

[66] For soils with high peat content (i.e., extremely high water-holding capacity), the method for determining the percent WHC described herein might be inaccurate and the results misleading. In such cases, the optimal moisture content may be estimated by eye (i.e., sample hydrated to a homogeneous, crumbly consistency with clumps approximately 1 to 3 mm in diameter) and the moisture content determined thereafter and reported as such (i.e., as moisture content instead of percent WHC).

[67] The use of purified water (i.e., deionized or reverse osmosis) to hydrate soils avoids the introduction of cations, anions, or trace metals into the soil (EC, 2012).

[68] 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, 2004). 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, 2004). Accordingly, the use of this alternate approach is not recommended here.

[69] If there is concern about volatilization of potential toxicants and/or changes in the nature of the toxicant of concern due to the drying process, alternative methods of drying the soil and/or the effects of drying the soil on the toxicity of the soil may be investigated.

optimal percentage of its WHC might be necessary. Upon completion of 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 (and the percent WHC that is optimal for biological testing) of a particular soil is generally unique to each soil type and/or horizon, 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 organic matter content) 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).[70] This method is outlined here.

For this method, ~130 g (wet wt)[71] of sample is placed in an aluminum pan or large petri dish (15 × 1 cm), and dried at 105°C until a constant weight is achieved (this usually takes a minimum of 24 h). The soil is then cooled for a minimum of 20 min. in a desiccator. Thereafter, 100 g of the oven-dried soil is placed into a 250-mL glass beaker with 100 mL of distilled or deionized water. The resulting slurry is mixed thoroughly with a glass stir rod. A folded filter paper (e.g., 185-mm diameter Fisherbrand™ P8 coarse porosity, qualitative creped filter paper; catalogue number 09-790-12G) is placed into a glass funnel (with a top inside diameter of 100 mm and a stem length of 95 mm). The folded filter paper should be level with the top of the glass funnel. Using a pipette, up to 9 mL of distilled or deionized 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 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.[72] 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 Equation 1).

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

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

---

[70] Some participants at the soil toxicity testing workshop sponsored by Environment Canada 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, 2004b). 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 Environment Canada's soil toxicity test methods when adjusting (if and as necessary) the moisture content of soil samples (Becker-van Slooten, *et al*., 2004).

[71] A larger amount of soil (i.e., for highly organic soils) might be necessary to obtain 100 g of soil (dry wt).

[72] In very organic soils, where humic compounds' hydrophobicity delays water uptake, WHC can be underestimated unless the length of the soil saturation period is extended.

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 hydration (i.e., the optimal percentage of the WHC) can be calculated as follows:[73]

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

---

[73] 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 a reproduction test with mites involving five replicates per treatment.

Assumptions

Soil #1: Negative Control (nc) Soil

W_{nc} = 2.4 g
D_{nc} = 1.9 g
WHC_{nc} = 80.3%
P_{WHCnc} = 60.0%
MC_{nc} = 26.3%
P_{Wnc} = 21.9%
M_{Dnc} = 93.7 g dry wt
V_{Wnc} = 20.5 mL
M_{Wnc} = 118.4 g wet wt

Soil #2: Contaminated (c) Soil

W_c = 7.1 g
D_c = 5.6 g
WHC_c = 77.1%
P_{WHCc} = 40.0%
MC_c = 26.8%
P_{Wc} = 4.04%
M_{Dc} = 31.3 g dry wt
V_{Wc} = 1.3 mL
M_{Wc} = 39.7 g wet wt

$MC = [(W - D) / D] \times 100$ [Equation 1]
$P_W = [WHC \times (P_{WHC} / 100)] - MC$ [Equation 2]
$V_W = (P_W \times M) / 100$ [Equation 3]
$M_W = (M_D \times W) / D$

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 a mite test using this example, it is assumed that a total mass of 125 g dry weight (wt) of soil is sufficient to satisfy the requirement for each treatment (i.e., 20 g dry wt per replicate × 5 replicates + 25 g dry wt extra soil for pH and electrical conductivity). To simplify the calculations, this example assumes that 20 g (dry wt) of either type of soil is sufficient to provide the 20 mL of soil volume to be added to each test vessel.

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:

= 125.0 g dry wt × (25/100)
= 31.3 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:

= 125.0 g dry wt × (75/100)
[or 125.0 g dry wt – 31.3 g dry wt]
= 93.7 g dry wt of negative control soil

Therefore, the final total mass of soil required, based on wet weight, is 138.9 g [118.4 g wet wt at the soil's initial moisture content (i.e., M_{Wnc}) + 20.5 mL of water] for the negative control soil, and 41.0 g [39.7 g wet wt at the soil's initial moisture content (i.e., M_{Wc}) + 1.3 mL of water] for the contaminated soil.

The final moisture content for each soil would be 48.2% {[(138.9 – 93.7) / 93.7] × 100} for the negative control soil, and 31.0% {[(41.0 – 31.3) / 31.3] × 100} for the contaminated soil.

The final moisture content of the negative control soil (i.e., 48.2%) represents 60% of that soil's water-holding capacity (48.2 ÷ 80.3 = 0.60). The final moisture content of the contaminated soil (i.e., 31.0%) represents 40% of that soil's water-holding capacity (31.0 ÷ 77.1 = 0.40).

where:

P_W = percentage of water to add to the soil (%)
WHC = water-holding capacity (%)
P_{WHC} = percentage of WHC desired (%)
MC = 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 (see footnote 73):

$$V_W = (P_W \times M)/100 \qquad \text{(Equation 3)}$$

where:

V_W = volume of water to add to the soil (mL)
P_W = percentage of water to add to the soil (%)
M = total mass of soil required for test (expressed as dry weight)[74]

Environment Canada (2012) describes various procedures that can be used to manipulate soil samples to render them testable to meet study objectives or DQOs when the conditions do not occur within the sample as collected. Detailed procedures for soil manipulations are described and include: washing, aging/weathering, adjusting soil pH, conditioning, adjusting soil *fertility,* and reducing indigenous soil microorganisms (EC, 2012). In general, samples of field-collected soil must not be adjusted or manipulated, except for research-oriented toxicity tests intended to determine the influence of a particular soil manipulation on sample toxicity. Soil horizons with high organic levels (e.g., *LFH horizons*), however, might require at least one or more freeze/thaw cycles in order to remove indigenous invertebrates before testing (see Section 5.6.6 of EC, 2012).[75] Studies intending to investigate the effect of a soil manipulation (e.g., pH adjustment) on sample toxicity should involve two side-by-side tests whereby one or more sets of treatments are adjusted, and one or more duplicate sets of treatments are not. Detailed, proper documentation of any soil manipulation procedures carried out must be made and reported.

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 vessels (see Section 4.1), and into the labelled containers required for the storage of subsample for subsequent physicochemical analyses. Any remaining portions of the homogenized sample that might be required for additional toxicity tests using mites or other test organisms (e.g., according to EC, 2004a, 2005a, 2013a, 2014a) should also be transferred to labelled containers at this time. Subsamples to be stored for future toxicity testing

---

[74] For tests with samples of field-collected soil, the amount of soil added to each 30-mL glass shell vial described in Section 3.2.2 is based on the wet weight of soil that is equivalent to a volume of ~ 20 mL (producing a soil depth of ~ 4cm). When the optimal percentage of the soil's WHC is determined, the equivalent wet weight (of ~20 mL within the test vessel) should be determined, and the sample analyzed for dry mass. Then, the total mass required per replicate and test concentration can be determined, based on dry mass equivalent. The "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 vessel on a dry-weight basis, a simple calculation is carried out. For example, assume that (for a given sample) 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 4.2 g and 2.8 g, respectively. The dry weight equivalent to a 30-g wet weight of this sample of soil can be calculated as follows:

(30 g × 2.8 g) ÷ 4.2 g = 20.0 g

Therefore, for the example provided here, the mass of this sample of soil required for each replicate (expressed as dry wt) is 20 g. The total mass ("M") can then be calculated simply by multiplying the dry mass required for each replicate (in this instance, 20 g dry wt) by the number of replicates to be used in the test (i.e., for this example, five replicates).

[75] To initiate a freeze/thaw cycle, the soil sample is placed in the freezer (≤ -20°C) for a minimum of three days. The soil is then removed from the freezer and allowed to thaw at ≥ 20°C for seven days. The cycle may then be repeated at least once more before testing is initiated (C. Fraser, Environment Canada, personal communication, 2013).

should be held in sealed containers with minimal air space, in darkness at 4 ± 2°C (Section 5.2) until tested. These storage conditions must be applied for subsamples collected for physicochemical analysis. Just before it is analyzed or used in the toxicity test, each subsample must be brought to room temperature and thoroughly remixed to ensure that it is homogeneous.

5.4 Special Considerations for the Collection, Handling, and Preparation of Soil from Canada's Ecozones

Specific guidance on sampling, handling, transporting, storing, and preparing soil from various Canadian ecozones is provided in EC (2012).

Previously published Environment Canada soil toxicity test methods (EC, 2004a, 2005a) were developed for the assessment of soils with neutral to near-neutral soil pH and organic matter content ranging from approximately 3% to 12%. These soils are generally characteristic of the Ah horizons of agricultural soils in Canada and soils from deciduous mixed forest eco-regions in the southeastern part of the country (i.e., prairies and mixed-wood plains ecozones). There are many other soil types in Canada with widespread distributions that have properties falling outside the ranges considered typical by EC's previously published standard methods, and therefore require special procedures for sampling, handling, transport, storage, and preparation. These soils include: boreal forest soils, taiga soils, stony/shallow soils, organic soils, cryosolic soils, and wetland soils, and are relevant for use with the test methodologies described in this test method document. Given that these soils cover most of Canada's land mass and that anthropogenic activities in these regions (e.g., mining, forestry, oil and gas production) have created or have the potential to create contaminated lands, specific guidance on sampling, handling, transporting, storing, and preparing soils from these various ecozones is provided in EC (2012). Guidance is also provided on the variability of the soils within each of the described ecosystems and special considerations for selecting the appropriate test species when testing soils from these various ecosystems (EC, 2012).

5.5 Test Observations and Measurements

A qualitative description of each field-collected test material should be made at the time that the test is being 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.

Section 4.6 provides 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 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 vessels may be set up at the beginning of the test (Section 4.1) 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 vessels, depending on the study's objectives. Measurements of chemical concentrations in the soil within these vessels may be made by removing aliquots of the soil for the appropriate analyses (see Section 5.2).

5.6 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, 2004a, 2005a, 2013a, 2014a). 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 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 the negative control soil must be used to judge the validity and acceptability of the test (see Section 4.4).

The biological endpoint for this method is reproductive success (a quantitative measurement) at the end of the test. Because of the different nature of the measurements involved, different statistical approaches are needed, and these approaches are further refined to reflect the objectives and design of the experiment. This section provides statistical guidance on data from single-concentration tests (i.e., soil samples from multiple sampling locations tested at full strength only). The simplest testing scenario involves the comparison of one test sampling location with one reference sampling location, whereas more complex designs might include a comparison of several sampling locations with a reference sampling location, or with each other. Only summary guidance is provided here for analyzing the reproduction endpoint as more extensive statistical guidance is available elsewhere (EC, 2005b). Standard statistical procedures are generally all that is needed for analyzing the results. Section 3 in EC (2005b) should be consulted for guidance when comparing the findings for single-concentration tests from multiple locations using parametric or non-parametric tests. As always, the advice of a statistician familiar with *toxicology* should be sought for the design and analysis of tests.

Guidance in Section 6 (including that in Section 6.2 for performing range-finding tests, and that in Sections 6.4 and 4.8 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 negative control soil or clean reference soil. Section 9 in EC (2005b) should be consulted when comparing such point estimates of toxicity for multiple samples of field-collected soil.

5.6.1 Variations in Design and Analysis

Environment Canada (EC, 2005b) provides detailed statistical guidance on the analysis of quantitative data in various test designs that examine multiple sampling locations. Choice of a specific statistical test depends on several considerations, including but not limited to:

- the type of comparison that is sought (e.g., complete series of pairwise comparisons between all sampling locations, or compare the response from each sampling location only with that of the reference site);
- if a chemical and/or biological response gradient is expected;[76] and
- the level and type (laboratory or field) of replication.

This guidance (EC, 2005b)[77] can be readily applied to measurements of mite reproduction (i.e., number of surviving progeny at the end of the test) in a multiple sampling location scenario. If test results at a single test sampling location are to be compared with test results at a reference sampling location, a *t*-test[78] is normally the appropriate statistical test (Section 3.2 in EC, 2005b). In situations where more than one test sampling location (treatment) is under study, and the investigator wishes to compare multiple sampling locations with the reference, or compare sampling locations with each other, a variety of ANOVA and multiple comparison tests

[76] In this case, the expected gradient is determined during the experimental design phase (a priori) not after the data has been collected. Section 3.3 in EC 2005b provides guidance on cases where a gradient effect is expected. If necessary, a statistician should be consulted for further guidance on analyses of data where a gradient is expected.

[77] Sections 3.2 and 3.3 in EC 2005b provide guidance on the analysis of quantitative measurements for a single location and quantitative measurements for multi-locations, respectively, and should be consulted for the analysis of reproduction data. Section 7.5 in EC 2005b provides additional guidance on multiple-comparison tests for hypothesis testing, and should be consulted for additional detail; however, the calculation of NOEC/LOEC is not recommended herein.

[78] The *t*-test assumes equal variance between groups; however, modification of the t-test that can accommodate unequal variance is also available (EC, 2005b).

(and non-parametric equivalents) exist (Section 3.3 in EC, 2005b). Choice of a specific test depends on the three conditions described above, in addition to assumptions of normality and homoscedasticity being met.

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 on designing more extensive studies. A preliminary evaluation might conceivably be conducted with samples from many locations, but without either field replicates or laboratory (within-sample) replicates. The objective might be to identify a reduced number of sampling locations deserving of more detailed and further study. In this case, opportunities for statistical analysis would be limited (EC, 2005b).

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. In these multi-location surveys, the type of replication would influence the interpretation of results (i.e., field replicates or laboratory replicates, or both). If both replicate samples (i.e., field replicates) and replicate vessels (i.e., laboratory replicates) have been tested, a statistician should be consulted for analysis options. If only laboratory replicates and no field replicates were tested, it is difficult to make statistically robust conclusions regarding differences between sampling sites (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, a number of tests are recommended, depending on whether the samples show a gradient and depending on whether there is an even or uneven number of replicates (see Section 3 in EC, 2005b).

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. Sections 3.1, 3.3, and 7.5 in EC 2005b provide further details, alternate tests, and non-parametric options, and the guidance therein should be followed.

5.6.2 Power Analysis

An important factor to consider in the analysis of toxicity tests with soils 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 for tolerating false positive results (Type I error), and usually set it at $p = 0.05$ or 0.01. Commonly, scientists following a specified test design will never consider the relationship between power, variability, and effect size, leaving the Type II error (β) completely unspecified. There are several factors that influence statistical power, including but not limited to:

- variability of replicate samples representing the same treatment;
- α (i.e., the probability of making a Type I error);
- effect size, (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, and in some cases, the allocation of those replicates).[79]

[79] If the experimental design requires the comparison of test samples with the reference sample only (e.g., using Dunnett's test or Williams' test), optimal power for the final reproduction endpoint is achieved by allocating a higher number of replicates in the reference treatment (Dunnett, 1955; Williams, 1972; OECD, 2006). As a

Environment Canada's guidance document on statistical methods for environmental toxicity tests (EC, 2005b) provides further information and guidance on Types I and II errors.

In research-based science, power analysis is most useful as part of a preliminary test design (Hoenig and Heisey, 2001; Lenth, 2007; Newman, 2008). Here, a preliminary experiment is run to determine the approximate standard deviation (variability), and to troubleshoot the execution of the experiment in general. Other factors in power analysis, such as effect size and number of replicates, can then be considered along with the standard deviation so that the final test design is optimized (e.g., number of replicates needed to detect a certain effect size is determined).

In the development of standardized test methods, the purpose of employing power analysis remains the optimization of test design or at least estimating the power of the current test design.[80] However, instead of a single estimate for variability and effect size, there would typically be a much richer data set to consider. For example, test method experts could collect a number of estimates of variability across different laboratories and different contaminant scenarios (Thursby *et al*., 1997; Van der Hoeven, 1998; Denton *et al*., 2011). Standardized tests are often used in monitoring or regulatory programs, which might specify the expected effect size (e.g., 25%) to be detected (AE, 2007).

Data from interlaboratory validation tests (ECCC, 2019) were used to estimate power for detecting a reduction in the number of surviving progeny (L. Van der Vliet, Environment and Climate Change Canada, personal communication, 2019). Variability estimates were collected from seven laboratories, from four different soils types, and from three rounds of testing. The most extensive data were available for artificial soil, and so the power analysis was focused on this soil type. Averaging across all labs, the variability among control replicates decreased from Round 1 (average CV = 38%) to Round 3 (average CV = 26%), suggesting that as labs gained experience, variability decreased. The CV from Round 3 was used in the power analysis, as it is assumed that laboratory personnel have been trained and have performed initial tests (see Section 3.2.1). Variability estimates were only available for replicate test vessels (laboratory replicates); variability among replicate samples (field replicates) is expected to be higher, and this expected increase would impact the power analysis. Effect sizes of 30%, 40%, and 50% reduction in number of surviving progeny were used. A one-sided, equal variance t-test with $\alpha = 0.05$ was assumed.

The power analysis showed that, given the conditions listed, for five replicates, an effect size of 40% or greater can be reliably detected (power $\geq$ 80%). This supports the requirement for a minimum of five replicate test vessels. If project or program requirements have specified a smaller effect size, more replicates are recommended. For example, if a 30% effect size is targeted, then eight replicates are recommended. That is, with eight replicates, a 30% effect size can be reliably detected (power $\geq$ 80%).

---

general rule, the number of reference replicates (n_o) can be related to the number of test sampling locations (k) and the number of test replicates (n) using: $n_o = n\sqrt{k}$ for Dunnett's test (OECD, 2006). A modified version is recommended if Williams' test is used, where $\sqrt{k}$ is replaced with a range between $1.1\sqrt{k}$ and $1.4\sqrt{k}$ (Williams, 1972). With the current test method, each sampling location should have a minimum of five replicates. If the investigator was interested in increasing the number of replicates beyond the minimum, extra replicates should be allocated to the reference samples to maximize power and minimize Type II error. As an example using Dunnett's formula, consider an experiment with one reference sampling location and four test sampling locations, and five replicates for each location. To maximize power, the optimal number of replicate samples at the reference sampling location would be $n_o = n\sqrt{k} = 5\times\sqrt{4} = 10$ replicates.

[80] In 2010, the USEPA introduced a data analysis approach termed the test of significant toxicity approach (TST; USEPA, 2010). The TST is a hypothesis testing approach based on bioequivalence, which is extensively used in pharmaceutical development and evaluation. It is included in the discussion here because power analysis and the TST share some similar goals (e.g., a priori statement of Type I and Type II error) and because of the similar context (application of standardized testing).

Section 6

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 the biological test method described in Section 4. Guidance in EC (1995) on 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 mites 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 3.3) with these substances. The spiking might be done with one or more chemicals or chemical products. Other options for toxicity tests with mites, performed using the procedures described herein, include the spiking of chemical(s) or chemical product(s) in reference soil (Section 3.5) or test soil (Section 3.6). Soil horizons collected separately must be treated as separate soil samples, as described in previous sections (4.1 and 5.3), and must be characterized and prepared (i.e., hydrated and spiked) separately prior to being tested (Section 6.2). 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 can then be used to determine statistical endpoints based on threshold concentrations causing specific *sublethal effects* (see Section 4.8.1).

In Section 6.2, procedures are described for preparing test mixtures of chemical-spiked soil. Section 6.3 describes making observations and measurements during and at the end of the toxicity test. Section 6.4 (and Section 4.8) 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 Sections 4.8 and 5, and especially 5.6). Multi-concentration tests with positive control soil (Section 3.4) or one or more concentrations of reference toxicant spiked in negative control soil (Section 4.9) 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.[81] Information should also be obtained for individual chemicals or chemical products (e.g., pesticides or

[81] Some studies might require the spiking (mixing) of one or more concentrations of chemical(s), chemical product(s), or test soils (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.

other commercial formulations), on their concentration of major "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 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 Safety Data Sheets 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 into negative control soil should be prepared, transferred to test vessels, and held overnight before adding the test organisms the next day (i.e., Day 0) (see Section 4.1). For some chemicals or chemical products (e.g., those that are very volatile, degrade easily, or might be metabolized), the addition of test organisms may be carried out immediately after preparation of the test soil. The dates of test soil preparation and test organism addition must be recorded and reported. 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 monitoring and/or physicochemical analyses (Sections 4.6 and 6.3) or the performance of other soil toxicity tests using mites or other soil organisms (e.g., those performed according to EC, 2004a, 2005a, 2013a, or 2014a).

The use of artificial soil (Section 3.3.2) to prepare each test mixture 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 USEPA, 1989; Wiles and Krogh, 1998; ISO, 1999; OECD, 2009). If used, the formulation for artificial soil provided in Section 3.3.2 should be followed. The quantity of artificial soil required for the test(s) should be prepared, hydrated to ~20% moisture content (which is ~28% of the soil's WHC), adjusted if and as necessary to a pH within the range of 6.0 to 7.5,[82] aged for a minimum three-day period, and stored until required (see Section 3.3.2). 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 3.3.2), for each treatment (concentration), or that which produces the optimal soil texture for testing (i.e., a homogeneous crumbly consistency

[82] 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) after the chemical(s) or chemical product(s) has (have) been added. Studies wishing to determine 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 conduct two side-by-side tests whereby 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.

with clumps ~1 to 3 mm in diameter; see Section 5.3).[83] The final moisture content of each mixture (treatment) included in a test should be as similar as possible.

Investigators may choose to use natural control soil (Section 3.3.1) rather than artificial control soil (Section 3.3.2) 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 (by hydrating or dehydrating the sample, as the case may be) using the guidance in Section 5.3. For natural soils, the weight of soil in each test vessel might also differ due to differences in bulk density of the various soils that might be used.

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

---

[83] The following example provides calculations that show the volume of both water (deionized 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 20 g (dry wt) of artificial soil (AS) is sufficient to provide the 20-mL aliquot of soil to be added to each test vessel (i.e., 30-mL glass shell vial) when performing a mite toxicity test involving five replicate test vessels per treatment.

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) = 3.2 g
Dry mass of AS = 2.7 g
Moisture content (MC) of AS
= [(3.2 – 2.7)/2.7] × 100
= 18.5% (initial moisture content)
Water-holding capacity (WHC) of AS = 72.1%
Percentage of WHC desired (P_{WHC}) = 70.0%
Dry mass of AS required for test (M_D)
= [20.0 g per rep × 5 reps] + 25.0 g extra
= 125.0 g dry wt
Wet mass of AS required for test (M_W)
= (125.0 × 3.2)/2.7
= 148.1 g wet wt

Calculations to prepare a treatment consisting of 200 mg boric acid per kg artificial soil (dry wt):
The stock solution consists of 0.4 g of H_3BO_3 in 100 mL of deionized water.

The amount of boric acid required on a dry-mass basis is:
H_3BO_3= (0.2 g H_3BO_3/1000 g soil dry wt)×125.0 g dry wt
= 0.025 g H_3BO_3

The amount of stock solution required, on a volume basis, is:
H_3BO_3 = 0.025 g H_3BO_3 / (0.4 g H_3BO_3/100 mL of water)
= 6.2 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 × (P_{WHC}/100)] – MC
= [72.1 × (70.0/100)] – 18.5
= 32.0%

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 × M_D)/100
= (32.0 × 125.0 g dry wt)/100
= 40.0 mL of water required

However, as part of this required volume, 6.2 mL of the stock solution is to be added for dosing; therefore, an additional volume of water of only 33.8 mL will be required (40.0 mL of water – 6.2 mL of stock solution).

Accordingly, the final total mass of soil required, based on wet weight, would be 188.1 g [148.1 g wet wt at the soil's initial moisture content (i.e., M_W) + 33.8 mL of water + 6.2 mL of stock solution], and the final moisture content of the soil, based on dry weight, would be 50.5% {[(188.1 – 125.0)/125.0] × 100}.

The final moisture content of this test treatment (i.e., 50.5% moisture) represents 70% of the test soil's water-holding capacity (50.5 ÷ 72.1 = 0.70).

product(s) and then mixing one or more measured volumes into *hydration water,* which is then added to artificial or natural negative control soil (Section 3.3).[84] The preferred solvent for preparing stock solutions is test water (i.e., deionized 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 (OECD, 2009). Surfactants should not be used.

If an organic solvent is used, the test must be conducted using a series of replicate test vessels 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 vessels containing only *solvent control soil* (ISO, 1999; OECD, 2009). For this purpose, a batch of solvent control soil must be prepared that contains the concentration of solubilizing agent that is present in the highest concentration of the test chemical(s) or chemical product(s) in the 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 reproduction of mites during the test. If this information is unknown, a preliminary solvent-only test, using various concentrations of solvent in the 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 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 consisting of a measured amount of a particular chemical-in-solvent mixture spiked in artificial soil (see Section 3.3.2). The solvent can then be 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) can be mixed thoroughly with the remaining quantity of pre-moistened sand and other ingredients required to make up artificial soil (Section 3.3.2). An amount of test water necessary to achieve a final moisture content of approximately 70% of the maximum water-holding capacity for this artificial soil can then be added and mixed with the soil/sand/peat mixture. The chemical-spiked soil can then be added to the test vessel.

For tests involving the spiking of natural soil, in which the chemical is insoluble in water, the following procedure may be used (R. Kuperman, US Army Edgewood Chemical Biological Center, personal communication, 2004). The chemical is dissolved in a solvent (e.g., acetone) and pipetted onto a 2.5-cm thick layer of soil to establish each chemical concentration in soil, ensuring that the volume of solution added at any one time does not exceed 15% (v:m) of the dry mass soil. The same total chemical:solvent solution volume at different concentrations is added to every treatment, equalling the volume required to dissolve the chemical at the highest concentration tested. The solvent is allowed to volatilize (usually requires a minimum of 18 h) in a dark chemical fume hood to prevent photolysis. Each amended soil sample is mixed until homogeneous (e.g., transferred into a fluorocarbon-coated high-density polyethylene container and mixed for 18 h on a three-dimensional rotary mixer). Other procedures for dissipation of solvent may be used depending on the nature of the test chemical and/or solvent.

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

[84] Adding the stock solution to the hydration water and then to the soil assists with homogenization and decreases the risk of having the contaminant bind to a very small area of soil.

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 that consists of 2.5 g of 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 deionized water necessary to achieve a final moisture content of ~70% of the maximum water-holding capacity can then be added and mixed in. The resulting mixture of chemical-spiked soil can then be added to the test vessels.

If the test chemical to be spiked in natural soil is insoluble in both water and any suitable (non-toxic) organic solvent, the test chemical can be added through dry-mixing. The following procedure may be used (Ritchie *et al*., 2017; EC, 2014b). A mixture of the natural soil and the quantity of test chemical necessary to achieve the desired concentration in the soil is prepared. This mixture is initially combined using an electric mixer, and then mixed over the course of several hours (e.g., 16 h), using a mechanical stirrer or mixer (e.g., rotary mixer) until homogeneous. The *spiked soil* can be mixed with test water (e.g., up to 50% of its optimal moisture content), prior to chemical spiking. Each concentration can be dry-mixed independently. Alternatively, a mixture of the test chemical and a portion of clean soil can be prepared at the highest test concentration, in a sufficient volume to meet the requirements of a test through dilution of the spiked soil with clean soil, following the initial spiking and mixing event. The efficacy of the dry-mixing procedures should be evaluated through chemical analysis of aliquots of soil.

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 (ISO, 1999; OECD, 2009). The assessment endpoints (e.g., ICps) are similarly expressed on a dry-weight basis (Sections 4.8 and 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, which could be for several minutes or as much as 24 h. During mixing, the 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 (e.g., "limit test").

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 (i.e., ICp; see Sections 4.8 and 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 seven test concentrations plus the appropriate control treatment(s) must be prepared for each multi-concentration test, and more (i.e., ≥ 10 plus controls) may be used (see Sections 4.1 and 4.8). If a range-finding test is conducted prior to definitive testing, fewer concentrations may be used in the definitive test, since more information on the effect concentration/dilution range will be available. In this case, a minimum of five test concentrations must be used. When selecting the test concentrations, an appropriate geometric dilution series should be used and may be selected such that each successive concentration of chemical(s) or chemical product(s) in soil is at least 50% of the previous one (e.g., 40, 20, 10, 5, 2.5, 1.25, 0.63 mg/kg). Test concentrations may also be selected from other appropriate

logarithmic dilution series (see Appendix F) or may be derived based on the findings of preliminary "range-finding" toxicity tests. The reader is referred to Section 4.1 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 Section 4.1) 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 a similar test substance) variance among test vessels within a treatment, might also be reduced for nonregulatory screening bioassays or research studies.

Depending on the test objectives, 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 consisting 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 3.3). Similarly, the degree to which the total organic carbon content (%) or organic matter content (%) of soil or soil horizons 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 mites 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 swath width, the nozzle size, the pressure, and the speed of coverage of the sprayer (G. Stephenson, AquaTerra Environmental Ltd., personal communication, 2001). The OECD (2009) provides some guidelines for applying test substances to the soil surface in preparation for reproduction tests with soil-dwelling organisms.

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 on 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.

Section 4.6 provides guidance on 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 the soil toxicity test described herein using one or more samples of chemical-spiked soil. For soils collected as soil horizons, these measurements must be made in each soil horizon tested.

Depending on the test objectives and experimental design, additional test vessels might be set up on Day -1 of the test (see Section 4.1) 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) or at the end of the test. These monitoring vessels would be set up on Day 0 if the test is initiated (i.e., organisms added to the test vessels) immediately after the preparation of the test soil due to concern over the volatilization, degradation, or metabolism of contaminants or chemicals in test soils (see Section 6.1). Test organisms might or might not be added to these extra test vessels, depending on study objectives. Measurements of chemical concentrations in the soil within these test vessels 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, depending on the nature of the toxicant and the objectives of the test.

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 Section 4.6. If analytical capabilities permit, it is recommended that the stock solution(s) be analyzed together with one or more subsamples 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,[85] in which instance, the endpoint values calculated (Sections 4.8 and 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 Section 4.8). Guidance on calculating an ICp (based on data showing reproductive inhibition) is given in Section 4.8.1. Section 5.6 provides guidance on calculating and comparing endpoints for single-concentration tests using samples of field-collected soil. This guidance 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, 2005b).

For any test that includes solvent control soil (see Section 6.2), the test results for mites held in that soil and in negative control soil must be examined to determine whether they independently meet the test validity criteria (see Section 4.4). If either of these controls fails to meet the test validity criteria, the test results must be considered invalid. If both controls meet the test validity criteria, the results from the solvent control should be used in statistical analysis.[86] If, however both controls meet the validity criteria but adult survival or reproduction in the solvent control differs significantly from the results of the clean control soil, this might be indicative of a potential solvent interference that would then require additional evaluation to determine the impact on the interpretation of the study. The USEPA (2008) provides guidance on what might be included in such an evaluation: (1) assess the relevance of the solvent control response (i.e., percent change relative to the response in control soil); (2) the degree of statistical significance associated with the difference between the two controls (i.e., highly significant difference versus marginally significant difference); (3) assess the breadth of the interference; (4) assess any other potential cause for the interference observed in the solvent control; and (5) assess the impact of the potential solvent control interference on uncertainty in the risk estimate.

---

[85] 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.

[86] Evidence to date with aquatic test organisms (Hutchinson *et al*., 2006) has shown that solvents rarely exert a direct effect on the test organism. However, if there was an effect of the solvent on the test organism, these effects would almost always be additive with the test substance, and the use of the solvent control compensates for this (Green, 2014). In addition, there could be an interaction between the test substance and the solvent that modifies toxicity. It is difficult to definitively show that this interaction is absent or present, because the test substance is not evaluated in the absence of the solvent. For this reason, the solvent control is the appropriate choice for comparisons (OECD, 2006).

Section 7

Reporting Requirements

Each test-specific report must indicate whether 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 that 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 (Section 7.2).

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, horizon, negative control soil) or coding, as provided to the laboratory personnel;
- information on labelling or coding of each sample;
- brief description of soil sampling, storage, and preparation (i.e., pretreatment) procedures;
- information on sample horizons as they were collected (i.e., number, relative depth of each soil horizon) for test, reference, and negative control soils, if applicable;
- type of negative control soil (natural or artificial) and, if applicable, reference soil;
- date of sample collection; date and time sample(s) received at test facility; and
- sample temperature and moisture content upon receipt at the test facility.

7.1.2 Test Organisms

- species and source of breeding stock and test organisms;
- age range of organisms 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
- person(s) performing the test (or each component of the test) and verifying results.

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., soil manipulation; 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 the 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 deviations from, or exclusion of, any of the procedures and conditions specified in this document;
- number of discrete samples per treatment; number of replicate test vessels for each treatment; number and description of treatments in each test including the control(s); test concentrations (if applicable);
- volume and depth of soil in each test vessel;
- number of organisms per test vessel and treatment;
- dates and times when test and control soils were prepared, test was started (i.e., organisms added to test and control soils), and test was ended;
- feeding regime and ration during the test;
- indication of test vessel aeration and assessment of soil moisture during the test;
- for each soil sample, any measurements of soil particle size, moisture content, water-holding capacity, pH, TOC, OM, CEC, and electrical conductivity; and
- for each composite sample or subsample taken at the same time from all replicates of each treatment, all measurements of temperature (air and soil), pH, moisture content, and water-holding capacity.

7.1.6 Test Results

- mean (± SD) percent survival of adult mites in each treatment, including control(s) and reference soils on Day 28; mean (± SD) number of surviving progeny in each treatment, including control(s) and reference soils on Day 28;
- any ICp (together with its 95% confidence limits) determined for the data on reproductive inhibition (i.e., number of surviving progeny in each treatment 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 and degree of difference from nominal strength;
- results for any 28-day IC50 (including its 95% confidence limits) or % reduction in progeny production relative to the control, for multi-concentration tests or positive controls, respectively, performed with the reference toxicant in conjunction with the definitive soil toxicity test; geometric mean value (± 2 SD) for the same reference toxicant, as derived at the test facility in previous tests with the reference toxicant using the procedures and conditions for testing with a reference toxicant described herein; and
- anything unusual about the test, any problems encountered, any remedial measures taken.

7.2 Additional Reporting Requirements

This section provides a list of items that must be either included in the test-specific report or the

general report, or held on file for a minimum of five years. Filed information must include the following, if available:

- a record of the chain of custody 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; 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.

7.2.1 Test Substance or Material

- name and signature of person(s) who collected and/or provided the sample;
- records of sample log-entry sheets;
- appearance (e.g., odour, colour) and conditions (e.g., in darkness, in sealed container) of sample upon receipt and during storage; and
- any additional records obtained for field (e.g., field records provided or maintained during sample collection) or chemical samples (impurities, additives, structural formulae, etc.).

7.2.2 Test Organisms

- records and methods used for taxonomic confirmation of test species;
- history and age of breeding stock for any culture used to provide test organisms;
- description of culture conditions and procedures for all lab cultures (including age-synchronized cultures), including temperature, lighting, type and amount of substrate and details on its periodic renewal, methods, and records for aeration and substrate hydration; measurements and records of substrate quality, density of mites, records of culture condition, health and performance indices; and any acclimation conditions and procedures (e.g., substrate and temperature), including rate of change;
- procedures used for preparation of age-synchronized cultures;
- 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 3.2.1);
- description of systems for providing lighting and for regulating temperature within the test facility;
- description of test vessels 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

- photoperiod and measurements of light intensity adjacent to the surface of the soil in test vessels;
- procedure for adding test organisms to test vessels;
- appearance of each sample (or mixture thereof) in test vessels; changes in appearance noted during test;
- records of the addition of test water on the surface of the soil in each test vessel throughout the test for increasing moisture content;
- record of any growth of bacteria or fungi, and the presence and estimated quantity of any uneaten food;
- description of procedures used for heat-extraction of test organisms and records of the time and temperatures achieved during heat-extraction;
- procedures used to assess and validate the efficiency of the heat-extraction procedure and records demonstrating the establishment and ongoing monitoring of the heat-extraction efficiency;
- any other physicochemical measurements (e.g., analyses of aliquots from the same batch to determine homogeneity; contaminant concentration, cations and anions, nitrogen, nitrate, nitrite, ammonia, phosphorus, potassium, C:N ratio, bulk density, total volatile solids, biochemical oxygen demand, chemical oxygen demand, total inorganic carbon, redox potential, soluble salts, sodium adsorption ratio) made before and during the test on test material (including negative control soil and reference soil) and contents of test vessels; including analyses of whole soil and porewater;
- 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 chemicals in stock solution(s) of the reference toxicant and, if measured, in test concentrations.

7.2.7 Test Results

- results for any range-finding test(s) conducted;
- number of surviving adult mites in each test vessel at test end; number of surviving progeny in each test vessel at test end; for regression analyses, 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 historical results for 28-day reference toxicity tests or positive control concentrations with the reference toxicant; CV for mean historical data derived for reference toxicity tests or positive control concentrations performed using the reference toxicant; and

- graphical presentation of data.

References

AAFC (Agriculture and Agri-Food Canada). 1998. The Canadian System of Soil Classification. 3rd ed. Soil Classification Working Group, Agriculture and Agri-Food Canada Publication 1646, National Research Council (NRC) Research Press, Ottawa. 187 p.

AE (Alberta Environment). 2007. Tier 2 Eco-contact Guideline Derivation Protocol, July 13, 2007 (Online Version). 33 p.

Adamski Z, Bloszyk J, Bruin J, and Ziemnicki K. 2007. Non-omnia Moriantur – Toxicity of Mancozeb on Dead Wood Microarthropod Fauna, Experimental and Applied Acarology, 42:47–53.

AESA (Alberta Environmentally Sustainable Agriculture Program). 2001. AESA Soil Quality Benchmark Study – Soil Organic Matter, Fact Sheet FS2001-ISQ, AESA Soil Quality Program, Conservation and Development Branch, Alberta Agriculture, Food and Rural Development, Edmonton, Alberta. 4 p.

Al-Assiuty AIM, and Khalil MA. 1995. The Influence of Insecticide-Pheromone Substitution on the Abundance and Distributional Pattern of Soil Oribatid Mites, Experimental and Applied Acarology, 19:399–410.

Al-Assiuty AIM, Khalil MA and Abdel-Lateif HM. 2000. Effects of Dry Sludge Application on Soil Microarthropod Communities in a Reclaimed Desert Ecosystem, Pedobiologia, 44:567–578.

AquaTerra Environmental Ltd. 1998. Development of a Reproduction Toxicity Test with *Onychiurus folsomi* for Assessment of Contaminated Soils, Report Prepared by AquaTerra Environmental, Ltd. (Orton, Ontario) for Environment Canada, Ottawa, Ontario. 253 p.

ASTM (American Society for Testing and Materials). 2004. Standard Guide for Conducting Laboratory Soil Toxicity or Bioaccumulation Tests with the Lumbricid Earthworm *Eisenia fetida* and the Enchytraeid Potworm *Enchytraeus albidus*, ASTM International, West Conshohocken, Pennsylvania, ASTM E1676-04. 27 p.

ASTM. 2014. Standard Guide for Use of Lighting in Laboratory Testing, ASTM International, West Conshohocken, Pennsylvania, ASTM E1733-95(2014). 12 p.

Battigelli JP, Spence JR, Langor DW, and Berch SM. 2004. Short-term Impact of Forest Soil Compaction and Organic Matter Removal on Soil Mesofauna Density and Oribatid Mite Diversity, Canadian Journal of Forest Research 34:1136–1149.

Becker-van Slooten K, Campiche S, and Tarradellas J. 2003. Research in Support of the Environment Canada Collembolan Toxicity Test Method with *Folsomia candida* for Assessment of Contaminated Soils, Report Prepared by Laboratory of Environmental Chemistry and Ecotoxicology (CECOTOX), École polytechnique fédérale de Lausanne (EPFL), ENAC-ISTE (Lausanne, Switzerland), for Environment Canada, Ottawa, Ontario.

Becker-van Slooten K, Campiche S, and Tarradellas J. 2004. 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), École polytechnique fédérale de Lausanne (EPFL), ENAC-ISTE (Lausanne, Switzerland), for Environment Canada, Ottawa, Ontario. 29 p.

Becker-van Slooten K, Auroy A, Stämpfli C, and Tarradellas J. 2005. Further Research in Support of the Environment Canada Collembola Toxicity Test Method with *Folsomia candida* and *Folsomia fimetaria* for Assessment of Contaminated Soils and Soils Amended with Substances, Report Prepared by Laboratory of Environmental Chemistry and Ecotoxicology (CECOTOX), École polytechnique fédérale de Lausanne (EPFL),

ENAC-ISTE (Lausanne, Switzerland), for Environment Canada, Ottawa, Ontario. 36 p.

Behan VM. 1978. Diversity, Distribution and Feeding Habits of North American Arctic Soil Acari. Doctoral thesis – McGill University. 428 p.

Behan VM, Hill SB, and Kevan DKM. 1978. Effects of Nitrogen Fertilizers, as Urea, on Acarina and Other Arthropods in Quebec Black Spruce Humus, Pedobiologia, 18:249–263.

Behan-Pelletier VM. 1999. Oribatid Mite Biodiversity in Agroecosystems: Role for Bioindication, Agriculture, Ecosystems and Environment, 74:411–423.

Blakely JK, Neher DA, and Spongberg AL. 2002. Soil Invertebrate and Microbial Communities, and Decomposition as Indicators of Polycyclic Aromatic Hydrocarbon Contamination, Applied Soil Ecology, 21:71–88.

Bonnell (Bonnell Environmental Consulting). 1994. Assessment of Soil Toxicity Test Species for Canadian Representativeness, Technical Report TS-28, Prepared by Bonnell Environmental Consulting for Environment Canada, Ottawa, Ontario.

Cannon RJC, and Block W. 1988. Cold Tolerance of Microarthropods, Biological Reviews Cambridge Philosophical Society, 63:23–77.

Carter MR (ed.). 1993. *Soil Sampling and Methods of Analysis*, Lewis Publishers, CRC Press Inc., Boca Raton, Florida.

Carter MR, and Gregorich EG (eds.). 2008. *Soil Sampling and Methods of Analysis*, *Second Edition*, Taylor & Francis Group, CRC Press, Boca Raton, Florida.

CCME (Canadian Council of Ministers of the Environment). 1994. A Framework for Ecological Risk Assessment at Contaminated Sites in Canada: Review and Recommendations, The National Contaminated Sites Program, Scientific Series 1999, Ecosystem Conservation Directorate, Evaluation and Interpretation Branch, Environment Canada, Ottawa, Ontario. 108 p.

CCME. 1996. A Framework for Ecological Risk Assessment: General Guide, CCME Subcommittee on Environmental Quality Criteria for Contaminated Sites, The National Contaminated Sites Program, Winnipeg, Manitoba. 32 p.

CCME. 1997. A Framework for Ecological Risk Assessment: Technical Appendices, CCME Subcommittee on Environmental Quality Criteria for Contaminated Sites, The National Contaminated Sites Program, Winnipeg, Manitoba. 67 p.

Coleman DC, Crossley DA Jr, and Hendrix PF. 2004. *Fundamentals of Soil Ecology*. Academic, London, United Kingdom.

Cortet J, Gillon D, Joffre R, Ourcival J-M, and Poinsot-Balaguer N. 2002. Effects of Pesticides on Organic Matter Recycling and Microarthropods in a Maize Field: Use and Discussion of the Litterbag Methodology, European Journal of Soil Biology, 38:261–265.

Courchesne F, Savoie S, and Dufresne A. 1995. Effects of Air-Drying on the Measurement of Soil pH in Acidic Forest Soils of Quebec, Canada, Soil Science, 160:56–68.

Crépin J, and Johnson RL. 1993. Soil Sampling for Environmental Assessment, In: *Soil Sampling and Methods of Analysis*, Carter MR (ed.), Lewis Publishers, CRC Press Inc., Boca Raton, Florida, pp. 5–18.

Crossley DA Jr, and Bohnsack KK. 1960. Long-Term Ecological Study in the Oak Ridge Area: III. The Oribatid Mite Fauna in Pine Litter, Ecology, 41:628–638.

Deitzer G. 1994. Spectral Comparisons of Sunlight and Different Lamps, In: *Proceedings of International Lighting in Controlled Environments Workshop*, Tibbits TW (ed.), Madison, Wisconsin, pp. 197–199.

de Lima e Silva C, Brennan N, Brouwer JM, Commandeur D, Verweij RA and van Gestel CAM. 2017. Comparative Toxicity of

Imidacloprid and Thiacloprid to Different Species of Soil Invertebrates, Ecotoxicology, 26:555–564.

Denton D, Diamond J and Zheng L. 2011. Test of Significant Toxicity: A Statistical Application for Assessing Whether an Effluent or Site Water is Truly Toxic, Environmental Toxicology and Chemistry, 30:1117–1126.

Dunnett CW. 1955. A Multiple Comparison Procedure for Comparing Several Treatments with a Control, Journal of the American Statistical Association, 50:1096–1121.

EC (Environment Canada). 1995. Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/30. 56 p.

EC. 1999. Guidance Document on Application and Interpretation of Single-species Tests in Environmental Toxicology, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/34. 203 p.

EC. 2004a. Biological Test Method: Tests for Measuring the Toxicity of Contaminated Soil to Earthworms (*Eisenia andrei*, *Eisenia fetida*, or *Lumbricus terrestris*), Environment Canada, Ottawa, Ontario, Report EPS 1/RM/43. 156 p.

EC. 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, British Columbia, Report published by Environment Canada, Ottawa, Ontario. 69 p.

EC. 2005a. Biological Test Method: Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/45, including June 2007 amendments. 131 p.

EC. 2005b. Guidance Document on Statistical Methods for Environmental Toxicity Tests, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/46. 241 p.

EC. 2007a. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia* 2nd edition, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/21. 100 p.

EC. 2007b. Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor,* 2nd edition, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/37. 141 p.

EC. 2010. Development and Standardization of New Toxicity Test Methodologies and Guidance for Assessment and Remediation of Impacts from Hydrocarbon and Brine Contamination on Boreal Forest, Taiga and Northern Soil using Organisms Representative of the Eco-zone Regions, Five Year Technical Progress Report, Environment Canada, Ottawa, Ontario. 167 p.

EC. 2011a. Biological Test Method: Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars), 2nd edition, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/27. 137 p.

EC. 2011b. Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows, 2nd edition, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/22. 100 p.

EC. 2012. Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/53. 134 p.

EC. 2013a. Biological Test Method: Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/56. 112 p.

EC. 2013b. Development and Standardization of Toxicity Test Methodologies for Assessment of Impacts from Contamination using Species and Microbial Communities Representative of Canadian Boreal and Taiga Eco-Zones, Technical

Report: 2010–2013, Environment Canada, Ottawa, Ontario. 245 p.

EC. 2014a. Biological Test Method: Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil, 2nd edition, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/47. 151 p.

EC. 2014b. Estimation of Chemical Bioavailability to Soil Organisms with CMP Substances from the Medium Priority List: Comparison of the Toxicity of Zinc Moieties, Technical Report, Environment Canada, Ottawa, Ontario. 148 p.

ECASG (Ecological Criteria Advisory Sub Group). 2006. Five-Year Review of the Canada-Wide Standards for Petroleum Hydrocarbons (PHC CWS); Ecological, Direct Soil Contact Guidance, Report to the Canadian Council of Ministers of Environment Soil Quality Guidelines Task Group. 5 p.

ECCC (Environment and Climate Change Canada) 2018. Development of a New Environment Canada Test Method for Measuring Reproduction in Soil Using the Oribatid Mite, *Oppia nitens*, Technical Report, Environment and Climate Change Canada, Ottawa, Ontario. 43 p.

ECCC. 2019. Inter-Laboratory Validation of Environment and Climate Change Canada's New Test Method for Measuring Reproduction of Oribatid Mites Exposed to Contaminants in Soil, Technical Report, Environment and Climate Change Canada, Ottawa, Ontario. 94 p.

EcoDynamics Consulting Inc. 2007. Sampling in the Swan Hills and Peace River Areas of Alberta and the Torch River Area of Saskatchewan, Prepared for Environment Canada, Ottawa, Ontario. 23 p.

EcoDynamics Consulting Inc. 2011a. Ecological Description of the Ontario Reference Site, Prepared for Environment Canada, Ottawa, Ontario.

EcoDynamics Consulting Inc. 2011b. Soil Sampling and Ecological Description at the Newfoundland Reference Site, Prepared for Environment Canada, Ottawa, Ontario. 8 p.

Erstfeld KM, and Snow-Ashbrook J. 1999. Effects of Chronic Low-Level PAH Contamination of Soil Invertebrate Communities, Chemosphere, 39:2117–2139.

Gainer A, Cousins M, Hogan N, and Siciliano SD. 2018. Petroleum Hydrocarbon Mixture Toxicity and a Trait-Based Approach to Soil Invertebrate Species for Site-Specific Risk Assessments, Environmental Toxicology and Chemistry, 37:2222–2234.

Gainer A, Hogan N, and Siciliano SD. 2019. Soil Invertebrate Avoidance Behaviour Identifies Petroleum Hydrocarbon Contaminated Soils Toxic to Sensitive Plant Species, Journal of Hazardous Materials, 361:338–347.

Gergócs V, and Hufnagel L. 2009. Application of Oribatid Mites as Indicators: A Review, Applied Ecology and Environmental Research, 7:79–98.

Green JW. 2014. Power and Control Choice in Aquatic Experiments with Solvents, Ecotoxicology and Environmental Safety, 102:142–146.

Greenslade P, and Vaughan GT. 2003. A Comparison of Collembola Species for Toxicity Testing of Australian Soils, Pedobiologia, 47:171–179.

Hansen RA. 2000. Effects of Habitat Complexity and Composition on a Diverse Litter Microarthropod Assemblage, Ecology, 81:1120–1132.

Hausenbuiller RL. 1985. *Soil Science –Principles and Practices, 3rd Edition*, W.C. Brown Publisher, Dubuque, Iowa.

Heethoff M, Laumann M, and Bergmann P. 2007. Adding to the Reproductive Biology of the Parthenogenetic Oribatid Mite, *Archegozetes longisetosus* (Acari, Oribatida, Trhypochthoniidae), Turkish Journal of Zoology, 31:151–159.

Hendershot WH, Lalande H, and Duquette M. 1993. Soil Reaction and Exchangeable Acidity, In: *Soil Sampling and Methods of Analysis*, Carter MR

(ed.), Lewis Publishers, Boca Raton, Florida. pp. 141–145.

Heneghan L, Coleman DC, Zou X, Crossley DA Jr, and Haines BL. 1999. Soil Microarthropod Contributions to Decomposition Dynamics: Tropical-Temperate Comparisons of a Single Substrate, Ecology, 80:1873–1882.

Hennessy R. 2010. Boreal Invertebrate Testing Towards: Development of Tier 2 Soil Eco-Contact Guidelines for Field Wide Application using Boreal Forest Species. Technical Report Prepared for WorleyParsons – Infrastructure & Environment, Calgary, Alberta. 46 p.

Hoenig JM, and Heisey DM. 2001. The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis, The American Statistician, 55:19–24.

Holst RW, and Ellanger TC. 1982. Pesticide Assessment Guidelines. Subdivision J. Hazard Evaluation: Non-Target Plants, Office of Pesticides and Toxic Substances, United States Environmental Protection Agency, Washington, DC, Report EPA–540/9-82-020. 60 p.

Huguier P, Manier N, Owojori OJ, Bauda P, Pandard P, and Römbke J. 2015. The Use of Soil Mites in Ecotoxicology: A Review, Ecotoxicology, 24:1–18.

Hutchinson TH, Shillabeer N, Winter MJ, and Pickford DB. 2006. Acute and Chronic Effects of Carrier Solvents in Aquatic Organisms: A Critical Review, Aquatic Toxicology, 76:69–92.

ISO (International Organization for Standardization). 1995. Soil Quality – Determination of Organic and Total Carbon After Dry Combustion (Elementary Analysis), Geneva, Switzerland, ISO 10694. 7 p.

ISO. 1999. Soil Quality – Inhibition of Reproduction of Collembola (*Folsomia candida*) by Soil Pollutants, Geneva, Switzerland, ISO 11267. 19 p.

ISO. 2005. Soil quality – Vocabulary. Geneva, Switzerland, ISO 11074. 113 p.

ISO. 2019a. Soil Quality – Identification of Ecotoxicological Test Species by DNA Barcoding, Geneva, Switzerland, ISO 21286:2019. 28 p.

ISO. 2019b. Soil Quality – Inhibition of Reproduction of Soil Mite (*Hypoaspis aculeifer*) by Soil Contaminants, Geneva, Switzerland, ISO 21285:2019. 38 p.

ISO. 2019c. Soil Quality – Test for Measuring the Inhibition of Reproduction in Oribatid Mites (*Oppia nitens*) Exposed to Contaminants in Soil, Geneva, Switzerland, ISO 23266:2019. 37 p.

Johnston JM, and Crossley DA Jr. 2002. Forest Ecosystem Recovery in the Southeast US: Soil Ecology as an Essential Component of Ecosystem Management, Forest Ecology and Management, 155:187–203.

Keddy C, Greene JC and Bonnell MA. 1995. Review of Whole-Organism Bioassays: Soil, Freshwater Sediment, and Freshwater Assessment in Canada, Ecotoxicology and Environmental Safety, 30:221–251.

Köhler HR, Alberti G, Seniczak S, and Seniczak A. 2005. Lead-Induced Hsp70 and Hsp60 Pattern Transformation and Leg Malformation During Postembryonic Development in the Oribatid Mite, Archegozetes longisetosus Aoki, Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 141:398–405.

Lebrun P, and van Straalen NM. 1995. Oribatid Mites: Prospects for Their Use in Ecotoxicology, Experimental and Applied Acarology, 19:361–379.

Lenth RV. 2007. Statistical Power Calculations, Journal of Animal Science, 85(E. Suppl.):E24–E29.

Li J, Verweij RA, and van Gestel CAM. 2018. Lanthanum Toxicity to Five Different Species of Soil Invertebrates in Relation to Availability in Soil, Chemosphere, 193:412–420.

Luxton M. 1981. Studies on the Oribatid Mites of a Danish Beech Wood Soil IV. Developmental Biology, Pedobiologia, 21:312–340.

Maraun M, and Scheu S. 2000. The Structure of Oribatid Mite Communities (Acari, Oribatida): Patterns, Mechanisms and Implications for Future Research, Ecography, 23:374–383.

Michael AD. 1884. *British Oribatidae*. Printed by Adlard and Son for The Ray Society, London, pp. 409–411.

Michereff-Filho M, Guedes RNC, Della-Lucia TMC, Michereff MFF, and Cruz I. 2004. Non-Target Impact of Chlorpyrifos on Soil Arthropods Associated with No-Tillage Corn Fields in Brazil, International Journal of Pest Management, 50:91–99.

Minor MA, and Norton RA. 2004. Effects of Soil Amendments on Assemblages of Soil Mites (Acari: Oribatida, Mesostigmata) in Short-Rotation Willow Plantings in Central New York, Canadian Journal of Forest Research, 34:1417–1425.

NERI (National Environmental Research Institute). 1993. Manual of SECOFASE: Development, Improvement and Standardization of Test Systems for Assessing Sublethal Effects on Chemicals on Fauna in the Soil Ecosystem, Løkke H, and van Gestel CAM (eds.), Report from a Workshop held in Silkeborg, January 18–19, 1993, Silkeborg, Denmark. 44 p.

Newman MC. 2008. What Exactly are you Inferring? A Closer Look at Hypothesis Testing, Environmental Toxicology and Chemistry, 27:1013–1126.

Norberg-King TJ. 1993. A Linear Interpolation Method for Sublethal Toxicity: the Inhibition Concentration (ICp) Approach (Version 2.0), Report 03-93 of National Effluent Toxicity Assessment Center, U.S. Environmental Protection Agency, Environmental Research Laboratory, Duluth, Minnesota. 25 p.

Norton RA. 1985. Aspects of the Biology and Systematics of Soil Arachnids, Particularly Saprophagous and Mycophagous Mites, Quaestiones Entomologicae, 21:523–541.

Norton, RA. 1994. Evolutionary Aspects of Oribatid Mites Life Histories and Consequences for the Origin of the Astigmata, In: *Mites: Ecological and Evolutionary Analyses of Life-History Patterns*, Houck MA (ed.), Chapman & Hall, New York, pp. 99–135.

Norton RA, Bonamo PM, Grierson JD, and Shear WA. 1988. Oribatid Mite Fossils from a Terrestrial Devonian Deposit Near Gilboa, New York, Journal of Paleontology, 62:259–269.

OECD (Organization for Economic Cooperation and Development). 1984a. Test No. 207: Earthworm Acute Toxicity Tests. OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, France. 9 p.

OECD. 1984b. Test No. 208: Terrestrial Plants, Growth Test. OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, France. 6 p.

OECD. 2005. OECD Guideline for the Testing of Chemicals: Proposal for a New Guideline – Collembola Reproduction Test (*Folsomia fimetaria* and *Folsomia candida*). 24 p.

OECD. 2006. Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application, OECD Environment Health and Safety Publications Series on Testing and Assessment, No. 54, Paris, France. 147 p.

OECD. 2008. Test No. 226: Predatory Mite (*Hypoaspis* (*Geolaelaps*) *aculeifer*) Reproduction Test in Soil. OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, France.

OECD. 2009. OECD Guideline for the Testing of Chemicals: Collembolan Reproduction Test in Soil, No. 232, Paris, France. 19 p.

OECD. 2016. OECD Guideline for the Testing of Chemicals: Earthworm Reproduction Test (*Eisenia fetida/Eisenia Andrei*), No 222, Paris, France. 21 p.

OMEE (Ontario Ministry of Environment and Energy). 1996. 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, Queen's Printer, Toronto, Ontario. 158 p.

Osler GHR, Westhorpe D, and Oliver I. 2001. The Short-Term Effects of Endosulfan Discharges on Eucalypt Floodplain Soil Microarthropods, Applied Soil Ecology, 16:263–273.

Owojori OJ, and Siciliano SD. 2012. Accumulation and Toxicity of Metals (Copper, Zinc, Cadmium, and Lead) and Organic Compounds (Geraniol and Benzo[a]pyrene) in the Oribatid Mite *Oppia nitens*, Environmental Toxicology and Chemistry, 31:1639–1648.

Parmelee RW, Wentsel RS, Phillips CT, Checkai RT, and Simini M. 1993. Soil Microcosm for Testing the Effects of Chemical Pollutants on Soil Fauna Communities and Trophic Structure, Environmental Toxicology and Chemistry, 12:1477–1486.

Princz, JI. 2014. Development of a New Test Suite of Ecologically-Relevant Species for the Assessment of Contaminants in Boreal Soils – Special Emphasis on Oribatid Mites. Doctoral thesis – University of Saskatchewan. 140 p.

Princz JI, Behan-Pelletier VM, Scroggins RP, and Siciliano SD. 2010. Oribatid Mites in Soil Toxicity Testing – The Use of *Oppia nitens* (C.L. Koch) as a New Test Species, Environmental Toxicology and Chemistry, 29:971–979.

Princz JI, Moody M, Fraser C, Van der Vliet L, Lemieux H, Scroggins R, and Siciliano SD. 2012. Evaluation of a New Battery of Toxicity Tests for Boreal Forest Soils: Assessment of the Impact of Hydrocarbons and Salts, Environmental Toxicology and Chemistry, 31:766–777.

Princz J, Jatar M, Lemieux H, and Scroggins R. 2018. Perfluorooctane Sulfonate in Surface Soils: Effects on Reproduction in the Collembolan, *Folsomia candida*, and the Oribatid Mite, *Oppia nitens*, Chemosphere, 208:757–763.

Prinzing A, Kretzler S, Badejo A, and Beck L. 2002. Traits of Oribatid Mite Species that Tolerate Habitat Disturbance Due to Pesticide Application, Soil Biology and Biochemistry, 34:1655–1661.

Proctor H. 1998. Acariformes. The "Mite-Like" Mites. Version 09 August 1998. The Tree of Life Project. URL: *http://tolweb.org/Acariformes/2563/1998.08.09*. Last accessed in May 2016.

Ritchie E, Boyd P, Lawson-Halasz A, Hawari J, Saucier S, Scroggins R, and Princz J. 2017. The Ecotoxicity of Zinc and Zinc-Containing Substances in Soil with Consideration of Metal-Moiety Approaches and Organometal Complexes, Environmental Toxicology and Chemistry, 36:3324–3332.

Rocchini RJ, Clark MJR, Jordan AJ, Horvath S, McLeay DJ, Servizi JA, Sholund A, Singleton H.J, Watts RG, and Young RH. 1982. Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish, British Columbia Ministry of Environment, Victoria, British Columbia. 18 p.

Römbke J, Jänsch S, and Scroggins R. 2006. Identification of Potential Organisms of Relevance to Canadian Boreal Forest and Northern Lands for Testing of Contaminated Soils, Environmental Reviews, 14:137–167.

Römbke J, Aira M, Backeljau T, Breugelmans K, Domínguez J, Funke E, Graf N, Hajibabaei M, Pérez-Losada M, Porto PG, Schmelz RM, Vierna J, Vizcaíno A, and Pfenninger M. 2016. DNA Barcoding of Earthworms (*Eisenia fetida*/*andrei* Complex) From 28 Ecotoxicological Test Laboratories, Applied Soil Ecology, 104:3–11.

Sager JC, and McFarlane C. 1997. Radiation, In: *Plant Growth Chamber Handbook,* Langhans RW, Tibbits TW (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, pp. 1–30.

Sengbusch HG, and Sengbusch CH. 1970. Post-Embryonic Development of *Oppia nitens* (Acarina: Oribatei), Journal of the New York Entomological Society, 78:207–214.

Seniczak A. 2006. The Effect of Density on Life-History Parameters and Morphology of *Archegozetes longisetosus* Aoki, 1965 (Acari: Oribatida) in Laboratory Conditions, Biological Letters, 43:209–213.

Seniczak A, and Seniczak S. 2002. The Effect of Cadmium on *Archegozetes longisetosus* (Acari, Oribatida) in Laboratory Conditions, European Journal of Soil Biology, 38: 315–317.

Seniczak A, Seniczak S, and Kobiersk M. 2006. Long-Term Effect of Cadmium on the Oribatid Mite *Archegozetes longisetosus* Aoki, 1965 in Laboratory Conditions, Biological Letters, 43:237–242.

Seniczak A, Ligocka A, Seniczak S, and Paluszak Z. 2009. The Influence of Cadmium on Life-History Parameters and Gut Microflora of *Archegozetes longisetosus* (Acari: Oribatida) Under Laboratory Conditions, Experimental and Applied Acarology, 47:191–200.

Seniczak S. 1975. Morphology of Juvenile Stages of Some Oppiidae (Acarina, Oribatei) II., Pedobiologia, 15:262–275.

Seniczak S, and Stefaniak O. 1978. The Microflora of the Alimentary Canal of *Oppia nitens* (Acarina, Oribatei), Pedobiologia, 18:110–119.

Sheppard SC, and Evenden WG. 1998. An Approach to Defining a Control or Diluent Soil for Ecotoxicity Assays, In: *Environmental Toxicology and Risk Assessment*, Little EE, Greenberg BM, DeLonay AJ (eds.), 7th Vol, ASTM STP 1333, American Society for Testing and Materials, Philadelphia, Pennsylvania, pp. 215–226.

Singh M, Jain KL, Mathur RB, and Dogra D. 1996. Laboratory Study of Food Preferences of Some Cryptostigmatic Mites and Their Contribution in Litter Degradation and Mineralization in Soils, *Annals of Biology,* 12:335–343.

Smit CE, Moser T, and Römbke J. 2012. A New OECD Test Guideline for the Predatory Soil Mite *Hypoaspis aculeifer*: Results of an International Ring Test, Ecotoxicology and Environmental Safety, 82:56–62.

SPAC (Soil and Plant Analysis Council Inc.). 1992. *Soil Analysis Handbook of Reference Methods*, Soil and Plant Analysis Council Inc., Georgia University Station, Athens, Georgia, 202 p.

Stämpfli C, Becker-van Slooten K, Campiche S, and Tarradellas J. 2005. Research in Support of Environment Canada's Toxicity Test Method with *Folsomia fimetaria* for Assessment of Contaminated Soils, Report prepared by Laboratory of Environmental Chemistry and Ecotoxicology (CECOTOX), École polytechnique fédérale de Lausanne (EPFL), ENAC-ISTE (Lausanne, Switzerland), for Environment Canada, Ottawa, Ontario.

Stark JD. 1992. Comparison of the Impact of a Neem Seed-Kernel Extract Formulation 'Margosan-O' and Chlorpyrifos on Non-Target Invertebrates Inhabiting Turf Grass, Pesticide Science, 36:293–299.

Stefaniak O, and Seniczak S. 1981. The Effect of Fungal Diet on the Development of *Oppia nitens* (Acari, Oribatei) on the Microflora of its Alimentary Tract, Pedobiologia, 21:202–210.

Stephenson GL. 2003. Unpublished Data. Stantec Consulting Ltd., Guelph, Ontario.

Stephenson GL, Koper N, Atkinson GF, Solomon KR, and Scroggins RP. 2000. Use of Nonlinear Regression Techniques for Describing Concentration-Response Relationships of Plant Species Exposed to Contaminated Site Soils, *Environmental Toxicology and Chemistry*, 19:2968–981.

Stephenson GL, Kuperman RG, Linder GL, and Visser S. 2002. Toxicity Tests for Assessing Contaminated Soils and Ground Water, In: *Environmental Analysis of Contaminated Sites*, Sunahara GI, Renoux AY, Thellen C, Gaudet CL, Pilon A (eds.), Ecological and Environmental Toxicology Series, John Wiley and Sons Ltd., Chichester, United Kingdom, pp. 25–43.

Stephenson G, Feisthauer N, Bessie K, and Scroggins R. 2008. Terrestrial Toxicity Testing in Support of Site-specific Risk Assessments: An Integrative Tool for Contaminated Site

Management, Proceedings of the Federal Contaminated Sites National Workshop, Vancouver, British Columbia, April 28–May 2, 2008, hosted by the Real Property Institute of Canada.

Sverdrup LE, Hagen SB, Krogh PH, and van Gestel CAM. 2007. Benzo[a]pyrene Shows Low Toxicity to Three Species of Terrestrial Plants, Two Soil Invertebrates, and Soil-Nitrifying Bacteria, Ecotoxicology and Environmental Safety, 66:362–368.

Taberly G. 1987. Recherches sur la Parthénogenèse Thélytoque de Deux Espèces D'acariens Oribatides: *Trypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). III. Étude Anatomique, Histologique et Cytologique des Femelles Parthénogénétiques, Acarologia, 28:389–403.

Thursby GB, Heltshe J, and Scott KJ. 1997. Revised Approach to Toxicity Test Acceptability Criteria Using a Statistical Performance Assessment, Environmental Toxicology and Chemistry, 16:1322–1329.

TOXSTAT. 1996. "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.]

USEPA (United States Environmental Protection Agency). 1989. Protocols for Short Term Toxicity Screening of Hazardous Waste Sites, Corvallis Environ. Research Lab., Corvallis, Oregon, Report EPA/600/3-88/029. 102 p.

USEPA. 1995. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms, Office of Research and Development, USEPA, Washington, DC, Report EPA/600/R-95/136. 661 p.

USEPA. 2002. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms, 4th ed., Office of Water, USEPA, Washington, DC, Report EPA 821-R-02-014. 464 p.

USEPA. 2008. Guidance for the Use of Dilution-Water (negative) and Solvent Controls in Statistical Data Analysis for Guideline Aquatic Toxicology Studies, Memo from Statistics Workgroup and Aquatic Biology Technical Team to Donald Brady, Director, Environmental Fate and Effects Division. Office of Pesticides Programs, Office of Chemical Safety and Pollution Prevention (OCSPP), Washington, DC.

USEPA. 2010. National Pollutant Discharge Elimination System Test of Significant Toxicity Implementation Document, Office of Wastewater Management, Washington, DC, Report EPA 833-R-10-003.

van Dam RA, Harford AJ, and Warne M St.J. 2012. Time to Get Off the Fence: The Need for Definitive International Guidance on Statistical Analysis of Ecotoxicity Data, Debate and Commentary, Integrated Environmental Assessment and Management, 8:242–245.

Van der Hoeven N. 1998. Power Analysis for the NOEC: What is the Probability of Detecting Small Toxic Effects on Three Different Species using the Appropriate Standardized Test Protocols?, Ecotoxicology, 7:355–361.

Van der Vliet L, Taylor LN, and Scroggins R. 2012. NOEC: Notable Oversight of Enlightened Canadians: A Response to van Dam *et al*. (2012), Letter to the Editor, Integrated Environmental Assessment and Management, 8:397–398.

van Gestel CAM, and Doornekamp A. 1998. Tests on the Oribatid Mite *Platynothrus peltifer*. In *Handbook of Soil Invertebrate Toxicity Tests*, Løkke H, van Gestel CAM (eds.), John Wiley & Sons, Chichester, United Kingdom, pp. 113–130.

Walter DE. 1985. The Effects of Litter Type and Elevation on Colonization of Mixed Coniferous Litterbags by Oribatid Mites, Pedobiologia, 28:383–387.

Walter DE, Gerald K, and Evert L. 1996. Acari. The Mites. Version 13 December 1996. The Tree of Life Web Project. URL: *http://tolweb.org/Acari/2554/1996.12.13.* Accessed in May 2016.

Webb NR. 1989. Observations on the Life Cycle of *Steganacarus magnus* (Acari: Cryptostigmata), Pedobiologia, 33:293–299.

Wiles JA, and Krogh PH. 1998. Tests with the Collembolans *Isotoma viridis, Folsomia candida,* and *Folsomia fimetaria,* In: *Handbook of Soil Invertebrate Toxicity Tests*, Løkke H, van Gestel CAM (eds.), John Wiley, London, United Kingdom, pp. 131–156.

Williams DA. 1972. The Comparison of Several Dose Levels with a Zero Dose Control, Biometrics, 28:519–532.

Woodring JP, and Cook EF. 1962. The Biology of *Ceratozetes cisalpinus* Berlese, *Scheloribates laevigatus* Koch and *Oppia neerlandica* Oudemans (Oribatei) with Description of All Stages, Acarologia, 4:101–137.

Yu L, Berry RE, and Croft BA. 1997. Effects of *Bacillus thuringiensis* Toxins in Transgenic Cotton and Potato on *Folsomia candida* (Collembola: Isotomidae) and *Oppia nitens* (Acari: Oribatidae), Journal of Economic Entomology, 90:113–118.

Zajdlik and Associates Inc. 2010. Statistical Guidance for Ecotoxicological Data Derived Using Environment Canada Standardized Biological Test Methods, unpublished report prepared by Zajdlik and Associates Inc. (Rockwood, ON) for Environment Canada (Ottawa, ON). 136 p.

Appendix A

Biological Test Methods and Supporting Guidance Documents Published by Environment and Climate Change Canada's Method Development and Applications Unit[a]

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 and May 2007
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 2nd Edition	February 2007	–
Test of Larval Growth and Survival Using Fathead Minnows	EPS 1/RM/22 2nd Edition	February 2011	–
Toxicity Test Using Luminescent Bacteria (*Photobacterium phosphoreum*)	EPS 1/RM/24	November 1992	–
Growth Inhibition Test Using a Freshwater Alga	EPS 1/RM/25 2nd Edition	March 2007	–
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 2nd Edition	February 2011	–
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	–

[a] These documents are available for purchase from the Publication Catalogue, Environment and Climate Change Canada, Ottawa ON K1A 0H3, Canada. Printed copies can also be requested by email from ec.enviroinfo.ec@canada.ca. These documents are available free of charge in PDF format at the following website: www.ec.gc.ca/faunescience-wildlifescience/default.asp?lang=En&n=0BB80E7B-1. For further information or comments, contact the Chief, Biological Assessment and Standardization Section, Environment and Climate Change Canada, Ottawa ON K1A 0H3.

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods (continued)			
Test for Survival and Growth in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*	EPS 1/RM/33 3rd Edition	September 2017	–
Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*	EPS 1/RM/37 2nd Edition	January 2007	–
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	June 2007
Tests for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil	EPS 1/RM/45	February 2005	June 2007
Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil	EPS 1/RM/47 2nd Edition	February 2014	–
Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region	EPS 1/RM/56	August 2013	–
B. Reference Methods[b]			
Reference Method for Determining Acute Lethality Using Threespine Stickleback	EPS 1/RM/10 2nd Edition	December 2017	–
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 2nd Edition	December 2000	May 2007 and February 2016
Reference Method for Determining Acute Lethality of Effluents to Daphnia magna	EPS 1/RM/14 2nd Edition	December 2000	February 2016
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	–
Reference Method for Measuring the Toxicity of Contaminated Sediment to Embryos and Larvae of Echinoids (Sea Urchins or Sand Dollars)	EPS 1/RM/58	July 2014	–
Reference Method for Determining Acute Lethality Using *Acartia tonsa*	STB 1/RM/60	June 2019	–

[b] 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 that is described precisely in a written document. Unlike other generic (multipurpose or "universal") biological test methods published by Environment and Climate Change 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 2nd Edition	December 2016	–
Guidance Document on Statistical Methods for Environmental Toxicity Tests	EPS 1/RM/46	March 2005	June 2007
Procedure for pH Stabilization During the Testing of Acute Lethality of Wastewater Effluent to Rainbow Trout	EPS 1/RM/50	March 2008	–
Supplementary Background and Guidance for Investigating Acute Lethality of Wastewater Effluent to Rainbow Trout	–	March 2008	–
Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing	EPS 1/RM/53	February 2012	–
Procedure for pH Stabilization During the Testing of Acute Lethality of Pulp and Paper Effluent to Rainbow Trout	STB 1/RM/59	March 2018	–
Supplementary Guidance for Investigating Acute Lethality of Pulp and Paper Mill Effluents due to Ammonia	–	March 2018	–

Appendix B

Environment and Climate Change Canada, Regional Environmental Testing Laboratories

Atlantic Laboratory for Environmental Testing
Environmental Science Building
443 Université Avenue, Université de Moncton
Moncton, New Brunswick
E1A 3E9

Pacific and Yukon Laboratory for Environmental Testing
Pacific Environmental Science Centre
2645 Dollarton Hwy
North Vancouver, British Columbia
V7H 1B1

Québec Laboratory for Environmental Testing
105 McGill Street
Montréal, Quebec
H2Y 2E7

Prairie and Northern Laboratory for Environmental Testing
Northern Forestry Building
5320 122 St NW
Edmonton, Alberta
T6H 3S5

For current regional laboratory contact information please contact:

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario
K1A 0H3
Email: ec.methodes-methods.ec@canada.ca

Appendix C

Members of the Inter-Governmental Ecotoxicological Testing Group (as of June 2019)

Federal, Environment and Climate Change Canada

Suzanne Agius
Marine Protection Programs Section
Gatineau, Quebec

Adrienne Bartlett
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Lee Beaudette
Biological Assessment & Standardization Section
Ottawa, Ontario

Rene Beaulieu
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Christian Blaise (Emeritus)
Centre St. Laurent
Montréal, Quebec

Patrick Boyd
Biological Assessment & Standardization Section
Ottawa, Ontario

Lorraine Brown
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Joy Bruno
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Julia Brydon
Marine Protection Programs Section
Gatineau, Quebec

Craig Buday
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Melanie Camplin
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Marshneil Chandra
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ajith Dias Samarajeewa
Biological Assessment & Standardization Section
Ottawa, Ontario

Heather Dillon
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ken Doe (Emeritus)
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Tamzin El-Fityani
National Guidelines and Standards Office
Ottawa, Ontario

François Gagné
Fluvial Ecosystem Research
Montréal, Quebec

Patricia Gillis
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Christina Heise
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Natasha Hostal
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Paula Jackman
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Stephanie Kvas
Biological Assessment & Standardization Section
Ottawa, Ontario

Christopher Le
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Heather Lemieux
Biological Assessment & Standardization Section
Ottawa, Ontario

Michelle Linssen-Sauvé
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Matthew Meier
Biological Assessment & Standardization Section
Ottawa, Ontario

Danielle Milani
Aquatic Ecosystem Impacts Research Division
Burlington, Ontario

Rachel Miliano
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Joanne Parrott
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Linda Porebski
Marine Protection Programs Section
Gatineau, Quebec

Juliska Princz
Biological Assessment & Standardization Section
Ottawa, Ontario

Rick Scroggins
Biological Assessment & Standardization Section
Ottawa, Ontario

David Taillefer
Marine Environmental Protection
Gatineau, Quebec

Sylvain Trottier
Quebec Laboratory for Environmental Testing
Montréal, Quebec

Graham van Aggelen
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Leana Van der Vliet
Biological Assessment & Standardization Section
Ottawa, Ontario

Brian Walker
Quebec Laboratory for Environmental Testing
Montréal, Quebec

Peter Wells (Emeritus)
Environmental Conservation Service
Dartmouth, Nova Scotia

Federal, Natural Resources Canada

Philippa Huntsman-Mapila
Ecosystem Risk Management Program
Mining & Mineral Sciences Laboratory
CANMET, NRCan
Ottawa, Ontario

Morgan King
Ecosystem Risk Management Program
Mining & Mineral Sciences Laboratory
CANMET, NRCan
Ottawa, Ontario

Carrie Rickwood
Ecosystem Risk Management Program
Mining & Mineral Sciences Laboratory
CANMET, NRCan
Ottawa, Ontario

Provincial

Lisa Kennedy (co-Chair)
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Jennifer Koene-Fenton
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Heather Osachoff
British Columbia Ministry of the Environment
Victoria, British Columbia

David Poirier (Emeritus)
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Éloïse Veilleux
Centre d'expertise en analyse environnementale du Québec
Ste. Foy, Quebec

Trudy Watson-Leung (co-Chair)
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Appendix D

Natural and Artificial Negative Control Soils Used for Method 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 that is essentially free of any contaminants that could adversely affect the performance of test organisms during the test (see Section 3.3). Before applying the test method described in this document as a standardized test to be conducted according to Environment and Climate Change 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. Fourteen types of negative control soils were used to develop this biological test method and to further assess its robustness with samples of soil that varied considerably in their physical and chemical characteristics (EC, 2010, 2013b, 2014b; Hennessy, 2010; Princz *et al*., 2010, 2012, 2018; Princz, 2014; Ritchie *et al*., 2017; ECCC, 2018, 2019). Of these, 11 soils were used to establish reasonable criteria for valid test results, based on control performance (ECCC, 2018). The 11 soils tested include 2 artificial soils, 4 natural agronomic soils, and 5 natural soils from the boreal and taiga ecozones. The physicochemical characteristics of the 2 artificial and 4 natural agronomic soils are summarized in Table D-1 of this appendix while the 5 natural boreal and taiga soils are characterized in Table D-2.

One of the artificial control soils used in this series of performance evaluation studies with diverse soil types was the same formulated soil as that recommended for use herein (see Section 3.3.2). It consists of 70% silica sand, 20% kaolin clay, 10% *Sphagnum* sp. Peat, and calcium carbonate (10−30 g $CaCO_3$/kg peat). The soil was formulated by mixing the ingredients in their dry form thoroughly, then gradually hydrating with deionized 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 ISO (1999) and OECD (2005). The second artificial soil is an augmented sandy soil that was collected in 2013 from Greely Sand and Gravel in Ottawa, Ontario. The soil consists of an 80:20 composition of City of Ottawa yard waste compost to red sand mixture. Once collected, the soil was dried, sieved (2-mm mesh), and homogenized. Some condensation had been observed on the interior of some of the pails in storage along with the occasional smell of mould; this was not unexpected in a matrix composed of high quantities of decomposing leaf litter. Aeration and rehomogenization were attempted as a means of disrupting the growth of undesirable microbial organisms and eliminating condensation. The physicochemical characteristics of the two artificial soils are presented in Table D-1.

The four agronomic soils used as negative control soil while developing this biological test method and establishing the test validity criteria (see Section 4.4) do not represent all Canadian soil types. However, they do vary greatly in their physicochemical characteristics and include agricultural soils with diverse textures. The soils originated from areas that had not been subjected to any direct application of pesticides in recent years. Three of the soils were collected in Canada with either a shovel or a backhoe, depending on the location and the amount of soil collected. Sampling depth depended upon the nature of the soil and the site itself. The fourth soil (LUFA 2.2) was obtained from LUFA Speyer (Germany).

Two of the four agronomic soils were sandy loam soils. The first sample of sandy loam soil (Sandy Loam 1), classified as an Orthic Dark Brown Chernozem, was collected in August 2014 from an undeveloped road allowance southeast of Vulcan, Alberta. The soil beneath the sod was collected to a depth of approximately 30 cm and placed into 20-L plastic pails after all large rocks and aggregates had been removed. The soil was then shipped to Environment and Climate Change Canada (Ottawa, Ontario) where it was kept between ~10°C and 25°C until needed. The second sample of sandy loam soil (Sandy Loam 2) was collected from St. Zotique, Quebec, in collaboration with McGill University; it is likely characterized as a Gleysol. The soil was collected to

a depth of 20–30 cm, air-dried, sieved (4-mm), and stored between ~10°C and 25°C until required. The physicochemical characteristics of the soils are presented in Table D-1.

The third agronomic soil was a sample of clay loam soil. It was collected in 2014 from a fallow agricultural field in Winchester, Ontario; the field had lain fallow for at least two years but had been tilled prior to collection. The tilled soil was collected in 20-L pails and shipped to Environment and Climate Change Canada (Ottawa, ON). Due to its clay-rich composition, clumps of the air-dried soil were passed through a soil grinder in order to facilitate processing. Thereafter, the soil was sieved (4-mm mesh) and stored between ~10°C and 25°C until needed. The physicochemical characteristics of the soil are presented in Table D-1.

The fourth agronomic soil was a loam sand (LUFA 2.2) from Hanhofen, Rheinland-Pflaz, Germany (purchased via LUFA Speyer, Germany). The soil is free from pesticides, biocidal fertilizers, or organic manure for at least five years prior to collection. The soil was sampled from a 0–20 cm depth, and sieved with a 2-mm mesh screen. The specific soil batch was purchased in 2017 and shipped to Environment and Climate Change Canada (Ottawa, Ontario), where it was stored at ~23°C until needed. The physicochemical characteristics of the soil are presented in Table D-1.

Table D-1. Physicochemical characteristics of candidate artificial and natural negative control soils[1]

Parameter	Artificial Soil	Sandy Soil - Artificial	Sandy Loam 1	Sandy Loam 2	Clay Loam	LUFA	Analytical Method
Source	formulated from constituents	compost / red sand (80:20)	field-collected from Alberta	field-collected from Quebec	field-collected from Ontario	standard soil from Europe	–
Soil Texture	Fine Sandy Loam	Sandy Soil	Sandy Loam	Sandy Loam	Clay Loam	Loamy Sand	as per Hausenbuiller (1985); based on grain size distribution
Sand (%)	77.3	90	77.0	60.5	39.8	77	gravimetric grain size distribution
Silt (%)	7.8	4	12.0	25.5	28.3	17	gravimetric grain size distribution
Clay (%)	14.9	6	11.0	14.0	31.9	5.9	gravimetric grain size distribution
Water-holding Capacity (%)	71.5	65	46.3	61.9	76.8	47.9	gravimetric analysis[3]
pH (units)	6	8.2	6.1	6.7	6.8	5.0	0.01 *M* $CaCl_2$ method
Total Carbon (%)	4.5	–[2]	1.5	–	–		Leco furnace method
Organic Carbon (%)	–	5.0	–	3.9	–	16000 (mg/kg)	Leco furnace Method
Organic Matter (%)	9	8.4	2.6	4.4	15.2	4	dichromate oxidation
Cation Exchange Capacity (Cmol+/kg)	19	30	11	14	34	< 10	barium chloride method
Total Nitrogen (%)	0.05	0.32	0.13	0.25	–	–	Kjeldahl method
NH_4-N (mg/kg)	–	< 1	< 1	< 1[4]	–	–	Kjeldahl method
NO_3-N (mg/kg)	–	30	< 1	3	–	78	Kjeldahl method
NO_2-N (mg/kg)	–	< 1.0	< 1.0	< 1.0	–	–	Kjeldahl method
Phosphorus (mg/kg)	23	59	17	89	18	270	nitric/perchloric acid digestion

Parameter	Artificial Soil	Sandy Soil - Artificial	Sandy Loam 1	Sandy Loam 2	Clay Loam	LUFA	Analytical Method
Potassium (mg/kg)	22	1693	395	310	140	360	NH_4 acetate extraction, colourimetric analysis
Magnesium (mg/kg)	149	567	205	220	707	560	NH_4 acetate extraction, colourimetric analysis
Calcium (mg/kg)	1848	4133	1300	2300	5467	1600	NH_4 acetate extraction, colourimetric analysis
Sodium (mg/kg)	67	93	40	30	67	–	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 method and associated criteria for test validity described herein (EC, 2010, 2013b, 2014b; Hennessy, 2010; Princz *et al*., 2010, 2012, 2018; Princz, 2014; Ritchie *et al*., 2017; ECCC, 2018, 2019).
[2] Not determined.
[3] Determined according to USEPA (1989) using a Fisherbrand™ P8 creped filter paper (see Section 5.3).
[4] NH_3-N (mg/kg).

Five types of negative control soils from boreal ecozones within Canada were also used to develop the biological test method described herein and to further assess the method's robustness 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 natural boreal forest soils included one from Newfoundland, one from Ontario, two from Saskatchewan, and one from Alberta. The physicochemical characteristics of the formulated artificial soil and the five (including horizons) forest soils are summarized in Table D-2.

The five natural soils used as negative control soil while developing this biological test method and establishing the test validity criteria (see Section 4.4) do not represent all Canadian soil types. However, they do vary greatly in their physicochemical characteristics and include boreal ecozone soils with diverse textures (see Table D-2). The soils originated from areas that had not been subjected to any direct application of pesticides in recent years. *Bulk soils* were collected as separate horizons, where possible. Sampling depth depended on the nature of the soil and the site itself. Once collected, all soil horizons were air-dried, sieved (4 to 8 mm), homogenized, and stored at room temperature (23°C) until required.

The Newfoundland soil (NL Podzol) was classified as a Gleyed Humo-ferric Podzol, developed on a stony, loamy-to-sandy, non-calcareous glacial till (EcoDynamics Consulting Inc., 2011b). The main canopy within the site was dominated by balsam fir and scattered black spruce. The understory consisted of sheep laurel (*Kalmia angustifolia*) and creeping snowberry (*Gaulteria hispidula*), regenerating trees, bunchberry (*Cornus canadensis*), with lesser amounts of spinulose woodfern (*Dryopteris spinulosa*), cinnamon fern (*Osmunda cinnamomea*), two-leaved solomonseal (*Maianthemum canadense*), and blue bead lily (*Clintonia borealis*). The ground surface was dominated by feathermosses (e.g., Shreber's moss [*Pleurozium schreberi*], stair-step moss [*Hylocomium splendens*], and knight's plume [*Ptilium crista-castrensis*]). Prior to sampling, woody debris and leaf litter were removed, and the underlying organic F and H horizons were collected together, followed by the separate collection of the Ahe (to a depth of 3 cm), Ae (to a depth of 25 cm), and Bf horizons. All four horizons were used in the establishment of test validity criteria for this method.

The Ontario soil (ON Podzol) was classified as a Gleyed Humo-ferric Podzol developed within a non-calcareous fluvial-lacustrine deposit (EcoDynamics Consulting Inc., 2011a). The site was a coniferous-dominant mixed-wood forest, with a mixture of both coniferous and deciduous species. The upper canopy consisted mainly of red pine (*Pinus resinosa*) and eastern white pine (*Pinus strobus*), with scattered sugar maple (*Acer saccharum*), and with a lower canopy consisting of a mixture of white birch (*Betula papyrifera*), eastern white cedar (*Thuja occidentalis*), black spruce (*Picea mariana*), white spruce (*Picea glauca*), red maple (*Acer rubra*), and eastern hemlock (*Tsuga canadensis*). The understory was dominated by regenerating tree species, with lesser amounts of speckled alder (*Alnus incana*), beaked hazelnut (*Corylus cornuta*), eastern leatherwood (*Dirca palustris*), wild raisin (*Viburnum nudum*), velvet blueberry (*Vaccinium myrtilloides*), and twinflower (*Linnaea borealis*). The ground surface was dominated by bunchberry (*Cornus canadensis*) and goldthread (*Coptis trifolia*). Three horizons were collected following the removal of the forest litter: the Ahe (to a depth of 2 cm), Ae (to a depth of 7 cm), and Bf horizons (to a depth of 20 cm). Only the Ahe horizon was used in the establishment of test validity criteria for this method.

The Alberta soil (AB01 Gleysol) was collected from a bog and consisted of a poorly drained Rego Humic Gleysol (Peaty Phase), with soil texture varying from loam to clay loam near the surface, and becoming clay-rich with depth (EcoDynamics Consulting Inc., 2007). The site was dominated by black spruce (*Picea mariana*), with an understory dominated by peat mosses (*Sphagnum* spp.), and haircap mosses (*Polytrichum* spp.). Two horizons were collected: a mixture of Of/Oh horizons, and the Ahg horizon (to a depth of 17 cm); however, only the Of/Oh layer was used in the establishment of test validity criteria for this method.

Two soils were collected from Saskatchewan. The first soil (SK01 Luvisol) was classified as a well-drained to moderately well-drained Dark Grey Luvisol, developed on stone-free, loamy-to-clayey glaciolacustrine materials (EcoDynamics Consulting Inc., 2007). The forest cover was a mixture of white spruce (*Picea glauca*) and

trembling aspen (*Populus tremuloides*), with an understory of aspen suckers, rose (*Rosa* sp.), willow (*Salix* spp.), bunchberry (*Cornus canadensis*), and twinflower (*Linnaea borealis*). Three horizons were collected: LFH (10 cm depth), Ahe (10 cm depth), and Bt (to a depth of 19 cm), but only the LFH horizon was used in the establishment of test validity criteria for this method.

The second soil collected from Saskatchewan (SK02 Brunisol) was classified as a rapidly drained Orthic Eutric Brunisol, developed in a stone-free, sandy glaciofluvial material (EcoDynamics Consulting Inc., 2007). The forest cover consisted of pure jack pine (*Pinus banksiana*), with an understorey dominated by aspen (*Populus tremuloides*), green alder (*Alnus crispa*), bearberry (*Arctostaphylos uva-ursi*), and reindeer lichens (*Cladina* spp.). The leaf litter was removed, and the FH was collected to a depth of approximately 6 cm; the Ah and Bm horizons were collected together to a depth of approximately 25 to 30 cm, as the Ah was discontinuous and thin (2 cm). Both horizons were used in the establishment of test validity criteria for this method.

Table D-2 Physicochemical characteristics of candidate artificial and natural negative control boreal soils and soil horizons[1]

			Soil type:	Artificial Soil	NFLD01 Podzol			
			Source:	In-house	Newfoundland			
			Soil classification:	n/a	Gleyed Humo-ferric Podzol			
			Horizon:	n/a	LFH	Ahe	Ae	Bf
Parameter	**Units**	**Analytical method**						
Soil texture[2]		n/a[3]		SL	-[4]	-	-	-
Sand	%	Particle size distribution (filter candle system)		76	64	73	48	72
Silt	%			12	< 1	16	44	20
Clay	%			12	36	11	8	8
Water-holding capacity	%	EC (2005a)		79.0	275.0	108.5	48.2	41.9
Optimal moisture content	%			62.5	92.5	70.0	50.0	55.0
pH	units	1:1 water method		7.4	3.9	3.6	3.7	4.2
Electrical conductivity	mS/cm	Saturated paste method		-	-	-	-	-
Organic carbon	%	Leco furnace method		5.5	-	-	-	-
Organic matter	%	Loss on ignition		4.6	82.6	26.7	2.9	4.6
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method		11	32	33	21	
Total nitrogen	%	Kjeldahl method		0.07	-	-	-	-
NH_3	mg/kg	2N KCL extractable		3	744	278	14	15
NO_3-N	mg/kg			5	< 10	< 10	< 10	< 10
NO_2-N	mg/kg			< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
Phosphorus (total)	%			0.03	0.08	0.04	0.04	0.04
Phosphorus	mg/kg	$NaHCO_3$ extractable		9	20	17	8	4
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis		11	160	90	20	20
Magnesium	mg/kg			77	110	90	20	20
Calcium	mg/kg			2000	400	300	100	< 100
Sodium	mg/kg			44	20	20	10	10
C/N				34	-	-	-	-
Sodium adsorption ratio		Saturated paste method		0.3	-	-	-	-

		Soil type:	**ON Podzol**			**AB01 Gleysol**	
		Source:	Ontario			Alberta	
		Soil classification:	Gleyed Humo-ferric Podzol			Rego Humic Gleysol	
		Horizon:	A	Ae	B	Of/Oh	Ahg
Parameter	**Units**	**Analytical method**					
Soil texture[2]		n/a[3]	LS	LS	LS	Peat	SL
Sand	%	Particle size distribution (filter candle system)	82	88	86	n/a	59
Silt	%		12	6	6	n/a	33
Clay	%		6	6	8	n/a	8
Water-holding capacity	%	EC (2005a)	41.0	181.9	40.9	248.1	73.9
Optimal moisture content	%		65.0	52.5	47.5	100.0	70.0
pH	units	1:1 water method	4.6	4.6	5.8	3.9	4.3
Electrical conductivity	mS/cm	Saturated paste method	-	-	-	0.38	0.1
Organic carbon	%	Leco furnace method	32.1	1.6	1.0	34.6	11.3
Organic matter	%	Loss on ignition	58.1	2.1	2.2	67.8	21.5
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method	26	9	12	27	39
Total nitrogen	%	Kjeldahl method	0.96	0.06	0.05	2	0.63
NH_3	mg/kg	2N KCL extractable	128	4	2	114	9
NO_3-N	mg/kg		< 1	< 1	< 1	3	9
NO_2-N	mg/kg		< 1	< 1	< 1	< 1	< 1
Phosphorus (total)	%		-	-	-	-	-
Phosphorus	mg/kg	$NaHCO_3$ extractable	16	2	< 2	28	33
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis	143	23	16	53	81
Magnesium	mg/kg		151	31	40	66	108
Calcium	mg/kg		765	184	191	462	570
Sodium	mg/kg		57	35	21	57	28
C/N			33.4	26	20.6	17.3	-
Sodium adsorption ratio		Saturated paste method	2.0	2.8	2.4	0.9	1.3

		Soil type:	**SK01 Luvisol**			**SK02 Brunisol**	
		Source:	Saskatchewan			Saskatchewan	
		Soil classification:	Dark Grey Luvisol			Orthic Eutric Brunisol	
		Horizon:	LFH	Ahe	Bt	FH	AB
Parameter	**Units**	**Analytical method**					
Soil texture[2]		n/a[3]	SL	L	L	SL	LS
Sand	%	Particle size distribution (filter candle system)	68	37	35	89	82
Silt	%		22	53	55	7	12
Clay	%		10	10	10	6	4
Water-holding capacity	%	EC (2005a)	287.7	68.6	42.1	174.1	39.5
Optimal moisture content	%		55.0	52.5	42.5	55.0	45.0
pH	units	1:1 water method	6.6	6.4	6.6	6.9	6.8
Electrical conductivity	mS/cm	Saturated paste method	-	-	-	-	-
Organic carbon	%	Leco furnace method	29.4	4.9	1.0	11.4	1.0
Organic matter	%	Loss on ignition	46.7	9.5	2.0	15.8	1.8
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method	43	22	11	22	6
Total nitrogen	%	Kjeldahl method	1.6	0.41	0.07	0.65	0.05
NH_3	mg/kg	2N KCL extractable	158	49	5	23	6
NO_3-N	mg/kg		15	7	3	86	< 1
NO_2-N	mg/kg		< 1	< 1	< 1	< 1	< 1
Phosphorus (total)	%		0.18	0.14	0.06	0.05	0.02
Phosphorus	mg/kg	$NaHCO_3$ extractable	56	62	9	24	16
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis	411	363	170	200	83
Magnesium	mg/kg		586	315	198	785	196
Calcium	mg/kg		7260	3540	1780	2860	795
Sodium	mg/kg		93	100	67	64	50
C/N			20.5	0.8	0.3	4	0.6
Sodium adsorption ratio		Saturated paste method	0.0	0.1	0.2	0.4	0.1

[1] Characteristics of the artificial and various negative control soils that have been used to develop the definitive biological test method and associated criteria for test validity described herein (EC, 2010, 2013b, 2014b; Hennessy, 2010; Princz *et al*., 2010, 2012, 2018; Princz, 2014; Ritchie *et al*., 2017; ECCC, 2018, 2019).

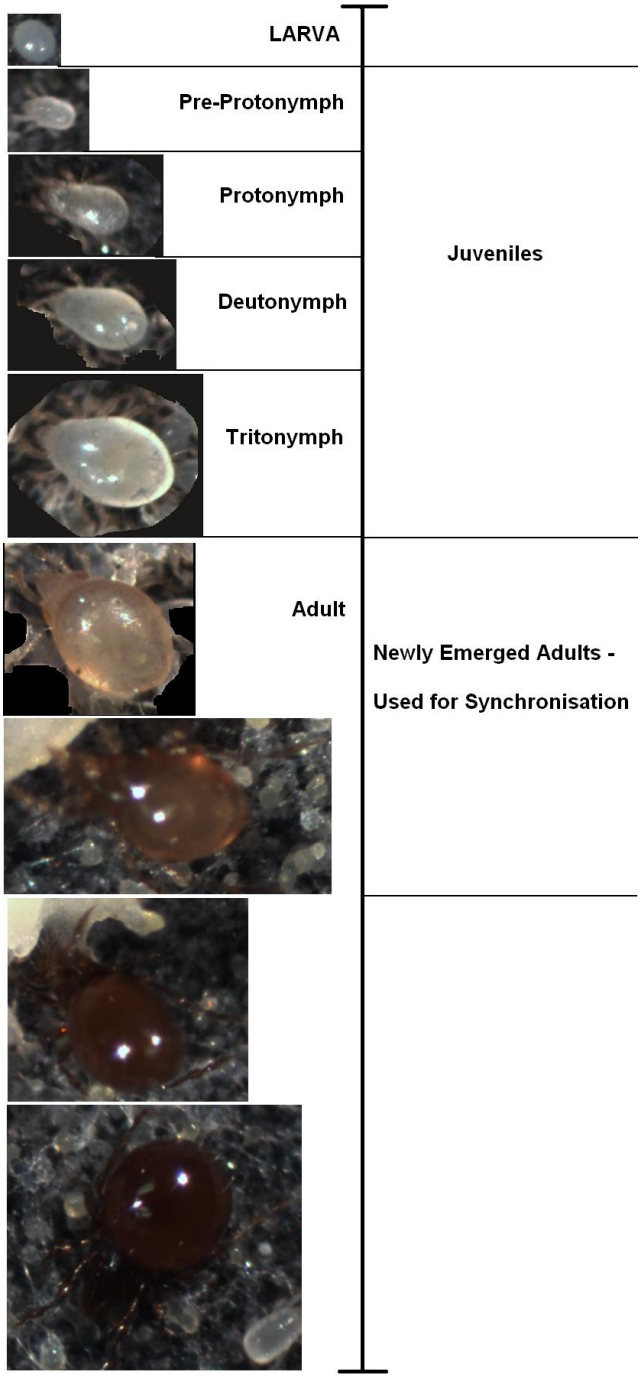
[2] SL = sandy loam; LS = loam sand; L = loam.

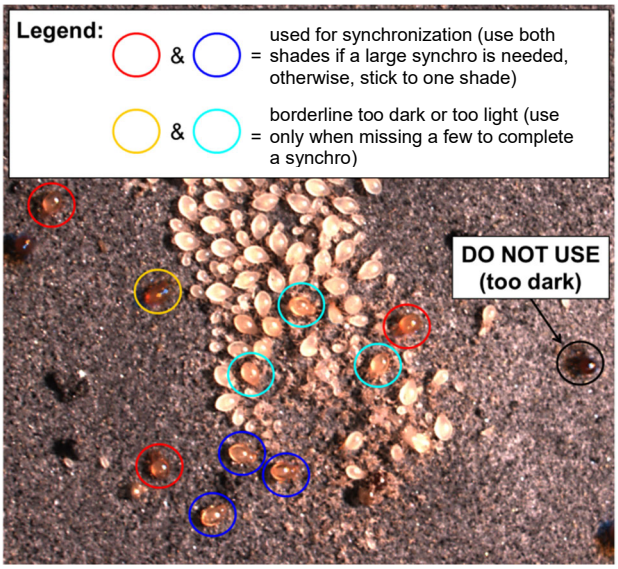
[3] Not applicable.

[4] Not determined.

Appendix E

Illustrative Photographs of *Oppia nitens*





Shades of *Oppia nitens* appropriate for use in age-synchronized testing.

Mites with an opaque "milky" appearance must not be used in testing.

Appendix F

Logarithmic Series of Concentrations Suitable for Toxicity Tests[a]

Column (Number of concentrations between 10.0 and 1.00, or between 1.00 and 0.10)[b]

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

[a] Modified from Rocchini *et al.* (1982).

[b] A series of successive concentrations may be chosen from a column. Midpoints between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed as a percentage by 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 2, which spans two orders of magnitude in concentration, 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. The finer gradations of columns 4 to 7 might occasionally be useful for testing chemicals that have an abrupt threshold of effect.

Appendix G

Heat-Extraction Procedures

Further details on the method described herein for the heat-extraction of *O. nitens* from test soils (Section 4.7) are provided in this appendix. Alternative methods for the heat-extraction of *O. nitens* from test soils are also provided in this appendix. These alternative procedures were provided by ECT Oekotoxikologie GmbH (Flörsheim am Main, Germany) and Eurofins Agroscience Services EcoTox GmbH (Niefern-Öschelbronn, Germany). Other heat-extraction procedures or variations of those provided herein may also be used, provided the extraction efficiency of the chosen method is validated (Section 4.7).

Environment and Climate Change Canada Procedure (see Section 4.7)

All equipment, apparatus, and construction materials must be made of non-toxic material, and the use of toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must be avoided. The facility should be well ventilated and free of fumes as well as isolated from any contaminants that might affect the test organisms and isolated from areas for sample preparation and storage.

Equipment and Reagents (Figure G-1)

- Personal protective equipment (e.g., lab coat, gloves, safety glasses)
- Plastic cups (e.g., Fisher cat #11-838-17)
- Round plastic needlework canvas (i.e. 3" diameter, #7 mesh)
- Parafilm®
- Scissors and/or X-acto knife (or similar cutting tool)
- Hot glue gun and non-toxic glue sticks
- Lamp (with 60-watt bulb)
- Ruler
- Thermometer(s)
- Plastic spray bottle
- Cheesecloth
- Deionized or distilled water
- Ethanol (70% v.v.)

Figure G-1 Materials needed to construct a heat-extraction vessel; polypropylene 4.5-oz cups (2 per vessel), plastic mesh circle 14.6 cm, glue gun, cheesecloth, and Parafilm®.

Creation of the Heat-Extraction Units

1. Cut the bottom off of a plastic cup (approximately 1 cm from the bottom) to create the top portion of the heat-extraction unit (Figure G-2A).
2. Trim a piece of plastic canvas so that it fits snugly into the ridge near the uncut end of the plastic cup (Figure G-2B). Secure the mesh in place with hot glue (Figure G-2C). Allow the glue to dry before assembling the extraction units.
3. To create the bottom portion of the heat-extraction unit, prepare a layer of plaster of Paris substrate in uncut plastic cups (about 1–2 cm depth).
4. Assemble the heat-extraction unit by inverting a top portion and placing it on top of a bottom portion so the mouths of the cups meet. Seal the unit by wrapping a strip of Parafilm® tightly around the seam between the two cups and secure with tape if necessary (Figure G-2D).
5. Place a small square of cheesecloth (single layer) onto the mesh so that no soil can drop through to the bottom portion (Figure G-2D).

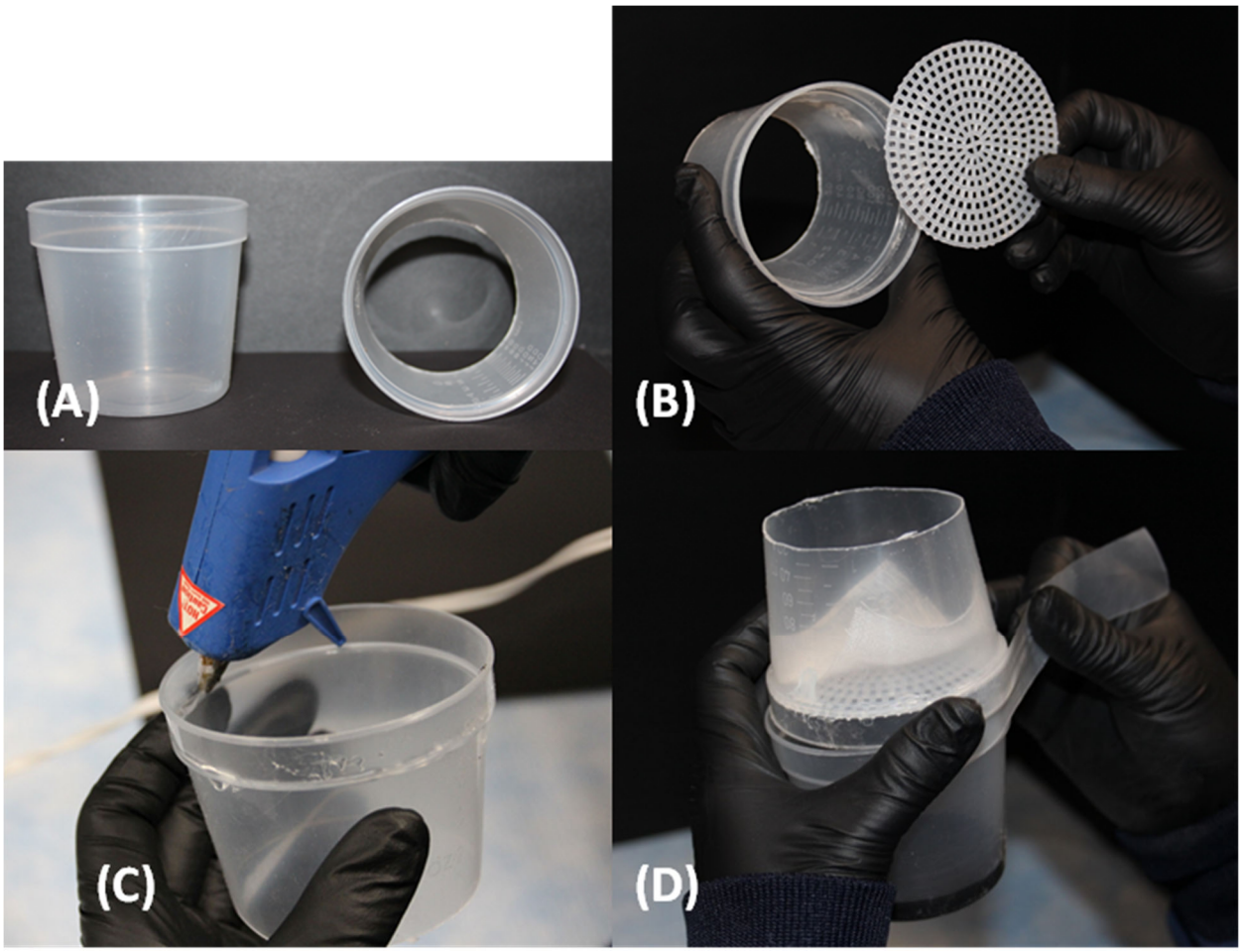


Figure G-2 Creation of the heat-extraction vessels. (A) The bottom is cut out from one of the plastic cups (right-hand photo). (B) The plastic mesh is cut to fit inside the plastic cup lip of the lid (~7 cm). (C) Hot glue is used to affix the mesh to the lip of the lid. (D) The bottom cup (filled with 1–2 cm of plaster of Paris substrate) is attached to the top of the unit (containing the cut bottom) with Parafilm® and tap as needed.

Performing the Heat-Extraction

The heat-extraction should be initiated on the last test day (i.e., Day 28). One heat-extraction unit should be prepared and used for each test vessel being heat-extracted.

1. Remove the lid from the test vessel. Transfer the soil from the test vessel to the corresponding heat-extraction unit by inverting the vessel onto the extraction unit (Figure G-3A). Take care to ensure all soil has been dislodged and transferred from the test vessel. Gently smooth out the soil evenly over the plastic mesh (lined with cheesecloth) using a spoon or scoopula.
2. Record the number of live and dead adults and juveniles (i.e., progeny) present in the test vessel that were not transferred to the heat-extraction unit. This can be done by adding deionized water to the empty test vessel; this will cause any remaining organisms to float, facilitating counting.
3. Place the heat-extraction units underneath a lamp fitted with a 60-watt bulb. Each lamp should have no more than 5–6 units beneath it, to keep heat and light consistent for each heat-extraction unit (Figure G-3B).
4. Place a thermometer probe directly into the soil of one of the heat-extraction vessels; one thermometer should be used per lamp. Record the time at which the heat-extraction began (i.e., when the bulbs were turned on).
5. Adjust the height of the lamp so that the bottom of the bulb is approximately 25 cm above the surface of the soil. As the extraction progresses, the lamp height does not need any adjustment, and the temperature should reach ~32°C after 48 hours.
6. Monitor and record the temperature within the soil throughout the extraction twice per day, or more frequently if necessary. Lamps may be raised during periods when technicians aren't present to monitor the units (e.g., overnight) to prevent the soil from drying completely.
7. Ensure the soil does not completely desiccate by wetting the surface with deionized water using a spray bottle as required. The soil should not be wetted 16 hours prior to the end of the extraction.
8. End the heat-extraction after 48 hours by turning off the lamps.
9. Count the organisms in the extraction vessel (bottom half of the cup) using flotation. If the organisms cannot be counted right away, the substrate can be moistened with deionized water and placed in cool storage (e.g., 4°C–10°C) until processing can continue, or the organisms can be preserved in 70% alcohol for enumeration at a later date. The flotation method described in EC 2014a can be used to enumerate the organisms. Deionized water (e.g., 1–2 cm depth) is added to the bottom half of the heat-extraction vessel containing the extracted organisms and the contents are swirled or stirred lightly with a brush to dislodge any organisms adhering to the plaster of Paris substrate or the sides of the container. The mites, both adult and progeny, float to the surface, and the cup is placed under a dissecting microscope for enumeration.
10. The heat-extraction vessels can be washed with soap and water, rinsed with deionized water (three times), and reused; however, the plaster of Paris substrate must be discarded and replaced.

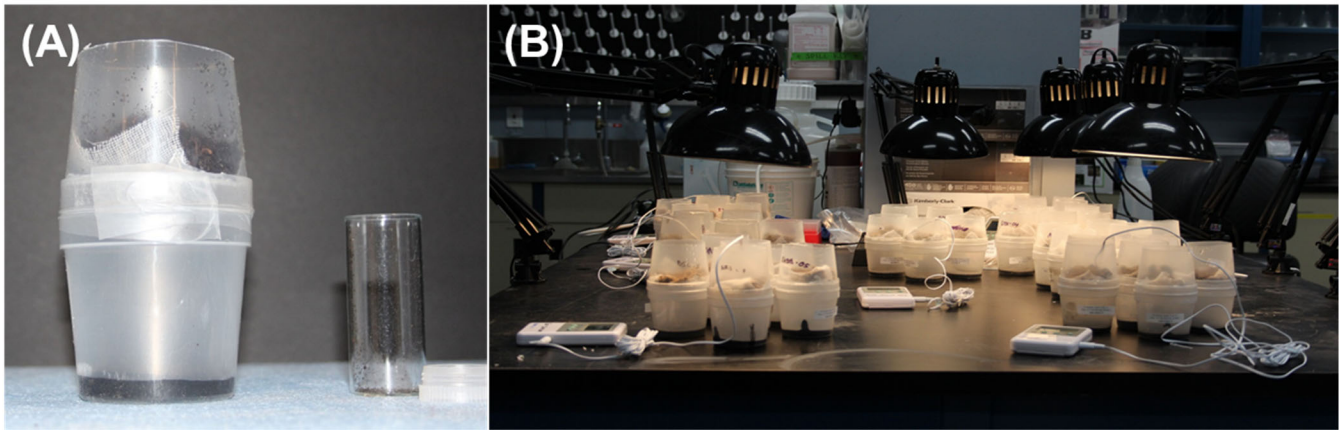


Figure G-3 (A) Soil is transferred from the test vessel to the heat-extraction unit. (B) The heat-extraction units are placed underneath the lamps for 48 hours.

ECT Oekotoxikologie GmbH Procedure

This method is a variation of that described above and in Section 4.7. Variations include:

- 25-watt bulbs are used and started at a height of approximately 30 cm above the surface of the soil (instead of 25 cm).
- On day 0 of the heat-extraction, the temperature should reach ~26°C.
- On day 1, the distance of the lamp to the soil is decreased to increase the temperature up to ~28°C.
- To achieve 32°C on the last day, the height of the lamp is adjusted to ~20 cm above the soil.

Eurofins Agroscience Services Ecotox GmbH Procedure

This procedure uses a MacFayden Extractor that generates a high light and heat gradient to slowly force the organisms out of a test soil through a mesh screen, dropping onto a collection vessel below.

Equipment and Reagents

- MacFayden Extractor (including plastic crucibles with metal mesh at the bottom)
- Thin gauze cut-outs (for covering the metal mesh)
- Plastic petri dishes (for covering the Units)
- Labels
- Plastic spray bottle with water
- Plastic spray bottle with Ethanol 70%
- Collecting vessels
- Ethanol (70% v.v.)

Description of the Extractor

The objective is to create a gradient of light, temperature, and (later) moisture. The apparatus consists of two chambers. In the upper chamber, a heat and light source is placed over the soil sample. In the lower chamber, cool air is provided by a cooling device.

The soil sample is placed in a (labelled) plastic crucible containing a mesh and thin gauze cut-out over a glass funnel. A collecting vessel is fixed below the funnel, which contains ethanol (70% v.v.) for preserving the animals. Furthermore, the soil sample is covered with a petri dish to prevent the soil from drying out.

Arthropods react to the heat and desiccation by moving downward (away from the heat) and fall through the screen at the bottom through the funnel into the collecting jar attached below the container.

Extraction Process

The extraction is carried out with stepwise daily temperature increases from test temperature (20 ± 2°C) to 32°C. Throughout the extraction process, the soil is moistened two to three times per day as needed with a spray bottle. The soil is not moistened for the final 24 hours of the heat-extraction.

Time (hours)	Temperature (°C)
0 (Start)	22 (Start)
8	24
24	26
48	28
72	30
96	32
120 (End)	32 (End)

An example of the duration and temperature gradient for the heat-extraction is provided in the table above. The duration of extraction and temperature gradient are guidelines only and should be chosen to optimize the heat-extraction efficiency. The temperature conditions can be recorded by a data logger.

Counting Test Organisms

After the heat-extraction is complete, the plastic crucibles are removed, and the funnels are rinsed two times with Ethanol 70% to ensure that all of the organisms are captured in the collection vessels (i.e., not remaining in the funnel). The rims of the collecting vessels are also rinsed with ethanol in order to transfer all organisms that might be stuck on the rims into the ethanol collection vessels (they might otherwise desiccate and become difficult to detect while counting). Organisms captured in the collection vessels are then counted and the results recorded.

Appendix H

Determining a Positive Control Concentration and Defining Warning Limits – Worked Example

1. Use a minimum of five valid (i.e., test validity criteria must be met) multi-concentration tests in one soil type (e.g., negative control soil or artificial soil), with the same reference toxicant. In this example, tests were conducted in a clean field soil (i.e., negative control soil) using boric acid as the reference toxicant.
2. For each test, tabulate the mean total number of progeny produced per treatment (Table H-1).
3. For each test, calculate and tabulate the percent reduction of progeny production relative to the control response (Table H-2) using the following formula:

$$\%\ Reduction = \left(\frac{(Control\ Response\ -\ Treatment\ Response)}{Control\ Response}\right) \times 100$$

4. Calculate the mean percent reduction of progeny production for each treatment (Table H-2). *Optional*: plot the data (Figure H-1).
5. Select a concentration where the data tend to be less variable (i.e., range of the data spans ~20%), but still show a partial effect (i.e., 30% to 70% reduction; see shaded cells in Table H-2 or circled data in Figure H-1).
6. Calculate the standard deviation (SD) and two standard deviations (2 SD) of the mean percent reduction for the selected test concentration (Table H-2).
7. Calculate the mean percent reduction ± 2 SD for the selected test concentration (Table H-2) and compare these values to the minimum and maximum percent reduction observed within that treatment to ensure that the proposed warning limits (i.e., mean percent reduction ± 2 SD) capture the response data. Use the mean percent reduction at that treatment to define the target effect level.
8. In this example, 95 mg H_3BO_3/kg dry soil produced a 41% mean percent reduction of progeny production (i.e., target effect level) with proposed warning limits of ≥ 27% and ≤ 56%. Based on these results, this is the test concentration of boric acid that a laboratory might choose and then run concurrently with each definitive test for the positive control treatment.
9. For tests where the positive control is included as part of the definitive reproduction test, the percent reduction of progeny production (i.e., effect) is compared with the established warning limits. This is carried out and documented following the same procedures as those used for comparing multi-concentration reference toxicity tests in reference toxicant warning charts (Section 4.9 and Figure H-2). If the percent reduction of progeny in a positive control run with a definitive test is within the established warning limits (i.e., mean % reduction ± 2 SD), the positive control is acceptable. If the response is outside of those limits, an investigation into the test conduct and sensitivity of the test population (i.e., in-house cultures) must be launched (see Section 4.9). This investigation might include, for example, determining if the positive control concentration was prepared properly, checking test calculations, confirming the positive control concentration analytically, investigating the negative control data, examining culture health data, investigating technician proficiency, or soil age quality (e.g., stored too long in buckets). In addition to maintaining warning charts of positive control data, a laboratory should monitor the variability of the positive control response over time by calculating the coefficient of variation (CV) of the response and evaluating it relative to a pre-defined acceptability limit (e.g., lab defines ≤ 30% CV as acceptable). In this example, the CV is 17.2% for six data points (Table H-2).

Table H-1 Mean number of *Oppia nitens* progeny produced upon exposure to boric acid in clean field soil

	Boric Acid Concentration (mg/kg)					
Test No.	**0**	**56**	**73**	**95**	**123**	**160**
1	175.4	81.0	207.8	122.4	33.8	10.2
2	38.0	39.4	46.0	21.0	9.8	0.2
3	173.2	180.0	173.6	109.6	53.0	9.8
4	114.4	117.6	120.4	57.2	33.6	1.2
5	139.5	129.8	117.4	75.6	29.0	6.2
6	124.6	117.2	85.8	75.4	46.2	2.6

Table H-2 Percent reduction in *Oppia nitens* progeny production, relative to the control response, upon exposure to boric acid in clean field soil

	Boric Acid Concentration (mg/kg)					
Test No.	**0**	**56**	**73**	**95**	**123**	**160**
1	0	53.8	-18.5	30.2	80.7	94.2
2	0	-3.7	-21.1	44.7	74.2	99.5
3	0	-3.9	-0.2	36.7	69.4	94.3
4	0	-2.8	-5.2	50.0	70.6	99.0
5	0	7.0	15.8	45.8	79.2	95.6
6	0	5.9	31.1	39.5	62.9	97.9
Mean	**-**	**9.4**	**0.3**	**41.2**	**72.9**	**96.7**
SD[a]				**7.1**		
2 SD				**14.3**		
Mean + 2 SD				**55.5**		
Mean – 2 SD				**26.9**		
%CV				**17.2**		

[a] Standard deviation

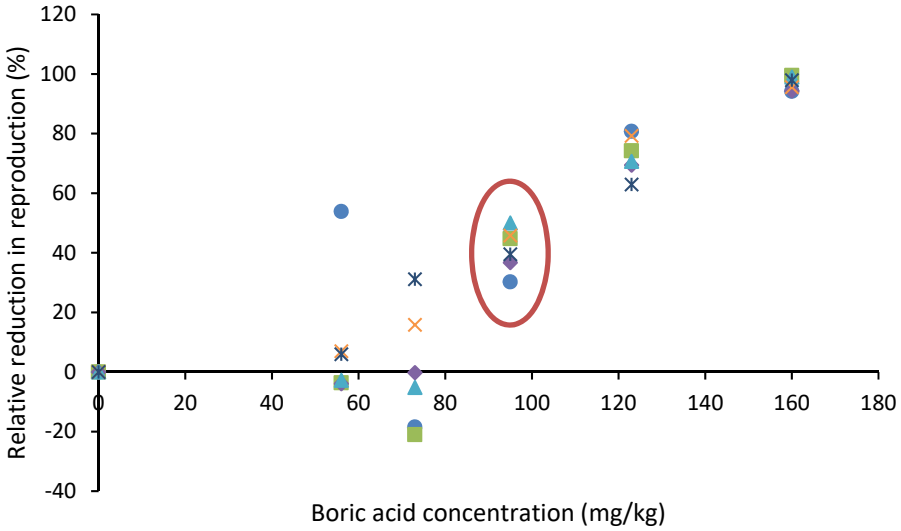


Figure H-1 ***Oppia nitens*** **reproduction (as a percentage of the control response) in clean field (i.e., negative control) soil spiked with boric acid (mg/kg). Each symbol represents different multi-concentration tests.**

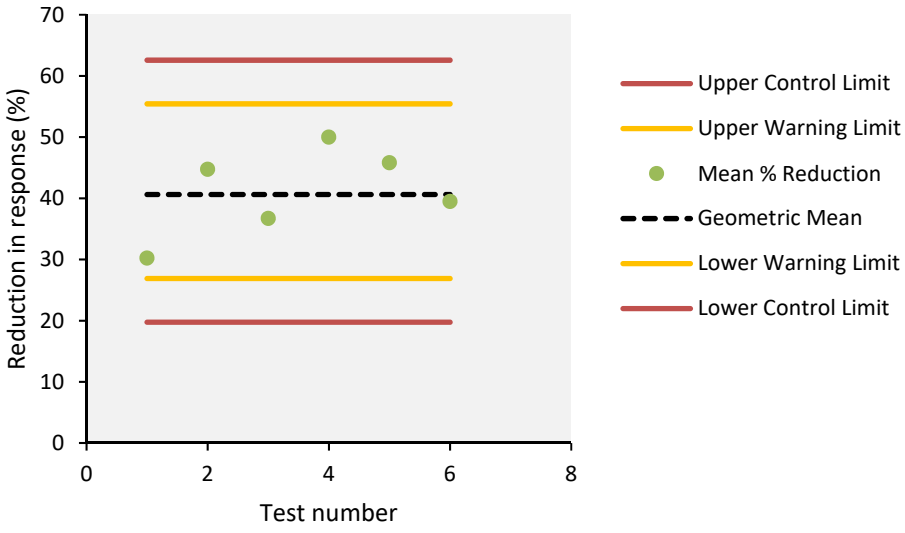


Figure H-2 **Warning chart depicting the percent reduction in** ***Oppia nitens*** **progeny production (as a percent of the control response) in positive control treatments (95 mg boric acid/kg soil) included with each definitive reproduction test. The lower and upper warning limits represent the mean percent reduction (calculated with each new test) ± 2 SD. The lower and upper control limits represent the mean percent reduction ± 3 SD.**