Environment Canada Environnement Canada

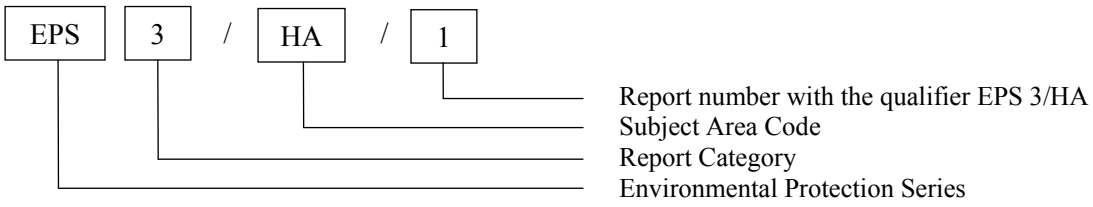
Biological Test Method: Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region

EPS 1/RM/56 – August 2013
Science and Technology Branch
Environment Canada

Canada

Environmental Protection Series

Sample Number:



Categories

1 Regulations/Guidelines/Codes of Practice
2 Problem Assessments and Control Options
3 Research and Technology Development
4 Literature Reviews
5 Surveys
6 Social, Economic and Environmental Impact Assessments
7 Surveillance
8 Policy Proposals and Statements
9 Manuals

Subject Areas

AG	Agriculture
AN	Anaerobic Technology
AP	Airborne Pollutants
AT	Aquatic Toxicity
CC	Commercial Chemicals
CE	Consumers and the Environment
CI	Chemical Industries
FA	Federal Activities
FP	Food Processing
HA	Hazardous Wastes
IC	Inorganic Chemicals
MA	Marine Pollutants
MM	Mining and Ore Processing
NR	Northern and Rural Regions
PF	Paper and Fibres
PG	Power Generation
PN	Petroleum and Natural Gas
RA	Refrigeration and Air Conditioning
RM	Reference Methods
SF	Surface Finishing
SP	Oil and Chemical Spills
SRM	Standard Reference Methods
TS	Transportation Systems
TX	Textiles
UP	Urban Pollution
WP	Wood Protection/Preservation

New subject areas and codes are introduced as they become necessary. A list of EPS reports may be obtained from Publications Catalogue, Environment Canada, Ottawa, Ontario, K1A 0H3, Canada.

Biological Test Method: Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region

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

Report EPS 1/RM/56
August 2013

Paper:
ISBN 978-1-100-22420-6
Cat. No.: En49-7/1-56E

PDF:
ISBN 978-1-100-22421-3
Cat. No.: En49-7/1-56E-PDF

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 Canada
335 River Road
Ottawa, Ontario
K1A 0H3

Lisa Taylor, Manager
Method Development and Applications Unit
Science and Technology Branch
Environment Canada
335 River Road
Ottawa, Ontario
K1A 0H3

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

Catalogue des publications
Environnement Canada
Ottawa (Ontario)
K1A 0H3

Review Notice

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

Abstract

This document provides detailed procedures, conditions and guidance for preparing for and conducting a biological test for measuring soil toxicity using terrestrial plants of the Canadian boreal region. Seven species options are provided and include trembling aspen (*Populus tremuloides*), bluejoint reedgrass (*Calamagrostis canadensis*), Canada goldenrod (*Solidago canadensis*), paper birch (*Betula papyrifera*), jack pine (*Pinus banksiana*), white spruce (*Picea glauca*), or black spruce (*Picea mariana*). The test is a 28-, 35- or 42-day test for effects on plant growth (measured as shoot and root length and shoot and root dry mass). The method is conducted as a static (i.e., no renewal) test, using one or more samples of contaminated or potentially contaminated soil or one or more concentrations of chemical(s) or chemical product(s) spiked in negative control (or other) soil. Soil samples are collected as distinct horizons and reassembled in test vessels in proportions correlated to the depths of each horizon, as collected in the field. Water is added to the test vessels to hydrate soils as needed, for the duration of the test.

The test is conducted at a mean temperature of 24 ± 3°C or with cyclical temperatures of 24 ± 3°C during the day and 15 ± 3°C at night. Test vessels (1-L polypropylene) contain a measured wet-weight equivalent to a volume of ~500 mL of test soil. Five or 10 seeds (i.e., number of seeds per test vessel is species-specific) are placed into each replicate test vessel. This test uses ≥ 5 replicated test vessels/treatment for a single-concentration test, and 3 to 6 replicated test vessels/treatment for a multi-concentration test. The options for test design in a multi-concentration test include an equal number of replicates per treatment (i.e., ≥ 4) or unequal replicates per treatment (i.e., 6 per treatment for each negative and other control; 4 replicates for each of the lowest 4 to 6 test concentrations; and 3 replicates for each of the highest 5 test concentrations). Following a 28-, 35-, or 42-day exposure (i.e., test duration is species-specific), the shoot and root lengths and the shoot and root dry weights of individual plants in each replicate are determined and the treatment means compared.

General or universal conditions and procedures are outlined for test preparation and performance. Additional conditions and procedures are stipulated that are specific to the intended use of each test. The biological test method described herein is suitable for measuring and assessing the toxicity of samples of field-collected soil or similar particulate material, or of negative control (or other) soil spiked (mixed) in the laboratory with test chemical(s) or chemical product(s). Instructions and requirements are included on test facilities, sample collection, handling and storing samples, seed source and storage, seed stratification, and handling, preparing soil or spiked-soil mixtures and initiating tests, specific test conditions, appropriate observations and measurements, endpoints and methods of calculation, and the use of a reference toxicant.

Résumé

Le présent document renferme des indications précises et décrit en détail les conditions et modes opératoires applicables à la préparation et à l'exécution d'un essai biologique visant à mesurer la toxicité d'un sol pour des plantes terrestres de la région boréale du Canada. Les espèces pouvant être utilisées sont les suivantes : peuplier faux-tremble (*Populus tremuloides*), calamagrostide du Canada (*Calamagrostis canadensis*), verge d'or du Canada (*Solidago canadensis*), bouleau à papier (*Betula papyrifera*), pin gris (*Pinus banksiana*), épinette blanche (*Picea glauca*), épinette noire (*Picea mariana*). L'essai, d'une durée de 28, 35 ou 42 jours, permet de mesurer les effets sur la croissance des plantes (d'après la longueur et le poids sec des pousses et des racines). Il s'agit d'un essai en conditions statiques faisant appel à un ou des échantillons de sol contaminé ou susceptible d'être contaminé, ou encore à une ou des concentrations d'une substance chimique dont on enrichit un sol témoin négatif (ou autre). Les échantillons de sol sont prélevés par horizon. En laboratoire, au moment du transfert des échantillons dans les récipients d'essai, on reconstitue ces horizons dans les mêmes proportions (profondeur) que celles observées sur le terrain. On ajoute de l'eau dans les récipients d'essai afin d'hydrater le sol au besoin pendant la durée de l'essai.

L'essai se déroule à une température moyenne de 24 ± 3 °C ou à des températures cycliques de 24 ± 3 °C le jour et de 15 ± 3 °C la nuit. Chaque récipient d'essai (de 1 L, en polypropylène) renferme un volume mesuré (poids humide) de sol d'essai équivalant à ~500 mL. Chaque récipient de répétition compte 5 ou 10 graines (selon l'espèce). Dans le cas d'un essai à concentration unique, on utilise ≥5 récipients de répétition par traitement; s'il s'agit d'un essai à concentrations multiples, on en utilise entre 3 et 6. Les options présentées pour le plan d'expérience faisant appel à un essai à concentrations multiples incluent un nombre (≥4) égal ou inégal de répétitions par traitement (soit 6 par traitement pour le sol témoin négatif ou autre sol témoin; 4 pour chacune des 4 à 6 concentrations expérimentales les plus faibles; 3 pour chacune des 5 concentrations expérimentales les plus élevées). Après une exposition de 28, 35 ou 42 jours (selon l'espèce employée), on détermine tant la longueur que le poids sec des pousses et des racines de chaque plante dans chacune des répétitions, puis on compare les moyennes obtenues.

Le présent document décrit les conditions et modes opératoires généraux ou universels applicables à la préparation et à l'exécution de l'essai. Il renferme aussi une description des conditions et modes opératoires supplémentaires propres à l'usage prévu des résultats de chaque essai. La méthode d'essai biologique présentée ici convient à la mesure et à l'évaluation de la toxicité d'échantillons d'un sol ou d'une matière particulaire semblable prélevés sur le terrain, ou encore d'échantillons d'un sol témoin négatif (ou autre) enrichi en laboratoire avec une ou des substances chimiques d'essai. La méthode contient aussi des instructions et des exigences relatives aux éléments suivants : installations d'essai; prélèvement, manipulation et entreposage des échantillons; source, entreposage, stratification et manipulation des graines; préparation du sol ou des mélanges de sol enrichi; mise en route de l'essai; conditions propres à l'essai; observations et mesures pertinentes; paramètres et méthodes de calcul; utilisation d'un toxique de référence.

Foreword

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

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

The different types of tests included in this series were selected because of their acceptability for the needs of environmental protection and management programs carried out by Environment Canada. These reports are intended to provide guidance and to facilitate the use of consistent, appropriate and comprehensive procedures for obtaining data on the toxicity to terrestrial or aquatic life of 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 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. Italics are also used as emphasis for these and other words, throughout the report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Chemical Formulae

AES	atomic emission spectrophotometry
ANOVA	analysis of variance
$CaCl_2$	calcium chloride
$CaCO_3$	calcium carbonate
$Ca(OH)_2$	calcium hydroxide
CCME	Canadian Council of Ministers of the Environment
CEC	cation exchange capacity
cm	centimetre(s)
C	carbon
°C	degree(s) Celsius
CV	coefficient of variation
d	day(s)
DQO	data quality objectives
EC50	median effective concentration
ES	effect size
g	gram(s)
h	hour(s)
H_0	null hypothesis
H_2O	water
HCl	hydrochloric acid
HNO_3	nitric acid
HPLC	high pressure liquid chromatography
ICAP	inductively coupled argon plasma
ICp	inhibiting concentration for a (specified) percent effect
KCl	potassium chloride
kg	kilogram(s)
L	litre(s)
lm	lumen(s)
LOEC	lowest-observed-effect concentration
LSD	least significant difference
m	metre(s)
M	mole(s) (concentration)
max	maximum
mg	milligram(s)
min	minimum
mL	millilitre(s)
mm	millimetre(s)
mo	month(s)
mS	millisiemens
n	sample size
N	nitrogen
nm	nanometre(s)
NOEC	no-observed-effect concentration
OM	organic matter
P	probability
PAHs	polycyclic aromatic hydrocarbons
QA/QC	quality assurance/quality control
s	second
SD	standard deviation
sp.	species (singular)
spp.	species (plural)
™	trademark
TOC	total organic carbon
v:v	volume-to-volume
WHC	water-holding capacity
wk	week(s)
wt	weight
µg	microgram(s)
µm	micrometres
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., shoot/root)
≈	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

Acuminate means tapering gradually to a sharp point.

Allelopathic refers to the ability of a species to produce chemicals that inhibit the growth of other species.

Alluvial soil is a type of azonal soil formed on the flood plains of river valleys and at river mouths (alluvial fans or deltas). New material is successively deposited on the surface when the land is subject to flooding.

Ament (catkin) is a cylindrical cluster of unisexual flowers that lack petals. Catkins are found on willows, birches and oaks.

Angiosperm is a term used in plant classification referring to plants that flower, and whose ovules (young seeds) are enclosed in an ovary. The ovary matures into a fruit with seeds, following fertilization. The Phylum Magnoliophyta (or Anthophyta) contains all angiosperms and is the largest and most diverse group within the Kingdom Plantae. Two Classes of angiosperms include the Class Magnoliopsida *(Dicotyledons)* and the Class Liliopsida *(Monocotyledons)*.

Auricles are claw-like appendages that occur at the junction of the leaf sheath and leaf blade and that clasp the stem.

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

Biserrate means having double serrations (i.e., having a second set of smaller serrations along each main serration).

Browse is the portion of woody growth (leaves and twigs) of shrubs, trees and woody vines that is available for animal consumption.

Canopy for the purpose of this method is the more or less continuous cover produced by the foliage of plants.

Chlorosis is a condition in which the green parts of plants have depressed concentrations of chlorophyll and the leaves are pale green or yellow in colour. This might result from disease, exposure to toxic substances, nutrient deficiencies or senescence.

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

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

Conifer (coniferous) is a cone-bearing tree of the Pine family, usually evergreen. (See also *deciduous.*)

Cool-season (C_3) plants are those that grow best at cooler temperatures, and are sometimes called C_3 plants because they have the C_3 photosynthetic pathway. This means that the first stable organic compound that is produced during the fixation of CO_2 from the atmosphere within the leaf cells contains three carbon atoms. Where both cool-season plants and *warm-season* plants grow together, the cool-season (C_3) plants grow earlier in the spring than the warm-season (C_4) plants. Most plants that grow in temperate regions are C_3 plants.

Cotyledon is a primary leaf of the developing embryo in seeds; there is only one in monocotyledonous plants, and two in dicotyledonous plants. In many dicotyledonous species, such as the bean, the cotyledons emerge above ground and appear as the first leaves.

Deciduous describes plants that shed their leaves at the end of the growing season. (See also *coniferous*.)

Defoliation is the condition in which a plant does not have a normal complement of leaves due to some internal or external cause.

Desiccation is a state of dryness or the process of drying. It can be applied to plants, portions of plants or soil.

Dicotyledon is a term used in the classification of plants that refers to those species having two seed leaves (*cotyledon*).

Dioecious describes a plant species having separate male and female flowers on different plants. (See also *monoecious.*)

Ecological risk assessment 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).

Emergence occurs following the *germination* of a plant, wherein the early *growth* of a seedling pushes the *epicotyl* through the soil surface. In this test method, a seedling is considered emerged if there is at least 3 mm of vegetative growth (measured as the stem from the surface of the soil to the tip of longest leaf) either vertically above the surface of the soil, or horizontally across the surface of the soil.

Epicotyl is that portion of an embryo or seedling containing the shoot. It is delineated anatomically by the tissue transition zone that separates the epicotyl from the *hypocotyl*.

Epigeal (germination) refers to a type of *germination* in which the *hypocotyl* is active and pulls the *cotyledon*s above ground during its growth. Germination begins with the *imbibition* of water and proceeds with *emergence* of the *radicle* from the seed to form the primary root and secondary roots; elongation of the active hypocotyl follows with the hypocotyl arch penetrating through the soil surface. Epigeal-emerging dicots (i.e., 90% of dicotyledonous plants) have the advantages of being able to commence photosynthesis as soon as the cotyledons emerge, and of being able to expand leaf area rapidly. (See also *hypogeal*.)

Forbs are herbaceous species other than graminoids that do not look like grasses, usually having broad leaves. (See also *Graminoid*.)

Germination refers to the process by which the plant embryo within the *seed* resumes *growth* after a period of dormancy and the *seedling* emerges from the seed. (See also *epigeal* and *hypogeal*.)

Graminoid refers to the grasses and other related species that have a grass-like appearance. The sedges are a common group of plants that are considered to be graminoids. (See also *Forb*.)

Growing degree days are a measure of the amount of heat available for plant *growth*. For each day, the number of degrees by which the mean temperature exceeds a base temperature (usually 5ºC) is calculated. Daily values are added as the season progresses to give a cumulative sum of growing degree days. The development of plants is usually closely related to the accumulation of growing degree days.

Growth is the increase in size or weight as the result of proliferation of new tissues. In this test method, growth refers to an increase in shoot and root length, as well as an increase in shoot and root dry and wet weights.

Gymnosperms are the conifers and their allies, a subdivision of the spermatophytes or seed-bearing plants. They are distinguished from *angiosperms* by having the ovules unprotected on the surface of the megasporophylls (scales of the cone). This is in contrast to the angiosperms which have the seeds enclosed in an ovary. (See also *angiosperm*.)

Herbs (herbaceous) are plants that have no woody tissue and die back to the ground at the end of the growing season. Herbs are divided into graminoids and forbs. (See also *Forb* and *Graminoid.*)

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

Hull is the dry outer covering of a *seed*.

Hypocotyl is that portion of an embryo or seedling containing the root or *radicle*. It is delineated anatomically by the tissue transition zone that separates the *epicotyl* from the hypocotyl.

Hypogeal (germination) refers to a type of *germination* in which the *hypocotyl* is inactive and the scutellum (*cotyledon*) remains below the ground. The *radicle* emerges first to form the primary root, followed by the *coleoptile*. *Emergence* is largely dependent on elongation of the coleoptile and the first internode. When the soil surface is reached, light inhibits further growth and true leaves emerge through the hollow sheath. All grasses (e.g., barley) are characterized by hypogeal germination.

Imbibition is the initial period of germination characterized by the rapid uptake of water by a dry seed.

Inflorescence is the part of the shoot of seed plants where flowers are formed.

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.

Lacustrine means of, or relating to, lakes. Lacustrine soils are formed from sediments that were laid down in a lake bottom and subsequently exposed.

Lanceolate means literally "lance shaped." The term is generally used to describe a leaf shape that is rounded at the base and gradually tapers to a point.

Ligule is a membranous projection on the side of a grass leaf facing the stem and that occurs at the junction of the leaf sheath and the leaf blade.

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/($m^2 \cdot s$)], the spectral quality of the light source must be known. Light conditions or irradiance are properly described in terms of quantal flux (photon fluence rate) in the photosynthetically effective wavelength range of approximately 400−700 nm. The relationship between quantal flux and lux or foot-candle is highly variable and depends on the light source, the light meter used, the geometrical arrangement and the possibility of reflections (see ASTM, 2008). The approximate conversion between quantal flux and lux, however, for full-spectrum fluorescent light (e.g., VitaLux® by Duro-Test®), is 1 lux ≈ 0.016 µmol/($m^2 \cdot s$) (Deitzer, 1994; Sager and McFarlane, 1997).

Malformation is a structural defect that occurs infrequently and is due to abnormal development.

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.

Monocotyledon is a term used in the classification of plants that refers to those species having a single seed leaf (*cotyledon*).

Monoecious describes a plant species having separate male and female flowers on the same plant. (See also *dioecious*.)

Mottling means marked with spots or streaks of different colours (e.g., blotched). This includes the discoloration of leaf margins.

Necrosis refers to dead tissue.

Panicle is a branched cluster of flowers in which the branches are *racemes*.

Perennial is a plant that, under natural conditions, lives for several to many growing seasons.

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.

Phytomass is the total weight (mass), either above and/or below ground, of a group of plants.

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.

Protocol is an explicit set of procedures for a test or an experiment, 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.

Raceme is an unbranched flower cluster, consisting of a single central stem, with individual flowers growing at intervals on small stalks.

Radicle is the end of a plant embryo that gives rise to the first root.

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 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).

Rhizomatous – see *rhizome*.

Rhizome is a fleshy, creeping, horizontal, underground stem that often sends out roots and shoots from its nodes. Certain plants reproduce vegetatively by means of their rhizome(s).

Risk assessment − see *ecological risk assessment*.

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

Root is usually the below-ground portion of a plant that serves as support, draws minerals from surrounding soil and sometimes stores food. There are two main types of root systems: the tap root system, in which there is a main primary root larger than the other branching roots, and the fibrous root system, in which there are many slender roots with numerous smaller root branches. (See also *shoot*.)

Seed is a fertilized and ripened plant ovule consisting of the plant embryo, varying amounts of stored food material and a protective outer seed coat.

Seed pretreatment is a coating (e.g., fungicide) applied to seeds before water *imbibition*.

Seedfall is the combined quantity of seeds falling from the plant community or from a particular species within the plant community.

Seedling is a young plant that is grown from a seed.

Seral (stage) refers to a plant community that is one of a stage of communities that develop on a site over time following a disturbance.

Serotinous refers to plants having an ecological adaptation in which seed is released slowly or in response to an environmental trigger. An example of such a trigger is fire.

Shoot is the usually above-ground portion of the plant such as the stems and leaves.

Staining is the discoloration of plant parts (roots, vegetative growth) caused by the test substance.

Stratification is the process by which water is imbibed into seeds from the surrounding medium during cool temperature storage. Stratification of seed improves speed and percent germination and may serve to improve synchronization of germination in some species.

Warm-season (C_4) plants are those that grow best at warmer temperatures and have the C_4 photosynthetic pathway. This means that the first stable organic compound that is produced during the fixation of CO_2 from the atmosphere within the leaf cells contains four carbon atoms. [See also *cool-season (C_3)* plants.]

Wilting occurs when plant tissues lose their turgidity and the plant becomes limp.

Withering is the process of drying; plants become limp and desiccated. This is frequently the result of root damage.

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.

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 also *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 enter the environment through spillage, application or discharge.

Chemical-spiked soil is natural or *artificial soil* to which one or more chemicals or chemical products have been added. (See also *spiked soil*.)

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

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

Consolidated soil 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 increased concentrations, often because of some direct or indirect human activity. The term is frequently applied to substances or materials present at concentrations having 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)*.]

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

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

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

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 de-ionized or distilled water, reverse-osmosis water, or de-chlorinated tap water. Depending on study design and intent, a surface water or groundwater from the *site* might be used instead of de-ionized or distilled water for the hydration of each test soil (including negative control soil). (See also *test water*, *de-ionized water* and *distilled water*.)

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

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

Negative control – see *control*.

Negative control soil is *clean* soil that does not contain concentrations of one or more contaminants that could affect the *emergence*, survival or growth 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 emergence and growth of the test plants during the test. The use of negative control soil provides a basis for interpreting data derived from toxicity tests using test soil(s).

Organic matter (OM) in soil consists primarily of plant and animal residues, at different stages of decomposition, including soil humus. The accumulation of OM within soil is a balance between the return or addition of plant and animal residues and their subsequent loss due to the decay of these residues by soil micro-organisms. For many types of soil, the following equation (AESA, 2001) is suitable for estimating the total OM content of soil from total organic carbon (TOC) measurements: % OM = % TOC × 1.78; 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 soil that contains known concentrations of one or more contaminants that adversely affect the *emergence*, survival or *growth* 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*. The test soil might be either field-collected *site soil* that is *contaminated* or potentially so, or *chemical-spiked soil*. Reference soil used in a test frequently exhibits physicochemical properties (e.g., *texture*, *organic matter* content, organic carbon content, *pH*, *conductivity*, and *fertility*) closely matching those of the test soil sample(s), except that it is free from the source of contamination being assessed. In tests involving samples of *site soil*, one or more samples of reference soil are often selected from the general location of test soil sampling, and thus might be subject to other sources of contamination aside from the one(s) being studied. Reference soil is used to describe matrix effects in the test, and may also be used as a diluent to prepare concentrations of the test soil. In tests involving *chemical-spiked soil*, one or more samples of artificial (formulated) soil with differing physicochemical characteristics might be chosen to investigate the influence of certain soil properties (e.g., soil *texture*, or percent *organic matter*) on the toxicity of a chemical mixed in each of these soil types. (See also *negative control soil*, *site soil*, *test soil*, *clean*, *artificial soil* and *chemical-spiked soil.*)

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

Reference toxicity test is a test conducted using a *reference toxicant* in conjunction with a soil toxicity test, to appraise the sensitivity of the organisms and/or the precision and reliability of results obtained by the laboratory for that chemical at the time the test material or substance is evaluated. Deviations outside an established normal range indicate that the sensitivity of the test organisms and the performance and precision of the test are suspect. A reference toxicity test with plants is often 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 it being *contaminated* or potentially contaminated by human activity.

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 plants. 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, invertebrates and plants) 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 consistence, and in chemical, biological, or mineralogical composition.

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

Stock solution means a concentrated solution of the substance(s) to be tested, 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 (*de-ionized* or *distilled water*, or equivalent), with or without the inclusion of an organic solvent.

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

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

Test soil is a sample of field-collected soil or *chemical-spiked soil* to be evaluated for toxicity to plants. In this method, test soil is 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*, dredged material) used in a soil toxicity test.

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

Texture is defined based on a measurement of the percentage by weight of sand, silt, and clay in the mineral fraction of soils. Classification as to texture confers information on the general character and behaviour of substances in soils, especially when coupled with information on the structural state and organic matter content of the soil. 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 (percentages) of sand, silt, and 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, 2012b). (See also *organic matter*.)

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

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 as to which *a priori* tests to select. (See also *post hoc.*)

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

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

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:

$$CV\ (\%) = 100 \times (SD \div mean).$$

Endpoint means the measurement(s) or value(s) that characterize the results of a test (e.g., IC25). It also means the response of the test organisms that is being measured (e.g., shoot/root length and weight).

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 to the control, in a *quantitative* (continuous) biological measurement such as length of shoots attained by individual plants at the end of the test.

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.

Phytotoxicity means unwanted detrimental deviations from the normal pattern of appearance, *growth*, and/or function of plants in response to the test material. Phytotoxicity might occur during *germination*, growth differentiation, and/or maturation of plants.

Post-hoc literally refers to something performed after the fact, or "after this." In the context of test design and statistics, *post hoc* tests are those that are decided on after the data have been collected. Used in a more general sense, the purpose of the *post hoc* test is to determine which treatment means are different from each other, while adjusting for the overall Type I error rate. (See also *a priori.*)

Precision refers to the closeness of repeated measurements of the same quantity to each other, i.e., the degree to which data generated from repeated 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, a seedling might fail to emerge from *contaminated* test soil. Generally, quantal effects are expressed as numerical counts or percentages thereof. (See also *quantitative*.)

Quantitative effects in a toxicity test are those in which the measured effect is continuously variable on a numerical scale. Examples would be shoot length of emerged seedlings or dry weight of roots at the end of the test. 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 unit; 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 4.8.2.1, 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, below the concentration or level of contamination that directly causes death within the test period.

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

Toxicant is a toxic substance or material.

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

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

Toxicology is a branch of science that studies the toxicity of substances, materials, or conditions. There is no limitation on the use of various scientific disciplines, field or laboratory tools, or studies at various levels of organization, whether molecular, single species, populations, or communities. Applied toxicology would normally have a goal of defining the limits of safety of chemical or other agents. (See also *environmental toxicology*.)

Treatment refers to a specific *test soil* (e.g., a *site soil*, *reference soil*, or *negative control soil*) from a particular *sampling 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*.)

Visual assessment represents the description of any visual damage to the test species based on observations of phytotoxicity (i.e., *malformation*, *chlorosis*, *necrosis*, *defoliation*, *desiccation*, *mottling*, *staining*, *wilting*, or *withering*) observed in test vessels with *contaminated* soil compared to the controls.

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

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

Acknowledgements

This document was co-authored by M. Moody and B. Godwin (Saskatchewan Research Council [SRC], Saskatoon, SK) and J. Miller (Miller Environmental Sciences Inc., King City, ON). R.P. Scroggins (Biological Assessment and Standardization Section, Environment Canada [EC], Ottawa, ON) acted as Scientific Authority and provided technical input and direction throughout the work. J.I. Princz (Soil Toxicology Laboratory, Biological Assessment and Standardization Section, EC, Ottawa, ON) and L. Van der Vliet (Method Development and Applications Unit, EC, Ottawa, ON) provided valuable advice and technical input throughout the work. The studies resulting in the biological test method defined herein were performed by the Toxicology section of the SRC, directed by M.J. Moody. Additional method development studies were carried out by the Soil Toxicology Laboratory, under the direction of J.I. Princz. Many thanks to D. Bolin, R. Suchan, C. Yackel and V. Ziegler (SRC, Saskatoon, SK) and H. Lemieux and M. Malette (Soil Toxicology Laboratory, Biological Assessment and Standardization Section, EC, Ottawa, ON) for their valued contribution to the development of this method.

The interlaboratory studies undertaken to validate the biological test method were coordinated by the Soil Toxicology Laboratory and performed by the following participating laboratories: Exova (Edmonton, AB), HydroQual Laboratories Ltd. (Calgary, AB), Maxxam Analytics (Burnaby, BC), SRC (Saskatoon, SK), and Stantec (Guelph, ON). Names of laboratory personnel participating in these tests are provided in the technical report on these studies (EC, 2013a).

We gratefully acknowledge the input provided by the participants of the boreal species priority ranking, including: B. Godwin, J. Thorpe, and M. Johnston (SRC, Saskatoon, SK), and D. Jiricka and M. McLaughlan (Saskatchewan Environment, Prince Albert, SK).

The Biological Assessment and Standardization Section would like to thank J. Nelson and the staff of EcoDynamics Consulting Inc. (Prince Albert, SK) for assisting both EC and SRC laboratories in identifying and collecting most of the field reference and contaminated boreal forest soils.

Finally, we are very appreciative of the unwavering funding support provided by the Program for Energy Research and Development (PERD) that has enabled the development and validation of this novel test method.

Introduction

1.1 Background

The *Method Development and Applications Unit (MDAU)* of Environment Canada is responsible for the development, standardization and publication (see Appendix A) of a series of biological test methods for measuring and assessing the *toxic* effect(s) on single species of terrestrial or aquatic organisms, caused by their exposure to samples of test *materials* or *substances* under controlled and defined laboratory conditions. In 1994, the MDAU, the Canadian Association of Petroleum Producers (CAPP), and the federal Panel for Energy Research and Development 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 new biological test methods that were applicable to diverse types of Canadian soil using terrestrial species that were representative of Canadian soil ecosystems. The initial phase of this multi-year program involved a comprehensive review of existing biological test methods, used globally to evaluate the toxicity of *contaminated* soils to plants and soil invertebrates. The resulting report recommended that Environment Canada support the development, standardization, and publication of a number of single-species biological test methods for measuring soil toxicity, including those using terrestrial plants (Bonnell Environmental Consulting, 1994). This recommendation was endorsed by both the headquarters and regional offices of Environment Canada (Appendix B) and the Inter-Governmental Ecotoxicological Testing Group (Appendix C).

Since 1994, a number of standardized toxicity test methods have been published by Environment Canada including Biological Test Method: Tests for Toxicity of Contaminated Soil to Earthworms *(Eisenia andrei, Eisenia fetida* or *Lumbricus terrestris)*, EPS 1/RM/43 (EC, 2004a); Biological Test Method: Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil, EPS/1/RM/45 (EC 2005a); and Biological Test Method: Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil, EPS/1/RM/47 (EC 2007a).

In 2003, Environment Canada's MDAU convened a three-day workshop on the toxicological assessment of Canadian soils and development of standardized testing tools. Based on pre-workshop background materials (a questionnaire), plenary sessions, and working group discussions, participants identified areas considered priorities for research and development. It was recommended that priority should be given to dedicating resources for the development of plant test methods using species that are more reflective of non-agricultural soils and/or habitats. With over 50% of Canada's total land mass being comprised of the boreal and taiga ecozones, and the contribution of resources within these ecozones to Canada's economy via oil and gas, mining and forestry industries, priority was given to the development of standardized tests applicable to the assessment of *contaminants* present in boreal forest soils. Since then, several years of research have been completed on the selection of suitable and sensitive test organisms for measuring soil toxicity to meet the needs of industry, Canadian regulatory and *monitoring* requirements, and on the development of appropriate biological test methods. A technical report was produced describing species selection criteria and recommending boreal forest plant species suitable for test method development (SRC, 2003). Subsequent technical reports described aspects of method development including *seed* acquisition, *germination* and *stratification* of seed, selection of suitable test durations, *growth* in reference soils, and testing of a *reference toxicant* (SRC, 2004, 2006). Additional method development included investigations of soil *pH* tolerance, seed source variability, testing in intact soil *cores*, improvement of paper birch *emergence*, and testing of boreal forest reference soils and soils contaminated by brine and hydrocarbons (SRC, 2007, 2008, 2009, 2012; EC and SRC 2007; EC 2010, 2013b). The methodology described in this report is based on Environment

Canada's Biological Test Method: Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil (EC, 2005a), with modifications required for testing with boreal plant species.

Numerous soil toxicity tests have been coordinated or supported by Environment Canada, using various plant species, native to the boreal forest region, exposed to samples of soil contaminated with metals, salts, hydrocarbons, or prospective reference toxicants. These studies (SRC, 2003, 2004, 2006, 2007, 2008, 2009; EC and SRC 2007; EC 2010, 2013b) focused on the development and standardization of biological test methods for determining the *sublethal* toxicity of samples of *contaminated* soil to native boreal forest plants. Based on the results of these studies, together with the findings of a series of interlaboratory method validation studies (EC, 2013a), Environment Canada proceeded with the preparation and finalization of a biological test method for conducting soil toxicity tests that measure *growth* inhibition of terrestrial plant species native to the Canadian boreal region, as described in this report.

A group of ecologists with long-term experience in the boreal forest (see Appendix D) served actively in reviewing and ranking candidate plant species based on specific criteria (SRC, 2003).

Detailed procedures and conditions for preparing and performing this biological test method are defined herein. Universal procedures for preparing and conducting soil toxicity tests using selected species of plants native to the Canadian boreal forest are described. Options for test species include: trembling aspen *(Populus tremuloides)*, bluejoint reedgrass *(Calamagrostis canadensis)*, Canada goldenrod *(Solidago canadensis)*, paper birch *(Betula papyrifera),* jack pine *(Pinus banksiana)*, white spruce *(Picea glauca),* and black spruce *(Picea mariana)*. 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 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 plant *growth* (measured as live *shoot* and *root* length and shoot and root dry mass) measured at the end of the test. Due to the insensitivity of the *emergence* endpoint, it is not included as a statistical endpoint though it is measured in all test vessels in order to calculate the mean shoot and root dry weights on a per plant basis and in the control treatments for test validity purposes.

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

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

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

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

(3) *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 most situations in which they will be applied. It is assumed that the user has a certain degree of familiarity with soil toxicity tests. Explicit instructions that might be required in a regulatory *protocol* are not provided in this report, although it is intended as a guidance document useful for that and other applications.

For guidance on the implementation of 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 in Environment Canada's Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology (EC, 1999a).

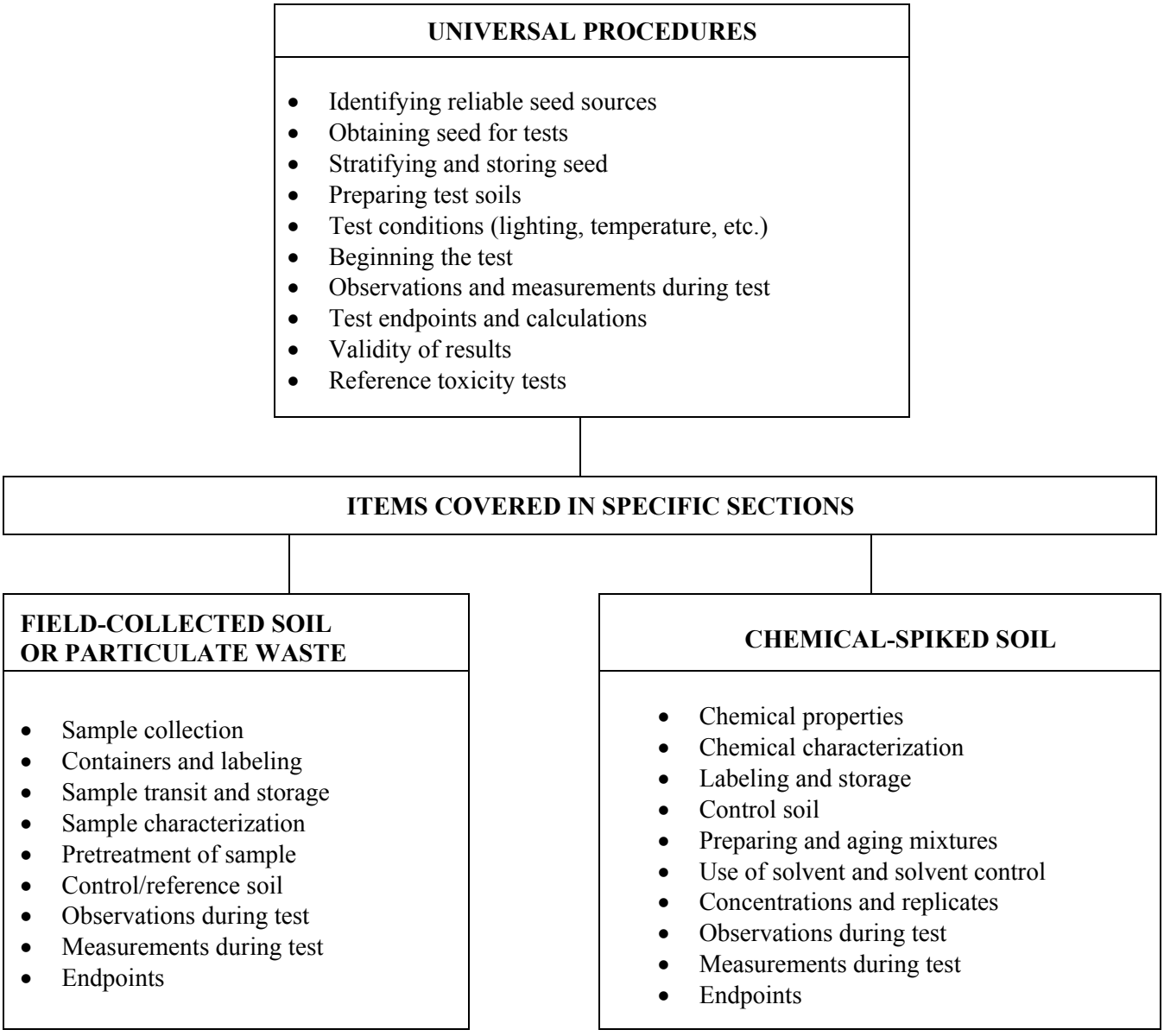


Figure 1 Considerations for preparing and performing soil toxicity tests using boreal forest plants and various types of test materials or substances

1.2 Selection of Test Species

Plant community data from upland boreal forest sites across Canada were reviewed with the aim of compiling a list of potential species from which ecologically relevant test species could be selected (SRC, 2003). The boreal forest is considered to include seven ecozones (AAFC). These ecozones are the Taiga Plains, Taiga Shield, Boreal Shield, Boreal Plains, Taiga Cordillera, Boreal Cordillera, and Hudson Plain. The provinces and territories with forest areas included within these ecozones are the Yukon Territory, Northwest Territories, the territory of Nunavut, and the provinces of British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec, and Newfoundland and Labrador. While New Brunswick, Nova Scotia, and Prince Edward Island fall entirely within the Atlantic Maritime Ecozone, many boreal plant species do occur in forests of this ecozone, and some forest areas within this region may be considered to be boreal in nature, just as most forest stands in the Aspen Parkland ecoregion of the Prairie Ecozone can be considered outliers of the boreal forest.

From an initial list of 219 candidate species, 40 were selected on the basis of 8 predetermined criteria of varying levels of importance (SRC, 2003), including those that were considered to be:

(1) extremely important:
- reproduction from seed (plants with microscopic seed or reproducing from spores were excluded),
- native origin (exotic species were excluded), and
- widespread occurrence (species that did not occupy at least half of the east-west extent of boreal forest were excluded, the border being between Manitoba and Ontario);

(2) moderately important:
- taxonomic concerns (species that were difficult to identify and might be wrongly identified by suppliers were excluded),
- growth rate (slow-growing plants unlikely to provide a reasonably sized *seedling* in the early growth stages were excluded), and
- stress tolerance (species of the family Ericaceae are largely stress tolerant and were therefore excluded); and

(3) less important:
- abundance (species that are rare, endangered, or infrequently occurring were excluded), and
- upland, boreal habitat (boreal species that were predominant in wetland or atypical habitats or species abundant primarily in other ecosystems were excluded).

This list of species was further reduced to 20[1] through rankings by a group of ecologists (see Appendix D) based on two factors:

(1) the species' ability for rapid growth from seed, and

(2) availability of seed from government or commercial suppliers or ease of wild seed collection.

Since many species received equal rankings, it was considered prudent to include both woody (six trees and five shrubs) and *herbaceous* species (seven *forbs* and two grasses) in the final test species options. Each species was assessed (SRC, 2004, 2006, 2007, 2008, 2009) based on criteria used to evaluate its amenability to the methods described in Environment Canada's test for measuring emergence and growth of terrestrial plants, already in use (EC, 2005a). Modifications to this standardized method were necessary, primarily because the seeds of some native species required *stratification* and longer test durations. Seed was obtained from Canadian commercial or government sources so that the genetics would reflect at least one region of Canada. Aspects of the test methodology were investigated in either scientific literature or laboratory experiments and included: stratification and *germination* of *seed*, growth of reasonably sized *seedlings* in artificial and forest reference soils, layering of natural *soil horizons* in tests, test durations and *biomass* at the end of the tests,

[1] The 20 plant species included: balsam fir *(Abies balsamea)*, black spruce *(Picea mariana)*, bluejoint (marsh) reedgrass *(Calamagrostis canadensis)*, chokecherry *(Prunus virginiana)*, cow parsnip *(Heracleum lanatum),* dewberry *(Rubus pubescens),* fireweed *(Chamerion angustifolium)*, jack pine *(Pinus banksiana)*, Lindley's aster *(Aster ciliolatus)*, paper (white) birch *(Betula papyrifera)*, pincherry *(Prunus pensylvanica)*, prickly rose *(Rosa acicularis),* purple oat grass *(Schizachne purpurascens)*, red baneberry *(Actaea rubra),* Saskatoon *(Amelanchier alnifolia),* tall lungwort *(Mertensia paniculata),* trembling aspen *(Populus tremuloides)*, white spruce *(Picea glauca),* wild raspberry *(Rubus idaeus)*, wild sarsaparilla *(Aralia nudicaulis).*

and ease of handling during processing. Fireweed *(Chamerion angustifolium)* was eliminated from testing as the fragility of its roots caused loss of biomass during processing at the end of the test. Canada goldenrod *(Solidago canadensis)* was added as an option for test species as it was found to be ecologically relevant and performed well in growth tests (SRC, 2004). Paper birch *(Betula papyrifera)* was added to provide an acid-tolerant *deciduous* tree species.

The performance of test species was evaluated in *growth* tests using eight forest reference soils from various Canadian locations, representing a broad range of soil types.[2] Testing of a *reference toxicant* (boric acid), and of salt- and hydrocarbon-contaminated soils demonstrated that the test species options carried a range of sensitivities to different *toxicants* (SRC, 2004, 2006, 2007, 2008, 2009; EC and SRC 2007; EC 2010, 2013b).

The seven boreal forest plant species selected for use in this test method are described in detail in the following subsections and summarized in Table 1. Physicochemical characteristics of the soils are compiled in Appendix F.

*1.2.1 Trembling Aspen (*Populus tremuloides *Michx.)*

Trembling aspen *(Populus tremuloides)*, a member of the willow family (Salicaceae), is known as quaking aspen in the United States but also has several other less commonly used names including mountain aspen, golden aspen, poplar, and trembling poplar. It is a *dioecious* (sexes on separate trees), broad-leaved, *deciduous* tree. Its small round leaves have an *acuminate* tip and a distinctive strongly flattened leaf petiole that causes the leaves to flutter in even slight winds. The bark is smooth and light green to grey in colour, often becoming black and furrowed near the ground, particularly with age. Trembling aspen is used for plywood, veneer, oriented strand board, waferboard and pulp, with some lumber derived from this species (Watson *et al*., 1980; Howard, 1996). It has higher protein levels than most other *browse*, and is therefore an important food source for large ungulates and smaller mammals, particularly in the fall and winter. In addition, trembling aspen stands provide important breeding habitat for many bird species (Watson *et al*., 1980; Howard, 1996).

Trembling aspen is a *dicotyledonous* plant with *epigeal germination*. It is a relatively short-lived tree. Aspen stands generally start to deteriorate at 80 to 100 years after establishment (Watson *et al*., 1980). The oldest trees reach a maximum age of about 200 years, but individual clones, which are regenerated trees from the root system of a single, earlier, seed-produced tree, may be thousands of years old (Perala, 1990). Aspens average 13 to 20 m in height and 20 to 25 cm in diameter (Watson *et al*., 1980), but may attain 30 m height and 60 cm diameter. Aspens have wide spreading roots, with suckers found 32 m from parent trees (Watson *et al*., 1980).

Trembling aspen can be grown from seeds collected in the same year, but seeds are short-lived. Seeds should be stored under refrigerated or freezer conditions (2°C to -40°C), and have been successfully stored for up to 2 years (Harrington, 2010). Germination rates have been reported as ranging from 50 to > 90% (Watson *et al*., 1980; Luna *et al*., 2001; Moench, 2001). Seeds are nondormant (i.e., no *stratification* required for *germination*), and germination begins within 24 hours after sowing, with light required for germination to occur (Luna *et al*., 2001).

In laboratory tests, germination of aspen begins quickly, and emergence reaches its maximum within 4 days when planted on the soil surface and using a cycle of 16 h light:8 h dark (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011). Germination was ≥ 80% in the majority of tests in artificial and reference soils but ranged from 60 to 100% in all tests that met the validity criteria (EC, 2013b). Aspen roots are fine and fibrous, requiring care when washing (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2012). Roots are generally at least twice as long as shoots (Figure 5), but generate less dry mass (EC, 2013b). Growth of both shoots and roots is reduced in low pH (< 3.7) soils such as the podzol reference

[2] The boreal reference soils included Gleyed Humo-ferric Podzols from Newfoundland, New Brunswick, and Ontario; a Dark Grey Luvisol, an Ortho Eutric Brunisol and Eluviated Dystric Brunisol from Saskatchewan; and a Rego Humic Gleysol and Rego Dark Grey Chernozem from Alberta.

Table 1 Characteristics of plant species

Plant	Phylogeny	Growth form	Germination	Life cycle	Rooting depth[a]	Soil type preference	Tolerance
Trembling aspen	angiosperm, dicotyledon	tree, deciduous	epigeal	perennial	2 m	Grows in a wide range of soil types	Low tolerance for drought, shade, standing water, anaerobic conditions; tolerates pH 3.7–>7.7[b].
Bluejoint reedgrass	angiosperm, monocotyledon	herb, graminoid	hypogeal	perennial	40 cm	Grows on fine to coarse textured wet to mesic soils, proliferates on disturbed sites and wetlands	Tolerates soil pH 3.3–>7.7[b], tolerates flood, drought and saline conditions; prefers moist, nutrient-rich soils
Canada goldenrod	angiosperm, dicotyledon	herb, forb	epigeal	perennial	30 cm (min)[c]	Grows in a wide variety of soil types, fine to coarse textures	Tolerates soil pH 5.0–>7.7[b]
Paper birch	angiosperm, dicotyledon	tree, deciduous	epigeal	perennial	60 cm (min)	Grows in a wide variety of soil types, fine to coarse textures	Intolerant of shade and salinity, tolerant of acid soils, tolerates soil pH 3.3–>7.5[b]
Jack pine	gymnosperm	tree, coniferous	epigeal	perennial	50 cm (min)	Grows in soils of fine to coarse textures	Tolerant of drier sites and lower fertility, tolerates soil pH 3.3–>7.8[b]
White spruce	gymnosperm	tree, coniferous	epigeal	perennial	55–100 cm	Grows in soils of fine to coarse textures	Intolerant of salinity, tolerates soil pH 3.2–>7.5[b]
Black spruce	gymnosperm	tree, coniferous	epigeal	perennial	40 cm (min)	Grows in soils of fine to coarse textures	Tolerates soil pH of 3.2–>7.6[b], high moisture requirement

[a] Root depth of mature plant in boreal forest.
[b] Value based on laboratory testing of reference soils (EC, 2013b).
[c] min = minimum.

soils collected in Ontario and Newfoundland (EC, 2013b). Aspen shows a strong concentration-response relationship to boric acid (SRC, 2006), metals (Wren *et al.,* 2012) and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012). Seeds are extremely small (0.5 × 1 mm) and amount to 5.5×10^6 to 8.0×10^6 seeds/kg (Perala, 1990). Aspen is wind-pollinated and generally flowers in April or May throughout its range (Perala, 1990). Initially, aspen grows rapidly, reaching 15 to 30 cm from seed in the first year, and 2 m in the second, with height growth continuing at a rapid pace for the first 20 years.

Aspen occupies a wide range of sites, from well-drained to poorly drained, and is also found on a wide range of soil *textures* (Watson *et al*., 1980). It has been found on acidic soils in Sudbury, Ontario (pH 3.2 to 4.5), which contradicts reports of its low tolerance of acid soils (Watson *et al*., 1980).

Trembling aspen is the most widely spread North American tree, occurring from Alaska to Newfoundland and south to northern Mexico in the west and to Virginia in the east (Howard, 1996). Worldwide, only the closely related European aspen *(Populus tremula)*, which can hybridize with trembling aspen, and Scots pine *(Pinus sylvestris)* have larger natural ranges. Following fire, the main disturbance event in the Western Canadian boreal forest region, trembling aspen frequently forms nearly pure stands along the forests' southern edge. Some of the best stands of aspen are found on glacial drift, rich in lime, from Saskatchewan through Manitoba and into the northern part of the Great Lakes states (Perala, 1990).

*1.2.2 Bluejoint Reedgrass [*Calamagrostis canadensis *(Michx.) P. Beauv.]*

Calamagrostis canadensis is a member of the grass family (Gramineae) and is commonly known as bluejoint reedgrass or marsh reed grass. Other common names include bluejoint, Canadian reedgrass, meadow pinegrass, marsh pinegrass, and marsh reedgrass. Bluejoint reedgrass is a relatively tall grass, growing from 50 to 110 cm tall, with leaf blades up to 10 mm wide and 40 cm long. Leaves are flat and lax with an *acuminate* tip, being ridged and scabrous (rough) on the upper surface and smooth or only slightly scabrous on the lower surface. Leaves have a prominent midrib. A long papery *ligule* is present and *auricles* are absent (Best *et al*., 1971). The *inflorescence* is an open *panicle* up to 20 cm long with single flowers in each spikelet (Rook, 1998). Flowering occurs between May and August. Bluejoint reedgrass has been used for stuffing mattresses and as a storage pit lining and cover for potatoes by the Woodland Cree (University of Michigan, 2003). Due to its low protein value, it is considered only fair to poor in its usefulness as forage to large ungulates, small mammals, and waterfowl (Tesky, 1992).

Bluejoint reedgrass occurs as a major *graminoid* in many of the aspen and balsam poplar plant community types of the Boreal Transition Ecoregion, but is a minor component or absent from the treed communities of the Aspen Parkland Ecoregion, which is the transition zone between the grassland and the boreal forest (Thorpe and Godwin, 2008). From a forestry perspective, bluejoint reedgrass is considered a problem species as a competitor for regenerating *conifer* species in the boreal forest (Winder, 1999), and also for aspen (Powell and Bork, 2004). Bluejoint reedgrass is also used for revegetation purposes (Wynia, 2006).

Bluejoint reedgrass is a *cool-season* (C_3) (Emery and Gross, 2007), *perennial*, *rhizomatous* grass with shallow fibrous roots. Like all grasses, it is a monocot. It is long-lived and once established, very dense stands may persist almost indefinitely. It flowers prolifically on disturbed sites and in wetlands. Seeds may be relatively short-lived; however, it has been reported that seed can be stored at least 2 years in cool dry conditions, and that it may remain viable in the soil for up to 5 years (Noller, 2001). Germination rates of seed from different sources and years apparently vary, with percentages ranging from low to over 95%, and germination times ranging from 3 to 14 days (Butler and Frieswyk, 2001; Tilley, 2010; Noller, 2001; Tesky, 1992). Germination is *hypogeal*. Seeds are non-dormant and it was reported that germination is not improved by *scarification*, *stratification* or light treatment (Wynia, 2006); however, stratification at 4 ± 2°C in moist peat for 2 to 6 weeks was found to improve both overall level and synchronization of germination (SRC, 2004).

In laboratory tests, emergence of bluejoint begins quickly and reaches its maximum within seven days when planted on the soil surface and using a cycle of 16 h light:8 h dark (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2012). Germination was ≥ 70% in the majority of tests in artificial and reference soils but ranged from 60 to 100% in all tests that met the validity criteria (EC, 2013b). Seeds measure 0.25 × 1 mm. Bluejoint roots are fine and fibrous, requiring care when washing (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011). The fine roots (see Figure 6) are similar in length to shoots, but generate less dry mass (EC, 2013b). Bluejoint is fairly tolerant of soils with pH as low as 3.3, although growth of shoots and roots was reduced in the Ontario (pH 3.7) and Newfoundland (pH 3.3) reference podzols but not in the New Brunswick (pH 4.1), Alberta AB01 (pH 3.5) and Saskatchewan SK03 (pH 3.7) reference podzols (EC, 2013b). Bluejoint shows a strong concentration-response relationship to boric acid (SRC, 2006) and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012).

Bluejoint reedgrass is a major species of many plant community types of the boreal forest, from wetland sites to moist upland forest sites (occupying imperfectly to moderately well-drained soils). It is tolerant of acidic soils to pH 3.5 (Rook, 1998) and of alkaline conditions to pH 8 (Wynia, 2006). The common occurrence of bluejoint reedgrass in northern forest communities is demonstrated by its inclusion as part of the descriptive name of 10 different plant community types described for the lower foothills subregion in Alberta (Lawrence *et al*., 2005). It occurs in wetlands that are flooded in spring and early summer but are too dry for the remainder of the summer to be dominated by sedges. It is also abundant in several plant communities where the dominant forest overstories are composed of trembling aspen (*Populus tremuloides*), balsam poplar (*Populus balsamifera*), and willow (*Salix alba*). Bluejoint reedgrass is found across the boreal forest from Alaska to Newfoundland and northwards onto the tundra. It extends southward to California, New Mexico, and across to Nebraska, Ohio, and Delaware (Scoggan, 1978).

*1.2.3 Canada Goldenrod (*Solidago canadensis *L.)*

A member of the aster (Asteraceae) family, Canada goldenrod *(Solidago canadensis)* is a long-lived *rhizomatous*, *perennial*, herbaceous, *dicotyledonous* plant native to and widespread in North America. It is also called common goldenrod. It occurs naturally in all political jurisdictions in Canada except for Nunavut (USDA-NRCS, 2010). Goldenrod grows and spreads rapidly from the *rhizomes*, commonly growing to 1 m in height, but reaching close to 2 m in ideal conditions (Coladonato, 1993). The plant is single-stemmed with serrated, *lanceolate*, 5- to 10-cm long leaves arranged spirally on the stem, and yellow floral heads arranged in a broadly pyramidal *panicle* occurring in late summer (Coladonato, 1993; Moss, 1994). The flowers are generally swept to the upper side of the panicle branches. The upper half of the stem is finely pubescent (hairy). Seeds are not considered to have dormancy and are reported to have a 50% germination rate (Wynia, 2002). In laboratory tests, emergence is generally complete within four days in *artificial soil* (see Table 2) and ranged from 60 to 100% in tests of artificial and reference soils (EC, 2013b). Goldenrod grows at pH values between 5 and 7.7 on coarse- to fine-textured soils. However, in laboratory tests of several reference soils (pH 5.8–7.7), emergence and growth endpoints were variable, making it difficult to recommend pH tolerance levels (EC, 2013b). Goldenrod shoot and root growth is reduced in low pH (≤ 4) soils such as the podzol reference soils collected in Saskatchewan SK03, Ontario and Newfoundland, and the Alberta AB01 reference soil (EC, 2013b). Seeds do not require treatment prior to planting (Wick *et al*., 2008), although Wynia (2002) suggests that *stratification* may improve germination. Stratification of seed in moist peat at 4 ± 2°C for 3 to 12 weeks is found to improve and synchronize germination (SRC, 2004). Seed viability is generally very low, with van der Grinten (2002) reporting a rate of only 12% live seed in collected batches. In contrast to this observation, during method development, goldenrod mean emergence was 82% in all tests that met the validity criteria and ranged from 60 to 100% (EC, 2013b). Seeds measure 0.25 × 1 mm. Goldenrod roots are fine and fibrous (see Figure 7), and can be expected to be significantly longer than shoots in most soils (EC, 2013b). Goldenrod shows a strong concentration-response relationship to boric acid

(SRC, 2006), metals (Wren *et al.,* 2012) and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012).

Canada goldenrod has historically been used for medicinal purposes by a variety of North American Aboriginal tribes (University of Michigan, 2003; USGS, 2006). It has low palatability to both browsing and grazing animals, with a high carbon-to-nitrogen ratio indicating relatively poor nutritional value (Coladonato, 1993; USDA-NRCS, 2010). However, white-tailed deer graze this species after *inflorescence* development (Coladonato, 1993).

Canada goldenrod also occurs naturally throughout the mainland United States except for the extreme southeastern states (USDA-NRCS, 2010). Although Canada goldenrod is a beloved garden plant in Europe, it is now considered an invasive species outside of North America (University of Michigan, 2003; Frankton and Mulligan, 1970). Canada goldenrod had spread to Europe by the 17th century and more recently to Asia, notably in China, and has become a problematic species. It is suggested that *allelopathic* effects may be responsible for its invasiveness (Sun *et al*., 2006). Canada goldenrod is not valuable economically but can be an important source of nectar for honey bees. It occupies a wide range of habitats, from forest to moist grassland, and frequently becomes abundant in disturbed habitats such as burned forests. It is typically found on moist, but not waterlogged, sites and can tolerate a wide range in soil fertility. It grows in generally open areas and is shade intolerant.

*1.2.4 Paper Birch (***Betula papyrifera** *Marsh.)*

Paper, white, silver, or canoe birch are all names commonly used for *Betula papyrifera* Marsh. Other scientific names used synonymously for *Betula papyrifera* are *B. cordifolia* Regel, *B. neoalaskana* Sarg., *B. neoalaskana* var. *kenaica* (W.H. Evans) Boivin, *B. alaskana* Sarg., and *B. papyrifera* var. *humilis* (Reg.) Fern. & Raup. Scoggan (1978), in the "Flora of Canada," separates the species into five varieties, while the more recent Flora of North America (2009a) has given some of these varieties the status of species. Paper birch is a member of the family Betulaceae, which includes the alders and hazelnuts. The birch species freely hybridize, causing frequent difficulty in identification of individual specimens (Flora of North America, 2009a).

Paper birch is a single- or multi-stemmed, *deciduous*, broad-leaved, hardwood tree, with simple, serrate, or *biserrate* alternate leaves about 4- to 7-cm long; and in mature trees, its usually white bark readily peels in sheets. It commonly grows to 21 to 24 m tall with 25 to 30 cm diameter trunks, but it can achieve 30 m in height, with trunk diameters of up to 75 cm. The flowers lack petals and occur in unisexual, pendulous spikes called *aments* or catkins, which occur in pairs on spur shoots. Both sexes of flowers occur on the same tree. Paper birch is a desirable firewood, and is used for furniture, flooring, veneer, plywood and oriented strand-board, pulpwood for paper, clothes pegs, and spools (Viereck and Little, 1972; Uchytil, 1991a; Flora of North America, 2009a). The sap has been used for syrup and the bark for baskets, storage containers, matts, baby carriers, snowshoes, and as a covering for the iconic birch bark canoe (Uchytil, 1991a; University of Michigan, 2003; Flora of North America, 2009a). Medicinally, birch served many functions for North American Aboriginal people (University of Michigan, 2003). It is an important winter food source for many large ungulates, small mammals and birds.

Paper birch generally does not live more than 140 years, but occasionally can exceed 200 years of age (Thorpe and Godwin, 1992; Ancient Forest Exploration and Research, n.d.). The trees are shade intolerant. Seed viability appears to vary by year, with germination measured at 77% during a heavy seed year and as low as 13% in a normal seed year (Bjorkbom, 1971). The period of greatest *seedfall* is between September and November (Bjorkbom, 1971). Seedfall was estimated at approximately 1 to 10×10^6 seeds per acre (0.405 hectares) in a Maine study area (Bjorkbom, 1971), but seed production as high as 36×10^6 seeds per acre has been reported in years of high production (Marquis, 1969). Flowering ranges from April in the southern portion of its range to as late as June in Alaska. Seed germination occurs best on disturbed mineral soil in the spring, with seedlings failing to grow on soils with a pH < 5.0. Near Sudbury, Ontario, however, paper birch has been found growing on acid soils of pH values as low as 3.2 (Watson *et al*., 1980). This observation is in agreement with laboratory tests in which paper birch grew well in all of the reference podzol soils with low pH (pH < 4) (EC, 2013b). Seed viability was found to vary between 49% and 65% for 3 different years of seed crops (Bjorkbom,

1971) while Watson *et al.* (1980) suggest viability percentages of 15 to 20%. Emergence of seed in laboratory tests ranged from 60 to 100% in tests that met the validity criteria (EC, 2013b). Seeds measure 2 × 2.5 mm, and can be stored for up to 2 years at room temperature and up to 10 years in the freezer (Watson *et al*., 1980). Seedlings can grow to a height of 5 to 12 cm during the first year. Seeds do not need *stratification* if germinated in the light, but stratification at 5°C for 60 to 70 days has been recommended (Watson *et al*., 1980). During method development testing, germination of paper birch in light was found to improve with stratification at 4 ± 2°C in moist peat for 4 weeks (SRC, 2004, 2012). Further improvements in emergence rates were observed when seeds were sorted using the ethanol floatation method, followed by hand-selection using a microscope lit from below (SRC, 2012; EC, 2013b). Non-stratified seeds germinate to higher percentages in light than in darkness; however, stratified seed germinated equally well in light and darkness (Bevington and Hoyle, 1981). Paper birch is a *dicotyledonous* plant. Asexual reproduction results following fire in the form of sprouts originating from stumps or the root collar. Paper birch shows a strong concentration-response relationship to boric acid (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011). In laboratory growth tests of 5 weeks, root lengths were longer than shoot lengths but shoot dry weight was greater than root dry weight (EC, 2013b). Leaf colour normally varies from green to red depending on soil and light conditions (see Figure 8).

Paper birch is found throughout the boreal forest region of North America (Alaska to Newfoundland) and extends southward from the boreal region as far as Oregon, Wyoming, Iowa, and Pennsylvania in the United States (Viereck and Little, 1972; USDI, 2006). It is used in reclamation and stabilization of severely disturbed sites and is important as an early colonizer of disturbed areas, including burned sites (Uchytil, 1991a). Paper birch is an early *seral*-stage species, seeding heavily on many sites following a fire, with gradual replacement by *conifers* throughout much of the boreal forest. In Labrador, birch stands begin to deteriorate at 75 to 100 years, with conifers becoming dominant by 125 years in the boreal mixed wood area.

*1.2.5 Jack Pine (*Pinus banksiana *Lamb.)*

Jack pine *(Pinus banksiana)*, a member of the pine family (Pinaceae), has also been called scrub pine, grey pine, Banksian pine, Hudson Bay pine, Banks pine, Eastern jack pine, black pine, black jack pine, and prince's pine (Carey, 1993; Earle, 2009b). It is a medium-sized *coniferous* tree (gymnosperm), with yellow-green needles (2 to 3.75 cm long) occurring in pairs rather than singly, as is the case with spruces and other non-pine coniferous evergreens native to Canada (Flora of North America, 2009b; Lakehead University, 2009). Jack pine employs *epigeal* germination, and has been used for pulpwood, lumber, telephone poles, fence posts, railroad ties, and is plantation-grown for Christmas trees (Carey, 1993). It was used by North American Aboriginal people for medicinal purposes. In addition, the sap was used for caulking and torches, the roots were used for sewing, the boughs were used for bedding, and smoke from the cones was used for tanning (University of Michigan, 2003). Jack pine is a food source for many small rodents as well as white-tailed deer, caribou, and snowshoe hares (Carey, 1993). Trees grow up to 27 m in height and 0.6 m in diameter. Seedlings grow to about 5 cm in the first year, and are 8 to 10–15 cm tall after two seasons, with roots 28 to 33 cm deep (Rudolph and Laidly, 1990). Open-grown trees are frequently crooked, but in dense stands, these trees grow straight (Flora of North America, 2009b). The bark is scaly at first, developing into scaly ridges as the tree matures. Female cones are curved upward along the branch. Cones generally open after many years on the tree or after fire. Jack pine's range overlaps that of lodgepole pine *(Pinus contorta)* in Alberta and British Columbia, where hybridization occurs (Rudolph and Laidly, 1990; Carey, 1993). Jack pine trees are *monoecious*.

The optimal temperature for jack pine germination is 25°C to 30°C (Qualtiere, 2008). Under light conditions, stratification reduced germination for some seed lots (Qualtiere, 2008), and germination is also reduced when light is excluded (Rudolph and Laidly, 1990). Dormancy of seeds ranged from non-dormant to 34% dormancy among 10 collections of jack pine from Saskatchewan (Qualtiere, 2008). Seed viability has been reported to be high with 1- to 6-year-old cones demonstrating 78 to 89% seed viability and 20-year-old seed demonstrating 50% viability (Carey, 1993). Emergence of seed in

laboratory tests ranged from 60 to 100% in tests that met the validity criteria (EC, 2013b). Seeds measure 2 × 3.5 mm. Following stratification of at least 2 weeks at 4 ± 2°C in moist peat, full emergence of seed in laboratory tests can be expected about 10 days after planting (SRC, 2006). Young seedling growth is characterized by production of a sturdy taproot (see Figure 9) approximately 2 to 3 times the length of the shoot (EC, 2013b), which grows primarily in number of leaves rather than length (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011). Jack pine demonstrates a strong concentration-response relationship to boric acid (SRC, 2006) and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012). In laboratory tests, jack pine tolerated soils with a wide range in pH: from 3.3 to 7.8 (EC, 2013b).

Jack pine height was found to be similar in western Quebec and Saskatchewan for the same-aged trees. Seedlings reach 1.3 m tall in 6 to 8 years, 6 m tall in 20 years, and 16 m tall in 60 years (Kabzems and Kirby, 1956). From a forestry perspective, jack pine matures at 70 to 80 years (Kabzems and Kirby, 1956), being one of the shortest-lived conifers of the boreal region. A 246-year-old tree found in Ontario (Earle, 2009b) may be the oldest documented tree. Jack pine develops a taproot, which in the first season will grow to a depth of 13 to 25 cm (Rudolph and Laidly, 1990). Trees may start to produce seed at 5 to 10 years of age (Carey, 1993). On well-drained soils the roots may penetrate to 2.7 m but the bulk of the root system is confined to approximately the upper 50 cm of the soil (Rudolph and Laidly, 1990).

Jack pine is most commonly found on dry, acidic, sandy soils, but is found on thin loam soil over bedrock, peat, and soil over permafrost (Carey, 1993). In general it occupies the less fertile and drier sites relative to other species that share its range.

Jack pine is the least shade-tolerant of the pines and only slightly more shade-tolerant than aspen, birch and tamarack. It rapidly colonizes after fire, forming evenly aged stands. The *serotinous* cones open in response to the heat from the fire, scattering the seed soon thereafter. It is only in the southernmost part of its range that the cones are non-*serotinous*. Jack pine seed germination (*epigeal* germination) is best on mineral soil, occurring within 15 to 60 days under favourable conditions. Most seedlings die if the organic soil layer is more than 1.3 cm thick (Rudolph and Laidly, 1990), and soil with < 0.5 cm thick humus provides the best seedbed (Carey, 1993). Jack pine is frequently replaced by other tree species with time, but on very dry sites such as deep sands, it may persist as the dominant species.

Jack pine occurs from the western Northwest Territories eastward through the boreal forest to eastern Quebec and into New Brunswick, Nova Scotia, and Prince Edward Island but does not extend into Newfoundland and Labrador. Its northern limit does not extend northwards to the tree-line along the northern edge of the boreal forest, and remains south of Hudson Bay in the East. It extends southwards into Minnesota, Wisconsin, Michigan, northern Illinois, and Maine (Flora of North America, 2009b).

*1.2.6 White Spruce [*Picea glauca *(Moench) Voss]*

Like jack pine, white spruce *(Picea glauca)* is a member of the pine family (Pinaceae), and has several other common names including Canadian spruce, eastern spruce, western white spruce, Black Hills spruce, skunk spruce, Alberta spruce, Porsild spruce, and cat spruce (Fowells, 1965; Earle, 2009a). It is a coniferous tree (gymnosperm) that employs *epigeal* germination. White spruce is an important source of pulpwood and lumber but is also used for the construction of log homes, musical instruments, paddles, and boxes. It was used by numerous Aboriginal groups for medicinal purposes, for construction of lodging, canoes, baskets, snowshoes, and bedding, as well as for food and beverages (University of Michigan, 2003). Being a dominant boreal forest species, white spruce is an important food source, and provides vital habitat to many wildlife species, including mammals, birds and invertebrates (Zasada *et al*., 1978; Nienstaedt and Zasada, 1990). A tall-growing conifer, white spruce can reach 40 m in height on favourable sites with diameters of up to 120 cm (Watson *et al*., 1980). In the southern boreal forest in the interior of the continent, white spruce growing on good sites average about 27 m in height at 100 years of age; however, at the same sites, 190-year-old trees are only 31 m in height (Kabzems, 1971). On poor sites, 100-year-old trees will be < 20 m tall. The leaves (needles) are green, four-sided, stiff, and sharp-pointed, and are 8 to 18 mm long (Hosie, 1979).

Commonly, white spruce is shallowly rooted (90 to 120 cm in depth) as a result of site conditions limiting root penetration; however, taproots can descend to 3 m. Large roots are frequently within 15 cm of the organic/mineral soil boundary (Nienstaedt and Zasada, 1990).

White spruce grows on a wide range of soils and site types. It is found on glacial, *lacustrine*, marine, and *alluvial* deposits, of sandy to clayey texture; however, it is usually a minor species on sandy soils. It does best on well to moderately well-drained soils and on moist, *alluvial* soils along streams (Nienstaedt and Zasada, 1990). In the northwest portion of its range (the Canadian territories and Alaska), white spruce is largely restricted to river floodplains and south-facing upland sites. White spruce grows under both acid and alkaline conditions but does best where pH ranges from 4.7 to 7 or higher (Nienstaedt and Zasada, 1990). In laboratory tests, white spruce tolerated soils with a wide range in pH from 3.2 to 7.5 (EC, 2013b). White spruce shows a strong concentration-response relationship to boric acid (SRC, 2006), metals (Wren *et al.,* 2012) and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012). In laboratory tests, germination of white spruce begins slowly, and emergence reaches its maximum within 18 days when planted at a depth of 3 mm (SRC, 2006). Germination was ≥ 80% in the majority of tests in artificial and reference soils but ranged from 53 to 96% in all tests that met the validity criteria (EC, 2013b). Young seedling growth is characterized by production of a sturdy tap root (see Figure 10) approximately double the length of the shoot (EC, 2013b), which grows primarily in number of leaves rather than length (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011).

Both sexes of cones occur on a single tree. Pollination occurs from May to as late as July, depending on geographic location and yearly weather conditions, with variation by as much as 4 weeks from year to year at the same locality (Nienstaedt and Zasada, 1990). Embryo development is completed in August, with maximum embryo development occurring after 635 *growing degree days* based on a 5°C threshold temperature in Alaska (Zasada *et al*., 1978). Cones ripen in August or September. Seeds measure 1.5 × 3 mm and can be collected 2 to 4 weeks before they ripen with seed quality improving in cool (4°C to 10°C), ventilated storage conditions. By age 30, trees are producing large quantities of seed. Good seed crops occur from 2 to 12 years apart, with good crops followed by a poor seed-producing year. Most seed dispersal occurs during dry weather in September. The optimum germination temperature is between 10°C and 24°C, with little germination below 10°C. *Stratification* at 2°C to 4°C is recommended for seed testing, but is not a prerequisite for germination (*epigeal* germination). In laboratory tests, stratification of at least 3 weeks at 4 ± 2°C in moist peat improved overall level and synchronization of germination (SRC, 2004). Cold storage of seeds may be successful for up to 10 years and for 2 years at room temperature (Watson *et al*., 1980). Forest soil *L* and *F* layers of > 5 to 8 cm restrict regeneration, especially in western North America with its generally drier climate. There are approximately 500 000 seeds/kg (Watson *et al*., 1980).

Optimum day/night temperatures for growing seedlings are 25°C/20°C at 400 lm/m^2 (Kabzems, 1971). However, light levels and temperature effects interact, with lower light levels requiring different temperature regimes for maximum seedling growth.

The longest-living trees occur at the cold climatic margins of the range, in both latitude and altitude. A tree on the Mackenzie Delta had 589 growth rings, and trees approaching 1000 years have been found above the Arctic Circle (Nienstaedt and Zasada, 1990). In general, on good sites, trees up to 250 years old are common.

In Canada, white spruce is found in all provinces and territories, ranging from the Atlantic Ocean in all the eastern provinces, almost to the Pacific coast in British Columbia and extending to the tree line along the tundra edge (Kabzems, 1971). Its southern range extends to Minnesota, Wisconsin, Michigan, New York, and Maine in the east and into isolated uplands in the west, such as the Cypress Hills of Saskatchewan, the Black Hills of Wyoming and South Dakota, and into the Rocky Mountains in Montana (Nienstaedt and Zasada, 1990).

*1.2.7 Black Spruce [***Picea mariana** *(Mill.) Britton, Sterns & Poggenb.]*

Black spruce *(Picea mariana)* is also a member of the pine family (Pinaceae), and has been called bog spruce, swamp spruce, and short-leaf spruce. It is frequently a small *coniferous* tree (gymnosperm), with a narrow crown. It employs *epigeal* germination. Black spruce characteristically has a dense grouping of branches at the top, giving the tree a clubbed appearance. Needles are short (approximately 12 mm) with four sides. It has greyish-brown bark, which when flaked off shows a yellowish tinge to the under-bark. Black spruce makes high-quality pulp and is used for lumber and Christmas trees (Viereck and Johnston, 1990). Like the white spruce, it was used widely by Aboriginal people across the North for a variety of purposes. Its stands provide important breeding habitat for several boreal nesting songbirds. Many species of birds and rodents feed on the seeds, whereas needles, bark and twigs form a major component of the snowshoe hare's diet in the winter (Viereck and Johnston, 1990; Rook, 2002).

Black spruce has an average maximum age of 200 years but can live to 280 years (Viereck and Johnston, 1990). Average height is 12 to 20 m on good sites at maturity, but black spruce can reach a height of 27 m in the Ontario Clay belt. Roots are concentrated in the upper 20 cm of the organic *horizons*. At the northern extent of its range, black spruce may be 3 to 6 m tall at ages of 100 to 200 years (Uchytil, 1991b). Black spruce is shade tolerant. It is a pioneer species after fire, producing seed at an early age with increased dispersal after the cones have been exposed to fire (Uchytil, 1991b).

Black spruce is most commonly found on organic soils in the southern range of the boreal forest as well as in the North. It also occurs as a widespread upland species across the northern boreal region on thin soils over bedrock, but also on sands through clay soils. It is better adapted to permafrost soils than other trees because of its shallow rooting habit. It occurs commonly as pure stands on organic soils, but is often in mixed stands with other conifers or even broad-leaved trees on mineral soils. The most productive stands occur on the better-drained sites (Viereck and Johnston, 1990).

The optimal germination temperature for black spruce is 20°C to 25°C (Qualtiere, 2008). Qualtiere (2008) tested 12 provenances of black spruce from Saskatchewan, and found that *stratification* did not improve germination. This is in contrast to another study by Baskin and Baskin (2002), in which cold moist stratification was used for 24 days followed by seeding at 20°C/10°C day/night temperatures for germination. In laboratory tests, stratification of at least 3 weeks at 4 ± 2°C in moist peat improved both overall level and synchronization of germination (SRC, 2004). Light did not appear to have an impact on germination, however, and at low temperatures (i.e., 10°C), long *photoperiods* (i.e., 22 h) were needed for germination (Qualtiere, 2008). In Qualtiere's study, germination rates ranged between 70 and 100% for all collections, whereas another study found seed viability to be only 53% for 1- to 5-year-old seed from northeast Ontario (Uchytil, 1991b). Seed germination was found to remain high for 12 years in stored seed and then declined rapidly (Viereck and Johnston, 1990). Black spruce seeds are the smallest of any North American spruce (1.5 × 2.5 mm). Seedling growth is slow, approaching 2.5 cm in the first year (Uchytil, 1991b).

In laboratory tests, black spruce tolerated soils with a wide range in pH from 3.2 to 7.6 (EC, 2013b). Black spruce shows a strong concentration-response relationship to boric acid (SRC, 2006), metals (Wren *et al.,* 2012), and weathered petroleum hydrocarbons and salts (Princz *et al*., 2012). In laboratory tests, germination of black spruce begins slowly, and emergence reaches its maximum within 16 days when planted at a depth of 3 mm (SRC, 2006). Germination was ≥ 80% in the majority of tests in artificial and reference soils but ranged from 60 to 100% in all tests that met the validity criteria (EC, 2013b). Young seedling growth is characterized by production of a sturdy tap root (see Figure 11) approximately double the length of the shoot (EC, 2013b), which grows primarily in number of leaves rather than length (M. Moody, personal communication, Saskatchewan Research Council, Saskatoon, SK, 2011).

Black spruce ranges from the west coast of Alaska to the northern Labrador coast, reaching tree-line across much of the continent. In the south it extends to central British Columbia, central Alberta and Saskatchewan, southern Manitoba, Minnesota,

Wisconsin, and Michigan, and Pennsylvania to northern New Jersey (Viereck and Johnston, 1990).

1.3 Historical Use of Boreal Plants in Toxicity Tests

The history of the development of biological test methods for soil toxicity testing and the use of whole-soil toxicity tests to characterize the effects of *toxicants* on plants is covered in detail in Environment Canada's Test for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil (EC, 2005a).

Plants are widely used as test organisms in single-species toxicity tests intended to measure the toxicity of pure chemicals, chemical products, or samples of soil contaminated or potentially contaminated with chemicals in the field or (for experimental purposes) in the laboratory. In Canada, results of soil toxicity tests are used to:

(1) derive national soil quality criteria,

(2) establish *site*-specific, *risk*-based, cleanup objectives (e.g., *remediation* targets), and

(3) assess the efficacy of remediation technologies (Stephenson *et al*., 2002).

Extensive reviews have been carried out on the use of plant toxicity tests as "ecological assessment tools" for appraising the toxicity of contaminated or potentially contaminated soils (Wang, 1991, 1992; Wang and Freemark, 1995; Kaputska, 1997; Meier *et al*., 1997; Saterbak *et al*., 1999). In some cases, standard methods have been modified or unique methods have been developed in order to obtain relevant data (Pfleeger *et al*., 1991; Sheppard, 1994; Chaineau *et al*., 1997). Database reviews have been summarized in reports discussing trends of plant toxicity to various contaminants (Kenaga, 1981; Miller *et al*., 1985; Boutin and Rogers, 2000). *Toxic* effects of plant exposure to contaminated soils have been documented in laboratory studies involving samples of soil spiked or contaminated with:

- pesticides (Fletcher *et al*., 1995, 1996; Boutin *et al*., 2000, 2004);
- metals (Godbold and Hüttermann, 1985; Kaputska *et al*., 1995; Kjaer and Elmegaard, 1996; Rader *et al*., 1997; Kjaer *et al*., 1998; Redente *et al*., 2002; Lock and Janssen, 2003; Feisthauer *et al*., 2006; Menzie *et al*., 2008; Sheppard and Stephenson, 2012);
- petroleum hydrocarbons (Chaineau *et al.,* 1997; Wong *et al*., 1999; Cermak *et al*., 2010; Angell *et al*., 2012); and
- other chemicals (Siciliano *et al*., 1997; Kalsch and Römbke, 1999; Velicogna *et al*., 2012).

Various plant species have been recommended for *phytotoxicity* testing by different agencies (OECD, 2006a; USEPA, 1989; ISO, 2012a, b; ASTM, 2009; EC, 2005a). The test species most commonly recommended among agencies include: lettuce, cabbage, cucumber, soybean, oat, perennial ryegrass, corn, tomato, rice, and carrot. Fletcher *et al*. (1985, 1988) reviewed the PHYTOTOX database and provided a summary of the most commonly used terrestrial plants. These plants included wheat, pea, tomato, oats, beans, apple, soybean, corn, and barley; they are most relevant to arable soils and regions. However, within Canada, the boreal zone covers over 50% of the landscape and is home to thousands of *sites* potentially impacted by forestry, mining, oil and gas and power generation industries CCME, 2008). Federal (CCME, 1994, 1996, 1997) and provincial (AE 2007 a, b, c) frameworks for *ecological risk assessment* (ERA) often rely on results of site-specific single-species toxicity tests in higher Tiers of risk assessment. Therefore, the use of a *test battery* utilizing species native to the boreal ecozones of Canada supports ecologically realistic site-specific ecological risk assessments and *remediation* programs (Princz *et al*., 2012).

The effects of salinity on boreal forest plants have been investigated to support reclamation of salt-impacted lands (Hettinger, 1982; Howat, 2000), particularly areas of oil sands development (Renault *et al*., 2000; Croser *et al*., 2001; Redfield, 2001; Franklin *et al*., 2002, Khasa *et al*., 2004). Toxicity testing using non-crop species, including boreal species, is increasing as risk of pesticide exposure is recognized and toxicity tests are developed (White *et al*., 2009).

Non-target and wild plant species have been used to measure aspects of variability in phytotoxicty testing such as ecotype and abiotic factors (Boutin *et al*., 2010) and the effects of pesticides (Dalton and Boutin, 2010). These woody plants of the boreal

region included in this standardized method (jack pine, white spruce, trembling aspen) have demonstrated genetic variability in salt tolerance (Khasa *et al*., 2002).

Environment Canada's standardized biological test method for performing a test that measures the *toxic* effects of prolonged exposure to *chemical-spiked soil* or *site soil* on the emergence and growth of terrestrial plants (EC, 2005a) provided the basis for methodology of testing using boreal forest plants, the new method described herein. The use of boreal forest species brought about changes to this test method in a number of areas including the requirement for stratification of seed, longer test durations and testing of soils in horizons, and use of wild or field-collected seed. The *emergence* endpoint has been removed due to insensitivity relative to the *growth* endpoints.

Test Organisms

2.1 Species

Test organisms to be used in this biological test method must be selected from the specified group of seven boreal forest plant species listed here. Acceptable choices for plant species include three *angiosperm dicotyledons*, one *angiosperm monocotyledon,* and three *gymnosperms*:

angiosperm dicotyledons

- trembling aspen (*Populus tremuloides*)
- Canada goldenrod (*Solidago canadensis*)
- paper birch (*Betula papyrifera*)

angiosperm monocotyledon

- bluejoint reedgrass (*Calamagrostis canadensis*)

gymnosperms

- white spruce (*Picea glauca*)
- black spruce (*Picea mariana*)
- jack pine (*Pinus banksiana*)

The selection of multiple test species should depend on the species' sensitivity to the substance or material being tested and to the pH of the test soil, if known. The number and type of species selected for comprehensive studies (i.e., multi-species effects) will depend on the purpose of the study and the regulatory requirements under which the tests are being conducted. The role of ecological relevance of test species selected will vary depending on the study objectives.[3] The ecological, economical, and physiological significance of the seven boreal plant species to be used in this biological test method are summarized in Section 1.2.

The test must be started using wild-collected *seed* (i.e., seed that is not selected for superior characteristics in a nursery or plantation) for all species except bluejoint reedgrass, the seed of which may be purchased from suppliers cultivating plants for reclamation purposes.[4] The plant seed used for testing must be from the same lot number for each of the individual plant species, and should be free of fungicides, insecticides, repellents, or other chemical pretreatment agents (i.e., *seed pretreatment*). Details on seed size, purchasing, *stratification*, storing, and condition are provided in Table 2 and in Sections 2.2, 2.3, 2.4 and 2.5.

Each plant species has unique characteristics that affect its performance in a toxicity test; therefore, certain test procedures and conditions (i.e., number of seeds used to initiate a test, test duration and test validity criteria) are modified on a species-specific basis to accommodate these requirements (see Sections 4.2, 4.3 and 4.4).

[3] Criteria to consider when selecting species for a *test battery* include: relevance to surrounding vegetation and habitat of site soils; phylogeny, evaluation of response to reference toxicant (shoot/root length and shoot/root mass); performance and sensitivity in *range-finding* tests (i.e., % emergence in *control*, test duration, ease of root separation, sufficient *biomass* at end of test, time to emergence, effect of soil on growth); type of *germination (epigeal/hypogeal)*; inclusion of monocotyledonous and/or dicotyledonous species; inclusion of *angiosperms* and/or *gymnosperms*; source availability and quality of seed; type of root formation; phenology and life history traits; and critical variable requirements (e.g., pH, nutrients) (ESG, 2002; SRC, 2003).

[4] Seed should be purchased from reputable suppliers that can reliably identify and obtain seeds of native, wild plants (or from cultivated plants for bluejoint reedgrass only). Suppliers should identify the geographical area in which seed was collected. Laboratories may collect wild seed if identification of the source plant(s) is assured by consultation with a trained plant ecologist or taxonomist. It is difficult to find commercial sources of viable trembling aspen seed, partly because the seed is ephemeral, i.e., viable for a short period of time and because commercial suppliers do not have facilities to keep seed frozen after collection. However, trembling aspen trees are widespread in Canada, and seed may be collected and cleaned according to Day *et al.* (2003). There are many commercial and government suppliers of white spruce, black spruce, jack pine, paper birch, and goldenrod seed that are skilled in collection, extraction, and storage of seed. It is recommended that they be used.

Table 2 Seed characteristics, source, and stratification

Plant	Seed size (mm)	Seeds/g	Stratification (min to max)[a]	Seedling emergence[b] (min)	Seed suppliers[c] (commercial and government)
Trembling aspen	0.5 × 1	5500–8000	none, keep seed frozen at -15°C to -30°C	4 d	ATISC, BCMF NTSC, SRC
Bluejoint reedgrass	0.25 × 1	5000–8000	2 wk to 6 mo	7 d	Acorus Restoration BrettYoung Ontario Seed Co. Sheffield's Seed Co.
Canada goldenrod	0.25 × 1	2000–4000	3 wk to 3 mo	4 d	Acorus Restoration ALCLA Native Plants Bedrock Seed Bank Gardens North
Paper birch	2 × 2.5	1700–5200	4 wk to 4 mo	7 d	BCMF, NTSC, Gardens North Jeffries Nurseries Ltd.
Jack pine	2 × 3.5	200–300	2 wk to 6 mo	10 d	ATISC, MFB OTSP, YPP
White spruce	1.5 × 3	300–600	3 wk to 4 mo	18 d	BCMF, MFB OTSP, YPP
Black spruce	1.5 × 2.5	400–1000	3 wk to 4 mo	16 d	BCMF, MFB, OTSP, YPP Sheffield's Seed Co.

[a] Minimum and maximum recommended duration of stratification.
[b] Estimated minimum number of days to full emergence of seed planted in artificial soil.
[c] See Section 2.2 for seed suppliers' contact information.

2.2 Source

Seeds used to initiate a soil toxicity test should be obtained from commercial seed companies or government seed banks. Certification of wild plant seed is not generally available, but the supplier should provide a statement that the seed has not been pretreated with any substances. The following information should be obtained when purchasing seed:

- species (Latin and common names),
- year of collection,
- geographic location where seed was collected,
- packet size (g or kg),
- lot number,
- rating for % germination,
- date of germination rating,
- date of purchase,
- shelf life, and
- name of supplier.

The date the seed package is opened should also be recorded. Seed should generally be purchased at least annually,[5] preferably within one year of the date of the seed germination rating; however, a given lot of seed may be used as long as the seed can meet the *control* performance criteria described herein (see Section 4.4), and provided that the sensitivity of the seed does not change significantly over time

[5] This recommendation is only a general guideline since seed viability will vary from year to year. It might be advantageous in certain instances, therefore, to use older seed with a relatively high (known) emergence rate versus newly purchased seed with a lower (and/or unknown) emergence rate, and to acquire seed from several different suppliers. Seed availability from seed suppliers may also vary from year to year.

as determined by the results of *reference toxicity tests* (see Section 4.9).

Commercial sources that have been used to secure quality seed for toxicity testing include:

Acorus Restoration
#722 6th Con. Rd. RR #1
Walsingham ON N0E 1X0
Tel.: 519-586-2603
Fax: 519-586-2447
Website: www.ecologyart.com
Email: info@ecologyart.com
Species: bluejoint reedgrass, Canada goldenrod

ALCLA Native Plants
3208 Bearspaw Dr. NW
Calgary AB T2L 1T2
Tel./Fax: 403-282-6516
Website: www.ALCLAnativeplants.com
Email: ALCLA@telus.net
Species: Canada goldenrod

Bedrock Environmental Services Ltd and Bedrock Seed Bank
P.O. Box 608
Sangudo AB T0E 2A0
Tel.: 780-448-1722
Website: www.bedrockseedbank.com
Email: bedrock@telusplanet.net
Species: Canada goldenrod

BrettYoung
P.O. Box 99
St. Norbert Postal Station
Winnipeg MB R3V 1L5
Tel.: 204-992-7129
Fax: 204-478-2232
Website: www.brettyoung.ca
Species: bluejoint reedgrass

Gardens North
P.O. Box 370
Annapolis Royal NS B0S 1A0
Fax: 902-532-7949
Website: www.gardensnorth.com
Email: seed@gardensnorth.com
Species: Canada goldenrod, paper birch

Jeffries Nurseries Ltd.
P.O. Box 402
Portage la Prairie MB R1N 3B7
Tel.: 204-857-5288
Fax: 204-857-2877
Website: www.jeffriesnurseries.com
Email: jeffnurs@mts.net
Species: paper birch

Ontario Seed Company (OSC Seeds)
P.O. Box 7
Waterloo ON N2J 3Z6
Tel.: 519-886-0557
Fax: 519-886-0605
Email: seeds@oscseeds.com
Species: bluejoint reedgrass

Sheffield's Seed Company Inc.
269 Auburn Rd, Route 34
Locke, New York 13092 USA
Tel.: 315-497-1058
Fax: 315-497-1059
Website: www.sheffields.com
Email: seed@sheffields.com
Species: black spruce, bluejoint reedgrass

Yellow Point Propagation (YPP)
P.O. Box 669, 13735 Quennell Rd.
Ladysmith BC V9G 1A5
Website: www.yellowpointpropagation.com
Email: ypprop@shaw.ca
Species: jack pine, white spruce, black spruce

The following government sources have been used to obtain quality seed during the development of this test method document (government sources might supply seed for research purposes only):

Alberta Tree Improvement and Seed Centre (ATISC)
Alberta Sustainable Resource Development
P.O. Box 750
Smoky Lake AB T0A 3C0
Tel.: 780-656-5073,
Fax: 780-656-2120
Email: donna.palamarek@gov.ab.ca
Species: trembling aspen, jack pine

British Columbia Ministry of Forests (BCMF)
BCMF Tree Seed Centre
Tree Improvement Branch
Tel.: 604-541-1683 ext. 2228
Fax: 604-541-1685
Website: www.for.gov.bc.ca/hti/treeseedcentre/index.htm
Species: trembling aspen, white spruce, black spruce, paper birch

Manitoba Forestry Branch (MFB)
(seed supplied through Pineland Forest Nursery)
P.O. Box 45
Hadashville MB R0E 0X0
Tel.: 204-426-3235
Fax: 204-426-2106
Website: www.pinelandforestnursery.com
Email: trevor.stanley@gov.mb.ca
Species: white spruce, black spruce, jack pine

National Tree Seed Centre (NTSC)
Natural Resources Canada
Canadian Forest Service – Atlantic
P.O. Box 4000
Fredericton NB E3B 5P7
Tel.: 506-452-3530
Fax: 506-452-3525
Website: http://cfs.nrcan.gc.ca/subsite/seedcentre
Email: dale.simpson@nrcan.gc.ca
Species: trembling aspen, paper birch (for research purposes)

Ontario Tree Seed Plant (OTSP)
Ontario Ministry of Natural Resources
Tel.: 877-861-8881, 705-424-5311
Fax: 705-424-9282
Website: www.mnr.gov.on.ca/en/Business/Forests/2ColumnSubPage/STEL02_166052.html
Email: ontariotreeseed@ontario.ca
Species: jack pine, white spruce, black spruce

Saskatchewan Research Council (SRC)
125-15 Innovation Blvd.
Saskatoon SK S7N 2X8
Tel.: 306-933-5400
Fax: 306-933-7817
Website: www.src.sk.ca
Email: info@src.sk.ca
Species: trembling aspen

A current list of seed suppliers can be obtained by contacting:

Method Development and Applications Unit
Science and Technology Branch
Environment Canada
335 River Road
Ottawa ON K1A 0H3
Email: methods@ec.gc.ca

2.3 Seed Stratification and Selection

Prior to toxicity testing, the seed of bluejoint reedgrass, Canada goldenrod, paper birch, white spruce, black spruce, and jack pine must be stratified.[6] *Stratification* is the process in which moisture is imbibed into the seed to overcome physiological dormancy, improve overall germination, and synchronize germination.

A method of stratification proven to be successful during the development of this method is as follows. Seeds are stratified by refrigerating an aliquot of dry seed in moist peat moss at 4 ± 2°C. The recommended duration for stratification of seed (i.e., minimum and maximum durations; see Table 2) varies by species and depends somewhat upon moisture content and age of the seed before stratification. When stored and stratified properly, seed that is several years old can germinate strongly. The weight of seed to be stratified can be calculated on the basis of expected needs for testing and the number of seeds per gram (see Table 2).

Paper birch seed should be sorted prior to stratification using an ethanol-floatation technique to separate empty, insect-damaged, and defective seeds from the plump, fertile seeds required for toxicity testing (Bjorkroth, 1973; Simpson and Daigle, 2011).[7] This method is recommended to improve emergence rates and is described in detail in Appendix E. In addition, removing the wings of paper birch seeds prior to ethanol separation may further improve emergence rates.[8]

Stratified seed may eventually germinate in the refrigerated peat moss, and fungal growth may be observed. Germination of this seed will decline, and it must be discarded if fungal growth appears or if ≥ 10% of the seed has germinated. Prior to testing, germination of stratified seed should be assessed.[9] It is recommended that stratified seed be assessed periodically to ensure that good-quality stratified seed is ready for testing. Since trembling aspen seed is not stratified prior to planting, seed germination

[6] Dry, sieved peat moss is placed in a zippered plastic bag, moistened, and mixed after the addition of sufficient de-ionized water. The peat is then squeezed until no water drips from it. Sufficient seeds are added, the mixture of seeds and peat shaken, and the bag rolled closed to eliminate air. The bag is then refrigerated for at least the minimum time suggested in Table 2.

[7] Some seed suppliers (i.e., British Columbia Ministry of Forests) might pre-sort paper birch seed using the ethanol floatation method. If seed is pre-sorted when purchased, no further ethanol separation is required.

[8] Low germination and emergence has been observed for paper birch by both the Saskatchewan Research Centre (SRC) and Environment Canada's Soil Toxicology Laboratory (STL) during the development of this method. As a result, SRC and STL undertook investigations to improve paper birch germination and emergence rates in soils (SRC, 2012; EC, 2013b). Results of these investigations indicated that de-winging and ethanol separation of paper birch seed, followed by a four-week stratification, and finally microscopic selection of healthy seed (illuminated from below) improved the emergence rates of paper birch (see Appendix E).

[9] The germination of stratified seed can be assessed by placing 10 to 20 seeds on moist filter paper, underlain with layer(s) of moistened glass-fibre filter paper, all inside of a Petri dish. The edges of the Petri dish are then sealed with sealing film and incubated under test conditions for up to three weeks. Seed germination is assessed weekly.

may be assessed by removing a small aliquot of the seed (10 to 20 seeds) from the freezer and allowing them to germinate on moist filter paper, under light, as described for stratified seed.

Seed is selected for toxicity testing by spreading an aliquot of stratified seed (and peat moss for all species except aspen) on a Petri dish and sorting the seed from the peat moss under low magnification of a stereomicroscope. It is important to select seed that is firm, regular in size, unblemished and undamaged, and showing no sign of germination. Seed batches of bluejoint reedgrass, Canada goldenrod and paper birch may contain vegetative debris and empty *hulls* and therefore should be selected using microscopic observation when lit from below. Microscopic selection also allows the selection of healthy, plump embryos that are more likely to be fertile. Seeds should be gently probed to distinguish between firm, good-quality seed and empty hulls and soft, degraded seed. It is normal (not harmful) to observe split seed coats in stratified seed of white spruce or black spruce. It is important that batches of stratified seed be removed from refrigeration for as short a period of time as possible to reduce the likelihood of premature germination.

2.4 Seed Storage

Purchased or collected seeds (i.e., prior to *stratification*) should be kept in their original packages and stored in the dark, in labelled, sealed containers (e.g., zip-lock bags). Aspen seed must be frozen immediately after collection and cleaning and stored at -15°C to -30°C. Dry seed of white spruce, black spruce, jack pine, bluejoint reedgrass, Canada goldenrod, and paper birch may be stored in the refrigerator (i.e., 4 ± 2°C) or freezer (i.e., -15°C to -30°C)[10] until an aliquot is stratified (in the refrigerator at 4 ± 2°C) for use in tests. The day of test initiation (Day 0), a portion of the stratified seed must be removed from the refrigerator (or freezer in the case of aspen) and brought to room temperature (for about 10 to 15 minutes) prior to use in the test.

2.5 Seed Condition

The sensitivity of each new lot of seed used in a *definitive* soil toxicity test must be measured using a 14-, 21-, 28-, or 35-day (i.e., depending on the species) *reference toxicity test* (see Section 4.9). Ideally, a reference toxicity test should be performed together with each definitive soil toxicity test.

However, laboratories routinely undertaking soil toxicity tests may choose instead to conduct routine reference toxicity tests (i.e., at least once every two months) using a portion of each lot of seed used for definitive soil toxicity tests. All tests with a *reference toxicant*(s) should be performed using the conditions and procedures outlined in Section 4.9. Species-specific, test-related criteria used to judge the validity of a particular definitive soil toxicity test (and, indirectly, the condition of the lot of seed used in the test), based on the performance of the test organisms in the negative control soil, are given in Section 4.4.

[10] Seed of white spruce, black spruce, jack pine, bluejoint, goldenrod, and paper birch can be safely frozen as long as the moisture content of the seed is < 10%. The National Tree Seed Centre (NTSC) uses 8% as the upper limit of seed moisture content. Frozen storage is best; continually bringing seed out of frozen storage, warming it to room temperature and putting it back in storage only to repeat the procedure again and again when more seeds are needed stresses the seed and will eventually kill it. The NTSC suggests that researchers place small quantities of seed in small vials, freeze all the vials and remove a vial or two at a time when seed is needed (D. Simpson, personal communication, National Tree Seed Centre, Fredericton, NB, 2011).

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).[11] 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 also be properly ventilated.

The test facility should be isolated from the area 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 sample handling and storage facilities, or those where chemicals are processed or tested, from contaminating the area of the laboratory where tests are conducted.

The mean air temperature of the test facility must be maintained at 24 ± 3°C. However, for those test facilities with growth chambers that can accommodate a cyclical temperature change, it is recommended that the temperature be maintained at a mean air temperature of 24 ± 3°C during the day and 15 ± 3°C at night. The relative humidity of the test facility should be maintained at a minimum of 50%.[12]

Any construction materials that might contact the organisms, water, or test vessels within this facility must be nontoxic 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 and pH) of the *test soil* and associated *test (hydration) water*. Additionally, the laboratory should be equipped to facilitate prompt and accurate analysis of the *moisture content* of test soils. Equipment requirements include a drying oven that can be set at 90ºC for drying test organisms and 105ºC for drying soils, a weighing balance accurate to the nearest 0.1 mg, a light meter, 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*, *test (hydration) water*, *stock solutions*, or test solutions must be clean and rinsed with *de-ionized* or *distilled* water (i.e., *test water*) before being used. All non-disposable materials should be washed after use. The following cleaning procedure is recommended (EC, 1997a, b, 2004a, 2005a, 2007a):

(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[13]) nitric (HNO_3) or hydrochloric acid (HCl) (metal-free grade) to remove scale, metals, and bases;
(4) rinse twice with de-ionized water (or other *test water*);

[11] Greenhouses are not considered acceptable alternatives for test facilities since they are typically too variable in terms of controlling light, temperature, and humidity conditions.

[12] The relative humidity of the test facility will not greatly affect the frequency of soil hydration required throughout the test as it is recommended that lids remain in place for the full duration of the test.

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

(5) rinse once with full-strength, pesticide-grade acetone to remove organic compounds and with reagent-grade (e.g., HPLC grade, ≥ 98.5% purity) hexane for oily residues (use a fume hood);[14]

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

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

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

3.2 Initial and Definitive Tests

3.2.1 Initial Tests

Before *definitive* plant 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 a *negative control soil* (see Section 3.4) 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 each test species can be achieved in the laboratory in a candidate natural or artificial negative control soil using the procedures specified in this report.

The conditions and procedures used to perform these initial tests with *negative control soil* should be identical and according to Section 4. The conditions and procedures used to perform these initial tests with one or more *reference toxicant*s should be identical and according to Section 4.9. Each set of initial tests with *negative control soil* or *reference toxicant(s)* should be performed on each boreal plant species intended for use in future definitive toxicity tests.

Data from the control performance tests ($n \geq 5$) must show that the criteria for test validity (see Section 4.4) can be met for the intended test species using a natural or *artificial soil* intended for use as negative control soil in a definitive soil toxicity test. Data from the initial reference toxicity tests ($n \geq 5$) 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 *seedling* growth for the duration of the test. It is important that the size, shape, colour, and composition of the vessel be appropriate for the plant species chosen. The vessels should have a sealable lid and should not interfere with light quality within. The test vessel recommended for growth tests with boreal forest plants described herein is a 1-L clear polypropylene container, with a clear polypropylene lid.[15] Alternatively, 1-L glass jars sealed with transparent lids (e.g., hinged glass lids) may be used.[16] The vessels should be covered for the full test duration. If a polypropylene container and lid is used, five holes may be made in the lid using a push-pin to reduce condensation. If plants reach the top of the container, the lid should be replaced with an inverted test unit or other suitable transparent container, taped in place, thereby allowing headspace for further plant growth.

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

[15] Advantages of using the 1-L polypropylene vessels include: 1) they do not need to be washed and are disposable; 2) they are clear so that water pooling on the bottom of the containers (i.e., soil saturation has been exceeded) can be observed when watering from above; 3) they are relatively inexpensive and readily available; 4) they can accommodate from 300 mL to close to 1 L of soil without changing the ratio of surface area to volume substantially, because of their tapered shape; 5) they come in a variety of sizes (e.g., 500 mL and 1000 mL) and have snap-on lids that are easily removed and replaced for watering; 6) the lids do not change substantially or interfere with light fluence; and 7) phytotoxic observations can be made without removing the lids.

[16] The 1-L glass jars proved to be a successful alternative in a study where a highly volatile test substance reacted with the 1-L polypropylene vessels, thereby compromising their integrity (Stephenson *et al*., 2001).

3.3 Lighting

Test vessels should be illuminated using full-spectrum fluorescent or equivalent lighting. The light fluence rate, measured adjacent to the level of the soil surface, must be 300 ± 100 µmol/(m^2 · s) (i.e., equivalent to 18 750 ± 6250 *lux*). Since light intensity tends to vary in a given space, it should be measured at several points within the testing area. The light fluence rate within the testing area should not vary by more than ± 15% of the selected light fluence rate.[17]

3.4 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 plants during the test. The use of negative control soil provides a measure of test acceptability, evidence of the health and performance of the test organisms, assurance as to the suitability of the test conditions and procedures, and a basis for interpreting data derived from the *test soils*.

A soil toxicity test may use *clean* (uncontaminated) natural soil and/or *artificial soil* as the negative control soil. The selection of an appropriate negative control soil depends on considerations such as the study design, physicochemical characteristics of the *test soil*(s), and the availability of suitable *clean* natural soil with acceptable properties.[18] Uncontaminated natural soil should be used as the negative control soil for *definitive* tests with field-collected soil. There should also be prior experimental evidence that the soil chosen for use as negative control soil will sustain seedling *emergence* and growth that consistently and reliably meet the criteria for test validity defined herein (see Section 4.4).

The biological test method described herein has been developed and tested using nine negative control soils with diverse physicochemical characteristics (SRC, 2009; EC 2010, 2013b). These *clean* soils included one *artificial soil* and eight natural boreal forest soils (i.e., Gleyed Humo-ferric Podzols from Newfoundland, New Brunswick, and Ontario; a Dark Grey Luvisol, an Ortho Eutric Brunisol, and Eluviated Dystric Brunisol from Saskatchewan; and a Rego Humic Gleysol and Rego Dark Grey Chernozem 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 were therefore considered to be "*clean*." The origin and physicochemical characteristics of these natural soils are further described in Appendix F. The test validity criteria for the various plant species described in Section 4.4 are based on the performance data for these species in negative control soil that were generated for each of these nine diverse soils (SRC, 2004, 2006, 2007, 2008, 2009; EC and SRC 2007; EC 2010, 2013b).

3.4.1 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. Before using a sample of *clean* field-collected soil as *negative control soil* in a definitive toxicity test, the test laboratory must have previous experimental evidence showing that natural soil from this source can meet the criteria that must be achieved for the results of a toxicity test to be considered valid (see Section 4.4).

Accordingly, initial tests involving a sample of this soil must be performed using the intended boreal forest test species, to confirm that the test organisms are able to meet the criteria for test validity

[17] Light intensity, and its control thereof, can be as important, if not more so, as the pH and temperature during plant toxicity tests. The light fluence rate throughout the entire test area should be checked before initiating the test. The distance between the plant *canopy* and the light source can be increased or reduced in order to achieve the appropriate lighting conditions. Alternatively, the portion of the test area that is within 15% of the selected light fluence rate can be "mapped out" to designate the boundaries of adequate versus inadequate light fluence rate (EC, 1999b, 2005a).

[18] 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 summary table of CCME's Environmental Quality Guidelines can be accessed directly at http://st-ts.ccme.ca/. These websites 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 phone (1-204-948-2090) or email (info@ccme.ca).

(see Section 3.2.1). Thereafter, and assuming that the preceding results for these preliminary tests are satisfactory, 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 (%)[19],
- *organic matter* content (%)[19],
- *pH*,
- *conductivity*,
- *moisture content* (%),
- *water-holding capacity*,
- nitrogen as total N, nitrate (NO_3^-), nitrite (NO_2^-), and ammonium (NH_4^+),
- phosphorus as total or plant-available,
- potassium as total or plant-available,
- C:N ratio, and
- *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^-).

In order 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,
- herbicides suite,
- metals suite,
- petroleum hydrocarbons (including PAHs),
- other site- or area-specific contaminants of concern.

Pesticide and metal concentrations should not exceed the Canadian Soil Quality Guidelines if such guidelines are available (see footnote 18). If seeds from a natural seedbank germinate in the sample(s) of natural soil at any time (i.e., during storage or testing), then these seedlings must be removed.

[19] *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.

If the results of both the preliminary biological tests and the physicochemical analyses are satisfactory, a larger sample of this natural soil can be collected, air dried to a *moisture content* of between 10% and 20%, coarse-screened (4 to 10 mm), transferred to clean, thoroughly rinsed plastic pails, and stored in darkness at 4 ± 2°C until required. Plastic pails should not be used for the collection and storage of soils if there are concerns about chemical constituents of the plastic leaching into the soil.

3.4.2 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 recommended for use in tests with a reference toxicant (Section 4.9).

In keeping with the formulation of artificial soil recommended by OECD (1984, 2004), USEPA (1989), ISO (1991, 1993, 1998), and ASTM International (2004) for earthworm testing, and that recommended in Environment Canada's soil toxicity tests using earthworms, plants, and springtails (EC, 2004a, 2005a, 2007a), the following ingredients should be used to prepare artificial soil to be used in the biological test method described herein:

- 10% *Sphagnum* sp. peat, air dried and sieved through a 2-mm mesh screen
- 20% kaolin clay with particles < 40 µm
- 70% "grade 70" silica sand

The ingredients should be mixed thoroughly in their dry form using a mechanical stirrer and/or gloved hands.[20] Reagent-grade calcium carbonate should be added to the dry mixture in a quantity sufficient to attain a pH for the *artificial soil* ranging within 6−7.5 once it is hydrated.[21] Thereafter, the mixture

[20] 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 has been well mixed. Personnel must take the appropriate precautions for protection to prevent the inhalation of and contact with these ingredients.

[21] 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 to 30 g of $CaCO_3$ for each kg of peat might prove adequate. A pH as low as 4.5 can occur when the soil is first formulated without the addition of $CaCO_3$. The initial pH adjustment should attempt to raise pH to range within 7.0−7.5, since the pH of artificial soil typically drops slightly (to 6.5−7.0) during the three-day equilibration period, before it stabilizes. The pH of stored samples of artificial soil

should be hydrated gradually using *test water* (i.e., de-ionized or distilled water) until its *moisture content* is ~20% (which is ~28% of the soil's *water-holding capacity*), while mixing further until the soil is visibly uniform in colour and texture. As necessary, reagent-grade calcium carbonate should be added to the hydrated mixture in a quantity sufficient to maintain a pH ranging within 6.5−7.5. 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 (see footnote 21). Thereafter, *artificial soil* can be stored at 4 ± 2°C. As and when required for a soil toxicity test, a suitable quantity of stored artificial soil should be hydrated further using test water until its moisture content is ~70% of the water-holding capacity.

3.5 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 boreal forest plants, 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 that is predictable based on earlier toxicity tests with this material. The positive control soil might be a sample of negative control soil that is spiked with a *reference toxicant* for which historic data are available on its toxicity to plants using specified test conditions and procedures. For the test method described herein, one or more reference toxicants must be used as a positive control soil when appraising the sensitivity of the test organisms and the *precision* and reliability of results obtained by the laboratory for that material (see Section 4.9). A test might also include a sample of negative control soil (natural or artificial; see Section 3.4) 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 the plant species used according to the biological test method described herein. In some instances, a test might include a positive control soil that is comprised of a highly contaminated sample of field-collected soil or sludge shown previously to be consistently toxic to boreal forest plants according to the biological test method described herein.[22]

3.6 Reference Soil

One or more samples of *reference soil* might be included in a soil toxicity test using boreal forest plants.[23] 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., *texture*, *organic matter* content, organic carbon content, *pH*, *conductivity*, and *fertility*) represent the sample(s) of test (*contaminated*) soil as much as possible. Ideally, the reference soil is collected near the *site*(s) where samples of *test soil* are collected, but it is removed from the source(s) of contamination. One or more samples of field-collected *clean* reference soil from sites removed from the test site(s) might also be chosen due to their known lack of toxicity in previous tests with plants, and their possession of physicochemical characteristics similar to the samples of test soil. As described for test soils (Section 3.7), reference soils must be collected as separate *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.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2001).

A mixture of formulated artificial soil can also be stored dry, followed by partial hydration to ~20% moisture content, storage at 20 ± 2°C for a minimum 3-day period, and subsequent hydration to ~70% WHC when required 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 artificial test soil. Storage of artificial soil that is partially hydrated, rather than dry, is considered a preferred approach since it enables laboratory personnel to more quickly hydrate to the desired moisture content (i.e., ~70% WHC) while ensuring pH equilibrium, and reduces any further delay in time associated with the dry storage of artificial soil.

[22] If the *positive control soil* is comprised of a highly contaminated sample of field-collected soil, it is important that its phytotoxic potential is stable over time (i.e., the sample is old enough that the bioavailability has stabilized).

[23] The use of field-collected *reference soil* might not be appropriate for certain toxicity tests such as those using samples of sludge (Section 5) or *chemical-spiked soil* (Section 6).

horizons, where possible. Soils collected in horizons are reassembled in test units in proportions correlated to the depths of each horizon, as collected in the field (see Section 4.1). The horizon sample(s) of field-collected reference soil used in a study could be tested for *toxic* effects at full strength only, or the horizon sample(s) could be mixed with those of test soil to prepare a range of concentrations to be included in a multi-concentration test[24] (see Sections 3.7, 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.

An investigator might choose to include one or more samples of *artificial soil* as reference soil in a particular test. For instance, these samples could be used in multi-concentration tests with *site soils* or *chemical-spiked soils* to investigate the influence of certain physicochemical characteristics (e.g., a number of artificial reference soils prepared to provide a range of differing values for *texture* and/or percent organic matter content; Sheppard and Evenden, 1998; Stephenson *et al.*, 2002) on the toxicity of a contaminated site soil or a chemical-spiked soil. Multiple samples of clean field-collected soil collected from various sites, which differ markedly with respect to one or more physicochemical characteristics, might also be used for this purpose. For such a study, a portion of each reference soil used to prepare a series of concentrations of the 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 3.4). 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.

3.7 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 boreal forest plants as test organisms. The sample(s) of test soil might be either field-collected soil from an industrial or other site of concern, or soil under consideration for possible land disposal. A sample of field-collected test soil might be tested at a single concentration (typically, 100%) or evaluated for toxicity in a multi-concentration test whereby a series of concentrations are prepared by mixing measured quantities with either *negative control soil* or *reference soil* (see Section 5).[25] 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, both reference and contaminated soils must be collected in separate horizons. Soils collected in horizons are reassembled in test units in proportions correlated to the depths of each horizon, as collected in the field (see Section 4.1). However, if the contaminants of concern have only been confirmed in one *soil horizon* (e.g., upper organic horizon), based on previous data characterizing the extent and nature of the contamination, then a decision must be made whether to conduct plant toxicity testing on this horizon alone or in the reassembled soil horizons from this site (see Section 5.3). Soils without distinct soil horizons (e.g., where the surface soil horizons have been mixed or disturbed due to human activity) are 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).

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

[25] To locate areas of contaminant(s) or to characterize a site, it may be useful to collect multiple smaller samples (e.g., intact soil cores). Although soil cores have been tested, the test methodology was not sufficiently developed at the time of publication of this method to include it as part of the standard guidance provided herein (EC, 2010). However, some guidance on the use of intact cores is provided in Section 4.1.

Universal Test Procedures

General procedures and conditions described in this section for toxicity tests with terrestrial boreal forest plants apply when testing the toxicity of samples of soil or chemical, and also apply to their associated reference toxicity tests. More specific procedures for conducting tests with field-collected samples of soil or other similar particulate material are provided in Section 5. Guidance and specific procedures for conducting tests with *negative control soil* or other soil spiked (amended) experimentally with chemical(s) or chemical product(s) are given in Section 6.

All aspects of the test system described in Section 3 must be incorporated into these universal test procedures. Those conditions and procedures described in Section 2 for seed storage, handling, and *stratification* in preparation for soil toxicity tests also apply. A summary checklist in Table 3 describes 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* or *negative control soil* (or other soil, *contaminated* or *clean*) spiked in the laboratory with one or more test chemicals or chemical products.

This biological test method uses terrestrial boreal forest plant seed as test organisms, and measures *seedling growth* (emergence, shoot and root length and dry mass) inhibition as the biological endpoints. Test organisms are chosen from a list of seven species approved for use in this test method (see Section 1.2). Soils are collected as distinct *horizons* and reassembled in the test vessels prior to testing. Test duration is 28, 35, or 42 days, depending on the species chosen and the *biomass* needed for determination of the endpoint measurement(s) (see Section 4.3). The test soils are hydrated during the test but not renewed. This definitive test method was applied and validated by several participating laboratories in three rounds of concurrent tests using black spruce *(Picea mariana)* in a field-collected reference soil and artificial soil, and jack pine *(Pinus banksiana)* and bluejoint reedgrass *(Calamagrostis canadensis)* in two different field-collected contaminated soils diluted with field-collected reference soils (EC, 2013a).[26]

4.1 Preparing Test Soils

Each test vessel (see Section 3.2.2) placed within the test facility must be clearly coded or labeled 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

[26] In the first phase of the interlaboratory validation tests, six laboratories participated in six-week control performance tests with black spruce exposed to a multi-horizon field-collected reference soil (RS) and artificial soil (AS). All laboratories met the interim validity criteria of ≥ 60% emergence, ≥ 22 mm root length, and ≥ 20 mm shoot length in both soils. Based on the data from all of the laboratories, the mean % emergence in the AS and RS were 94% and 80%, respectively. The mean shoot length was the same for both soils (i.e., 28 mm), and the mean root lengths were 52 mm for the AS and 33 mm for the RS. Mean shoot weights for the six laboratories were similar for the two soils (i.e., 6.1 mg and 5.3 mg), and mean root weights were 1.8 mg and 0.6 mg for AS and RS, respectively. The interlaboratory variability, expressed as the coefficient of variation (CV) was low for emergence in both soils (4% and 10% for AS and RS, respectively), and slightly higher for the shoot length (11% and 20% for AS and RS, respectively) and root length (24% and 33% for AS and RS, respectively) endpoints. As expected, the variability was higher with the dry weight endpoints, with CVs ranging from 26% for shoot dry weight in AS to 85% for root dry weight in RS.

Six laboratories participated in the second phase of the interlaboratory validation tests. These were five-week definitive tests with jack pine exposed to a brine-contaminated, field-collected soil diluted with a field-collected reference soil. Only four laboratories met all three interim validity criteria established for jack pine (i.e., ≥ 60% emergence, ≥ 60 mm root length, and ≥ 43 mm shoot length), and therefore were the only tests valid for this round. All IC25s are reported as percentages of contaminated soil (e.g., 14% represents a mixture of 14% contaminated soil and 86% reference soil). The mean emergence of seeds in the reference soil for all four laboratories was 87%, with a CV of 16%. The mean IC25 for shoot length was 15% with IC25s for individual laboratories ranging from 8% to 20%. The IC25s for root length ranged from 3% to 13% with a mean IC25 of 6%. IC25s for shoot and root dry weight ranged from 14% to > 25% and 3% to > 25%, respectively for the various laboratories involved. The interlaboratory CVs were not calculable for all of the endpoints due to a number of "greater than" results (i.e., > 25%); however, the CVs based on IC25s for shoot and root length produced by the various laboratories were 36% and 79%, respectively.

The third phase of the inter-laboratory validation tests were four-week definitive tests with bluejoint reedgrass exposed to a field-collected petroleum hydrocarbon-contaminated soil diluted with a field-collected reference soil. This validation round was unsuccessful as four out of five laboratories failed one or more of the interim validity criteria established for bluejoint reedgrass. Following an investigation and some follow-up testing, the reasons for the failures could not be found (EC 2013a).

test vessels should be positioned such that observations and measurements can be made easily. *Treatments* should be positioned randomly within the test facility (EC, 1997a, b, 2004a, 2005a, 2007a) and rotated regularly (e.g., weekly while watering, or during observations of test vessels).

On the day of the start of the test, which is the day the seeds are initially exposed to samples of test material or substance (i.e., Day 0), each sample or subsample of *test soil* or similar particulate material, including *negative control soil* and, if used, *reference soil*, should be mixed thoroughly[27] (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). Each *horizon* must be prepared separately. If field-collected soils were collected and intended to be tested as *consolidated* samples they must remain intact for the duration of the test.[28]

Test soils for boreal forest plant testing are prepared on the day of test initiation (i.e., Day 0). The quantity of each test *soil horizon*, mixed as a *batch*, should be enough to set up the *replicates* of that *treatment* (see Table 3) plus an additional amount for the physicochemical analyses to be performed (Section 4.6) 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* might be made by taking aliquots of the mixture for measurements such as particle size analysis, *total organic carbon* (%), *organic matter* content (%), *moisture content* (%), and concentration of one or more specific chemicals.

Immediately following the mixing of a *batch*, an identical wet weight of test soil is transferred to each *replicate* test vessel. The volume of each *horizon* to be layered in the test vessel should, where possible, correlate proportionally to the depths of each horizon, as collected in the field. The total volume of soil in each test vessel is still ~500 mL (i.e., a wet weight equivalent to a volume of ~500 mL);[29] however, it may be composed of two or more layers with depths (i.e., a suitable wet weight of each *horizon* soil is added to each test vessel) that proportionally correspond to the depths of the *soil horizons* chosen to best illustrate the contaminated horizons at the test soil collection site (see Figure 2). The soil added to each test vessel should be smoothed (but not compressed) using a spoon, by gently shaking the vessel back and forth horizontally, or by gently tapping the vessel ≥ 3 times on the benchtop or with a hand. Care must be taken during the re-layering process to ensure the soil horizons do not mix with each other. For a multi-concentration test, each horizon of test soil is mixed with the same horizon of negative control or reference soil (see Section 5) at the appropriate test concentration (e.g., 25%). In some cases, it may not be possible to collect the same horizons of reference or negative control soil and test soil. For example, negative control soils may 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 may be mixed. In this case, test concentrations are prepared by mixing suitable

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

[28] The testing of soil cores has not be standardized in this test method document; however, for some investigations soil cores may be collected as described in EC (2012). *Consolidated* cores, collected using a slide hammer core sampler (e.g., Halltech Environmental Inc.), are encased in plastic sleeves that must remain in place during the test. The soil cores are refrigerated (i.e., 4 ± 2°C) until test initiation. Prior to test initiation, the soil cores are brought to room temperature, the lids are removed, and any vegetation or roots on the surface of the soil cores are removed. The soil surface is moistened, cut level with the top of the sleeve, and scored or loosened to allow planting of two seeds (for jack pine, white spruce, black spruce, bluejoint reedgrass, trembling aspen, paper birch) or one seed (for Canada goldenrod). Cores are then placed in a test vessel (e.g., 1-L polypropylene container) and covered to reduce moisture loss. All other aspects of the test are the same as those described for the definitive test (see Section 4 and Table 3). Preliminary use of this methodology is described in technical reports (SRC, 2007, 2009).

[29] The wet weight of soil required to achieve a total volume of ~500 mL depends on the moisture content, bulk density, and other characteristics of the soil, and will vary from sample to sample and horizon to horizon. Accordingly, the wet weight of each horizon required to achieve the appropriate volume should be determined by transferring the amount of sample required to fill a preweighed (or tared) 1-L test vessel to desired volume mark scribed on its side, followed by smoothing the surface and gently tapping the container on the bench top, three times. Thereafter, the wet weight of that quantity should be determined and recorded, and an identical wet weight added to each *replicate* test vessel. This process is repeated with subsequent layers until an overall total volume of ~500 mL is achieved.

weights of test soil into the available horizons of negative control soils at the appropriate test concentration (e.g., 25%). These mixtures are then layered in the test vessels at depths appropriate to represent field conditions or to fulfill study objectives.

Figure 2 Test units showing layering of soil horizons: (L) an Ontario podzol (Ah, Ae, and B horizons) and (R) a Saskatchewan brunisol (FH and AeB horizons)
(photo: H. Lemieux)

For a single-concentration test [e.g., *site soil* tested at 100% concentration only; a particular concentration of test soil; or a chemical 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 an identical wet weight of the same *batch* to each replicate vessel (for a total volume of ~500 mL). For site soils, *replicate vessels* should represent *replicate samples* (i.e., field replicates) collected individually from a given sample location (see Section 5.1). For a multi-concentration test, either equal or unequal replication across *treatments* can be used. If replication is equal across treatments, at least four replicate test vessels must be set up for each treatment. If replication is unequal across treatments (see Section 4.8), a minimum of six replicate vessels should be prepared for the negative control soil, four replicate vessels should be prepared for the lowest four to six test concentrations, and three replicate vessels should be prepared for the highest five test concentrations.[30] For any test that is intended to estimate the ICp in a definitive soil test (see Section 4.8), at least nine concentrations plus a negative control soil must be prepared and more (≥ 11) are recommended to improve the likelihood of bracketing each endpoint sought.[31]

Concentrations should be chosen to span a wide range, including a low concentration that obtains effects like the negative control, and a high concentration that results in "complete" or severe effects. It is a common mistake to anticipate the endpoint and bracket it with a closely spaced series of concentrations, all of which 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 geometric series should be used. Additional guidance on selecting test concentrations that applies here is found in EC (2005b).

4.1.1 Range-Finding Test

In the case of appreciable uncertainty about sample toxicity, it is often beneficial to run a *range-finding* test for the sole purpose of establishing more closely the concentrations to be used for the *definitive* test. Conditions and procedures for the range-finding test are similar to the definitive test (see Table 3); however, the experimental design differs.

The *range-finding* test may be a shorter-term test using ≥ 6 concentrations of test chemical or test soil covering a broader range,[32] and only duplicate

[30] A greater number of replicates can be used, and the distribution of replicates across treatments can be balanced (i.e., ≥ 4 replicates per treatment); however, the unbalanced nature of the regression design (i.e., unequal replicates among treatments) was developed to keep the level of effort comparable to that of an analysis of variance (ANOVA) design in terms of the total number of test vessels per test (Stephenson, 2003a; EC, 2005b).

[31] The large number of test treatments is needed to show the shape of the concentration-response relationship and to choose the appropriate linear or nonlinear regression model. Also, they contribute to the success of the computer calculations of the ICp and increase the probability of deriving a value (EC, 2005b).

[32] For range-finding tests, a suitable range of test concentrations for a test chemical spiked in soil might be 1000, 100, 10, 1, 0.1, and 0.01 mg/kg, or some common multiple thereof. For a multi-concentration test with a sample of contaminated (or potentially contaminated) field-collected soil, a suitable range of concentrations to use in a range-finding test might be 100, 50, 25, 12.5, 6.25, and 3% (Stephenson et al., 2001; EC, 2005b).

vessels (i.e., two replicates) per treatment. The test species must be the same as that to be used in the definitive test (see Section 2.1), and the number of seeds per replicate should be the same as those used in the definitive test (see Table 3 and Section 4.2).

Table 3 Checklist of required and recommended conditions and procedures for conducting definitive tests of soil toxicity using boreal forest plants

Universal		
Test type	–	whole soil toxicity test; no renewal (static test)
Test duration	–	· 28 days for trembling aspen or bluejoint reedgrass · 35 days for Canada goldenrod, paper birch, or jack pine · 42 days for white spruce or black spruce
Approved test species	–	· gymnosperms: white spruce *(Picea glauca)*, black spruce *(Picea mariana),* jack pine *(Pinus banksiana)* · angiosperm dicotyledons: trembling aspen *(Populus tremuloides)*, Canada goldenrod *(Solidago canadensis)*, paper birch *(Betula papyrifera)* · angiosperm monocotyledons: bluejoint reedgrass *(Calamagrostis canadensis)*
Number of concentrations	–	minimum of 9, plus negative control; recommend ≥ 11, plus negative control
Number of replicates	–	for single-concentration test (e.g., site soil tested at 100% concentration only): · ≥ 5 replicates/treatment for multi-concentration test: · ≥ 4 replicates/treatment for equal replicate test design; or regression design: unequal replicates among test treatments: · 6 replicates for negative control soil, · 4 replicates for lowest 4 to 6 test concentrations, and · 3 replicates for highest 5 test concentrations
Number of seeds per test vessel	–	· 5 seeds/vessel for trembling aspen, paper birch, Canada goldenrod, or bluejoint reedgrass, and · 10 seeds/vessel for black spruce, white spruce, or jack pine
Negative control soil	–	depends on study design and objectives; *clean* field-collected soil if testing site soils; soils collected in separate horizons are relayered in test vessels to represent field depths; natural control soil for tests with chemical(s) or chemical products(s) spiked in soil
Test vessel	–	polypropylene cups (1 L), covered for full duration of the test, lids replaced by inverted test unit or other suitable container if plants reach top
Amount of soil/test vessel	–	identical wet wt, equivalent to a volume of ~500 mL; ~350 g dry wt if artificial soil
Moisture content test soils	–	for soil preparation, hydrate to the optimal % of its water-holding capacity (WHC) if field collected soil (see Section 5.3), or to ~70% of WHC if artificial soil; each horizon prepared separately; during test, hydrate to maintain moist, crumbly texture as at start of test or if soil appears dry
Air temperature	–	daily range, constant 24 ± 3°C; alternatively, day: 24 ± 3°C, night: 15 ± 3°C
Humidity	–	test area may be kept at ≥ 50%, but this is not critical as test units closed for entire duration of test
Lighting	–	full spectrum fluorescent: mimic natural light spectrum (e.g., VitaLux® by Duro-Test®); 300 ± 100 µmol/(m^2 · s) adjacent to the level of the soil surface; 16 h light:8 h dark

Watering	–	hydration water sprayed conservatively over soil surface to maintain original moist, crumbly texture, and whenever soil appears dry
Measurement during test	–	soil moisture content in each treatment/concentration at start; pH in each treatment/concentration at start and end; temperature in test facility, daily or continuously; light intensity once during test; pH and moisture content measurements to be taken at each soil horizon
Observations during test	–	number of emerged seedlings at end of test in each test vessel; shoot/root length and shoot/root dry mass at test end; observations of emerged plants at test end showing an atypical appearance (e.g., chlorosis, lesions)
Biological endpoints	–	number of seedlings at end of test in each test vessel; length of longest shoot and longest root at test end; dry weight of entire shoot and root structures (oven-dried at 90°C until constant mass) at test end; optionally, wet weight of shoot and root at test end
Statistical endpoints	–	mean (± SD) percent emergence in control soil (for test validity) at test end (Day 28, 35, or 42); mean (± SD) length of longest shoots and roots in each treatment at test end (Day 28, 35, or 42); mean (± SD) dry wt of shoots and roots in each treatment at test end (Day 28, 35, or 42); if multi-concentration test: 28-, 35-, or 42-day ICp for each of mean shoot length, root length, shoot dry wt, and root dry wt in each concentration at test end
Test validity	–	invalid if **any** of the following occurs in negative control soil at test end: · mean % emergence is < 60% for trembling aspen, bluejoint reedgrass, Canada goldenrod, paper birch, jack pine, white spruce, or black spruce · mean root length is: < 35 mm for trembling aspen; < 17 mm for bluejoint reedgrass; < 80 mm for Canada goldenrod; < 53 mm for paper birch; < 62 mm for jack pine; < 36 mm for white spruce; < 24 mm for black spruce · mean shoot length is: < 10 mm for trembling aspen; < 35 mm for bluejoint reedgrass; < 7 mm for Canada goldenrod; < 26 mm for paper birch; < 44 mm for jack pine; < 26 mm for white spruce; < 20 mm for black spruce
Test with reference toxicant	–	must perform at least once every two months, or in conjunction with definitive test(s) with soil samples; use boric acid; prepare and test ≥ 5 concentrations plus a negative control, using artificial soil as a substrate; ≥ 3 replicates/concentration; 5 or 10 seeds per replicate (i.e., species-specific); follow procedures and conditions for a reference toxicity test described in Section 4.9; determine % emergence in controls and 14-, 21, 28-, or 35-day (species-dependent) ICp for root length (including 95% confidence limits); express as mg boric acid/kg, dry wt

Field-Collected Soil

Transport and storage	–	seal in plastic and minimize air space; transport in darkness (e.g., using an opaque cooler, plastic pail or other light-tight container); do not freeze or overheat during transportation; 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 (preferred) or artificial soil, for which previous plant tests have shown that all criteria for test validity could be regularly met
Reference soil	–	one or more samples for tests with field-collected soil or artificial soil; ideally taken from site(s) presumed to be clean but near sites of test soil collection; characteristics including percent organic matter, particle size distribution, and pH similar to test soil(s)
Characterization of test soils	–	at least particle sizes (% sand, % silt, % clay), TOC (%), OM (%), pH, conductivity, moisture content (%), WHC, nitrogen, phosphorus, potassium, C:N ratio and CEC and optionally, major cations and anions and contaminants of concern [e.g., metals, polycyclic aromatic hydrocarbons (PAHs), pesticides]
Preparation of test soils	–	if necessary, remove debris and indigenous macro-organisms using forceps; if necessary, press through a sieve of suitable mesh size (e.g., 4 to 10 mm); mix; determine soil moisture content; hydrate with de-ionized or distilled water (or, if and as necessary, dehydrate) to optimal percentage of its WHC (see Section 5.3); mix; dilute with control or reference soil if multi-concentration test; ensure homogeneity

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

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

The *range-finding* test may be a shorter-term test using ≥ 6 concentrations of test chemical or test soil covering a broader range (see footnote 32), and only duplicate vessels (i.e., two replicates) per treatment. The test species must be the same as that to be used in the definitive test (see Section 2.1), and the number of seeds per replicate should be the same as those used in the definitive test (see Table 3 and Section 4.2). Negative control soil, air temperature, lighting conditions, percent moisture of soils, watering, and measurements during the test are the same as those described for the definitive test (Table 3). Shoot length and root length can be used to predict where the *sublethal* endpoints for growth will be in the definitive test.[33] In most cases, the endpoints for growth in the definitive test will be at lower concentrations than those observed for the *range-finding* test, due to the longer test duration in the definitive test. The number of emerged seedlings at the end of the range-finding test should also be observed and recorded to determine whether the test validity criteria for seedling *emergence* in the definitive test are likely to be met (see Section 4.4).

[33] The effect on seedling growth can be visibly reflected in the above-ground biomass of the seedling; however, root endpoints are often more sensitive than shoot endpoints. Both shoot and root lengths, therefore, should be measured in a range-finding test in order to provide a strong likelihood that the concentration range chosen for the definitive test will bracket both shoot and root endpoints.

4.2 Beginning the Test

Following the addition of *test soil* to each test vessel, 5 or 10 stratified seeds, depending on the species (see Section 2.1), are selected under magnification (illuminated from below) from an aliquot of seeds stratified for an appropriate length of time. Aspen seeds are not stratified but are kept frozen until just before planting (e.g., 10 to 15 minutes). Seeds are planted in or upon the soil within each test vessel, in order of increasing test concentration. For species requiring only 5 seeds (i.e., trembling aspen, bluejoint reedgrass, paper birch, and Canada goldenrod), 4 seeds are distributed equally around one seed within the centre of the soil in each test vessel. For jack pine, white spruce, and black spruce, which require 10 seeds per test vessel, 9 seeds are distributed equidistant around one center seed. Using fine forceps, each seed of jack pine, white spruce, and black spruce should be planted to a depth that is twice the diameter of the seed itself. The seeds are covered with the surrounding test substrate by tapping the test substrate with a stainless steel spatula or glass rod.[34] Seeds of trembling aspen, bluejoint reedgrass, Canada goldenrod, and paper birch require light for germination and therefore must be pressed onto the surface of the soil. After the seeds have been added to each test vessel, the vessels are hydrated by spraying the soil surface with *hydration water* using a fine-mist spray bottle. Enough water is added to bring the moisture content of the soils close to saturation (i.e., near-saturation; see Section 4.5). Following hydration, lids (see Section 3.2.2) should be placed on the test vessels, to minimize loss of moisture.

[34] To avoid variability in the planting depth of seed, which can lead to variability in percent emergence, the following procedure may be used for planting jack pine, white spruce, and black spruce seed. When preparing each replicate, add all but a portion (~10%) of the soil to the test vessel. The surface of the soil in the test vessel is gently flattened, and the seed is placed on the soil surface in the species-specific pattern described in this section. A plastic template, with pre-marked holes that are just slightly larger than the diameter of the seed, is useful to achieve a uniform distribution of very small seeds within the test vessel. After the seed has been equally distributed on the surface of the soil according to the pattern described herein, the remainder of the soil (i.e., the ~10% that was not originally added to the replicate vessel) is used to cover the seed to a uniform depth. This is repeated for each of the test replicates (G. Lazarovits, personal communication, Agriculture and Agri-Food Canada, London, ON, 2004).

4.3 Test Conditions

- This is a 28-, 35-, or 42-day soil toxicity test, during which the soil in each test vessel is not renewed. The test duration for trembling aspen and bluejoint reedgrass is 28 days; for Canada goldenrod, paper birch, and jack pine it is 35 days; and for white spruce and black spruce (i.e., species that produce less *phytomass* and/or take longer to germinate) the test duration is 42 days.
- The test vessel is a 1-L clear polypropylene container. Its contents (i.e., test soil horizons totaling a 500-mL volume) are covered with a clear polypropylene lid (see Section 3.2.2).
- For a single-concentration of field-collected test soil (site soil tested at 100% only), five *replicate samples* (i.e., field replicates) collected from each *sampling location* should be tested. If only a single replicate sample is collected from a given sampling location, or for a particular concentration of test soil or chemical, at least five *replicates* must be set up for each test soil (i.e., each treatment). For a multi-concentration test, the use of an unequal number of replicate test vessels per test concentration and control, depending on concentration and treatment, is recommended. For equal replication across treatments, a minimum of four replicates per treatment are prepared; for unequal replication across treatments, a minimum of six replicates for controls, four replicates in the lowest four to six test concentrations, and three replicates in the highest five test concentrations, should be prepared (see Section 4.1 and footnote 30).
- The test must be conducted at a constant mean air temperature of 24 ± 3°C; or a daily mean air temperature of 24 ± 3°C and a nightly mean air temperature of 15 ± 3°C for those facilities that can accommodate daily changes in test temperatures (see Section 3.1).
- Test vessels must be illuminated with a 16-h light and 8-h dark daily *photoperiod*. Full-spectrum fluorescent lights or equivalents that mimic a natural light spectrum (e.g., VitaLux® by Duro-Test®) should be used. Light intensity adjacent to the surface of the soil in each test vessel must be 300 ± 100 µmol/($m^2 \cdot s$) (i.e., equivalent to 18 750 ± 6250 *lux*) (see Section 3.3).

4.4 Criteria for a Valid Test

For a valid test, each of the following three test criteria must be achieved:[35]

(1) The mean percent *emergence* for each plant species grown in negative control soil at test end must be:

≥ 60% for any species.

(2) The mean root length for each plant species grown in negative control soil at test end must be:

≥ 35 mm for trembling aspen;
≥ 17 mm for bluejoint reedgrass;
≥ 80 mm for Canada goldenrod;
≥ 53 mm for paper birch;
≥ 62 mm for jack pine;
≥ 36 mm for white spruce; or
≥ 24 mm for black spruce.

(3) The mean shoot length for each plant species grown in negative control soil at test end must be:

≥ 10 mm for trembling aspen;
≥ 35 mm for bluejoint reedgrass;
≥ 7 mm for goldenrod;
≥ 26 mm for paper birch;
≥ 44 mm for jack pine;
≥ 26 mm for white spruce; or
≥ 20 mm for black spruce.

4.5 Hydration of Test Soil During the Test

Test soils are hydrated to "near-saturation" as needed, throughout the test. Hydrating to near-saturation means, in this instance, that water is added to the surface of the soil until ~0.5 cm of water is temporarily (≤ 1 h) visible pooling at the bottom of the test vessel following its addition. *Hydration water*, at 24 ± 3°C, should be sprayed onto the surface of the soil using a fine-mist spray bottle on Day 0, just after the seed has been added to the test vessels.[36] Thereafter and for the duration of the test, water should be added conservatively (e.g., only if needed), and as required to maintain the original moist, crumbly texture of the soil (see Section 4.6).[37] It is recommended that the condition of the soil be checked a few days after planting, particularly in tests of species planted on the soil surface (trembling aspen, bluejoint reedgrass, Canada goldenrod, and paper birch) to ensure that seeds have adequate moisture to support germination; misting may be needed.

The location of the test vessels in the environmental chamber or the testing area should be randomly varied each time that moisture condition of the soil is checked or when water is added to test vessels, so that the test organisms within these vessels are randomly exposed to any slight variations in test conditions (i.e., lighting, temperature, humidity, or ventilation) that might exist in the testing area.

[35] The test validity criteria presented here are based on control data generated in many studies carried out during the development of the method. These studies involved various toxicants including salt- and hydrocarbon-contaminated soils, boric acid as well as the control data of nine different types of negative control soil (see Appendix F) (SRC, 2004, 2006, 2007, 2008, 2009; EC 2007b, 2010, 2013b). The coefficient of variation (CV) based on all of the data used to develop the test validity criteria for percent emergence ranged from 10% for paperbirch to 21% for goldenrod. For shoot length in control soils, the CV ranged from 10% for white spruce to 61% for goldenrod and for root lengths, the CV ranged from 26% for goldenrod to 39% for bluejoint reedgrass. Most of these levels of variability are considered to be acceptable in terms of intra-and inter-laboratory precision. Some of the higher CVs calculated for shoot and root length data for some of the test species are due to the variability in growth observed among the eight substantially different natural soils and the artificial soil used to develop the validity criteria (EC, 2013b).

[36] It is also acceptable to add hydration water to the soil surface before planting. In this case, seeds of aspen, bluejoint reedgrass, Canada goldenrod and paper birch are pressed into the moist soil surface but not covered with soil. Seeds of white spruce, black spruce, and jack pine are planted as described in Section 4.2.

[37] The rate of watering depends on the rate of water loss from the soil and might vary between test vessels. The rate of water loss is influenced by soil type, its water-holding capacity, organic matter content, root mass, the development of shoot *canopy* during the test, and the humidity of the test facility. Over-watering of clay soils can result in a breakdown of the soil structure, whereas highly organic soils may require more frequent watering. The watering regime changes over the duration of the test. Since it is not required to remove the lids from the test vessels during the test, watering might not be necessary. If watering is necessary, however, it can be accomplished by gently spraying the soil surface using a spray bottle, and subsequently observing (over one hour) to see if a condition of "near-saturation" has been achieved (as per guidance herein). At this time, water might still be added to the soil surface by spraying. The judicious use of a spray bottle ensures that the surface soils do not get unduly disturbed (i.e., the integrity of the surface-soil structure is maintained) and that the water is distributed uniformly over the surface of the soil column within each test vessel. As the root mass becomes appreciable, there is less water lost directly from the soil via evaporation, but the roots can take up more water. At this point, water may be added by gently pouring it over the soil surface while being careful not to exceed the saturation of the soil column.

4.6 Observations and Measurements During the Test

The biological endpoints for the test are seedling root and shoot length, and root and shoot dry mass at the end of the test (i.e., on Day 28, 35, or 42, depending on the test species). Depending on the study objectives, root and shoot wet mass might also be determined at the end of the test; however, these endpoints are optional. Throughout the test, observations should be made and recorded of the number of emerged and the state or condition of the emerged plants (e.g., weekly or more frequently, as needed).

In order to determine whether the test validity criterion for percent seedling *emergence* in negative control soil has been met (see Section 4.4), the number of emerged seedlings in the control test vessels must be counted at the end of the test. Seedling *emergence* is measured visually by counting the number of seedlings that have emerged ≥ 3 mm, vertically from or horizontally across the soil surface (Figures 3 and 4). The emergence measurement includes the stem from the surface of the soil to the tip of longest leaf. The approximate number of days for complete emergence of the test species is species-specific (4 d for trembling aspen and goldenrod, 7 d for bluejoint reedgrass and paper birch, 10 d for jack pine, 16 d for black spruce, and 18 d for white spruce).

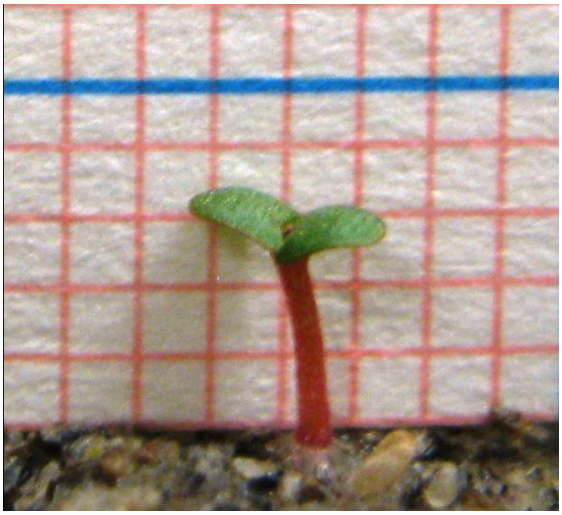


Figure 3 Early emergence of aspen to 4 mm (against mm ruled paper, photo: D. Bolin)

Figure 4 Early emergence of black spruce (against mm ruled paper, photo: D. Bolin)

A *visual assessment* of the health and condition of the plants (e.g., *phytotoxicity*) in each test vessel should also be made and recorded when the plants first appear, and thereafter each time the moisture condition of soil in a test vessel is checked (i.e., weekly) and/or the soil is watered.[38] Observations might include:

- *chlorosis* (loss of pigment),
- *necrosis* (localized dead tissue),
- *defoliation* (loss of leaves),
- *desiccation* (dried leaves or stems),
- *malformation* (structural defects),
- *mottling* (marked or spotted),
- *staining* (discolouration),
- *wilting* (limp),
- *withering* (in the process of drying),
- discoloured or deformed leaves or stem,
- overt signs of delayed emergence, or
- impaired development and/or growth.

It is normal for aspen and paper birch plants to display a range of colours from red to greenish red to

[38] Each time a test vessel is watered or the moisture condition of the soil is checked (i.e., weekly), the general condition of the plants therein should be assessed and observations recorded. These observations can be useful for interpretation of the results. For example, it is important to note when symptoms of stress (e.g., wilting, discoloration) first occur, and whether they get progressively worse or not. It is also important to record any signs of phytotoxicity for the plants in the negative control soil treatment, again for interpretation of the test results at the end of the test. Phytotoxic signs can be indicative of effects from physical factors such as lights too close to the foliage *canopy*, soils too dry between watering, or nutrients in soil are either limited or in excess (G.L. Stephenson, personal communication, Aquaterra Environmental, Orton, ON, 2002).

green depending on soil and light conditions. Figure 8 illustrates normal leaves of paper birch that have a red tint when grown in artificial soil and are bright green in the New Brunswick podzol reference soil. *Conifer* leaves may also display reddish tips.

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

The light fluence rate must be measured at least once during the test period at points approximately the same distance from the light source as the soil surface and at several locations in the test area (see Section 3.3).

In at least one replicate of each *treatment* (including the *negative control soil* and, if used, *reference soil*), the pH must be measured and recorded at the beginning and end of the test, and the *moisture content* must be measured and recorded at the beginning of the test only.[39] These measurements must be made in each separate soil horizon. The initial (Day 0) measurements should be made using subsamples of each *batch* of *test soil* used to set up replicates of a particular *treatment* (see Section 4.1).[40] The final (i.e., Day 28, 35, or 42) measurements should be made using subsamples of the replicates of each treatment (whether composed of a single soil or a number of horizons) to which plants were exposed, following the end-of-test observations of plant emergence, condition, and growth.

Soil pH should be measured using a calcium chloride ($CaCl_2$) slurry method (modified from Hendershot *et al*., 1993).[41] For these analyses, 4 g of hydrated soil[42] is placed into a 30-mL glass beaker (~3 cm in diameter and ~7 cm high) with 20 mL of 0.01 *M* $CaCl_2$.[43] The suspension should be stirred intermittently for 30 min (e.g., once every 6 min). The slurry should then be left undisturbed for ~1 h. Thereafter, a pH probe is immersed into the supernatant and the pH recorded once the meter reading is constant.

The *moisture content* of each horizon of *test soil* is measured by placing a 3 to 5 g subsample of each test soil into a pre-weighed aluminum weighing pan, and measuring and recording the wet weight of the subsample. Each subsample should then be placed into a drying oven at 105ºC 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:

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

[39] The moisture content of test soils can be quite variable at the end of the test, depending on how much time has lapsed between hydrating the test soils and processing the vessels at the end of the test. This, in addition to the fact that test vessels are watered to near-saturation throughout the test, means that the measurement of moisture content at the end of the test is not required.

[40] Additional soil for each batch should be prepared for physicochemical analyses of the test-initiation (Day 0) conditions. One or more additional replicates of each test soil should be prepared and placed into a test vessel within the test facility. These replicates (with seed added) should be reserved for physicochemical analyses of test-end (i.e., Day 28, 35, or 42) conditions.

[41] The method by Hendershot *et al.* (1993) includes a step that involves air drying the sample for 48 h before analyzing the pH. The experience of Environment Canada investigators is that this step is needlessly time-consuming (K. Doe, personal communication, Atlantic Environmental Science Centre, Environment Canada, Moncton, NB, 2004; J. Princz, personal communication, Soil Toxicity Laboratory, Environment Canada, Ottawa, ON, 2004), and does not appreciably modify the pH relative to that for hydrated (i.e., as per the toxicity test) soil (Courchesne *et al.*, 1995; J. Princz, personal communication, 2004). Becker-van Slooten *et al*. (2004) assessed three different soil slurry methods for measuring pH. The need for this testing was identified during Environment Canada's soil toxicity workshop in Vancouver, B.C. (February, 2003), during which certain participants recommended that a commonly used and "universally standardized" method for measuring soil pH be incorporated into each of Environment Canada's soil toxicity test methods (EC, 2004b). The following three methods for measuring soil pH were compared: 1) 1 *M* KCl in water; 2) 0.01 *M* $CaCl_2$ in water; and 3) water only. Results of this investigation showed that there were advantages and disadvantages with each of these methods for measuring pH. However, based on practical considerations and the recommendations of the workshop participants (i.e., that a widely used method for characterizing soil pH be applied), the 0.01 *M* $CaCl_2$ method was recommended as the most appropriate for Environment Canada's soil toxicity test methods (Becker-van Slooten *et al*., 2004).

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

[43] To prepare 0.01 *M* $CaCl_2$, dissolve 2.940 g of calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$) with distilled water, in a 2000-mL volumetric flask. The conductivity of the $CaCl_2$ solution should be between 224 and 240 mS/m at 25ºC, and the pH should range within 5.5 to 6.5 at 25ºC (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 conductivity is not within the acceptable range, a new solution must be prepared.

It is important that the moisture content (%) calculation be based on dry weight (**not** wet 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 strengths tested, as a minimum (see Section 6.3). Aliquots for these analyses should be taken from each soil horizon as described previously for pH and moisture content; analyses should be according to proven and recognized (e.g., SPAC, 1992; Carter, 1993) analytical techniques.

4.7 Ending the Test

The test is terminated after 28 days of exposure for trembling aspen and bluejoint reedgrass; after 35 days of exposure for Canada goldenrod, paper birch, and jack pine; and after 42 days of exposure for white spruce and black spruce. At that time, the number of plants surviving in each test vessel should be determined and recorded, and any abnormal patterns in morphology, growth, and development (i.e., relative to the plants in the negative control soil) also recorded.[44] Photographs might be taken to visually record the concentration-response relationship in the above-ground *phytomass*. Even if no shoots are visible above the soil surface, the soil should be checked for root material in case roots developed from the seed but no shoot material was produced. These observations are for qualitative purposes only [i.e., for this test method, a seedling must emerge 3 mm vertically from or horizontally across the soil surface to be considered "emerged" (see Figures 3 and 4)] and, if roots develop, where no shoot material was produced, it should be noted. Thereafter, each test vessel must be processed separately to keep the seedlings within each replicate isolated from those in each of the other replicate test vessels.

The plants must be carefully separated from the test soil and from the roots of the other plants. This can be achieved by gently loosening the soil and root matrix from the test vessel and removing all soil that can be easily removed without disturbing the root matrix. In some cases, roots can be more easily separated from the soil after the soil is first saturated with water and allowed to soak for several minutes. The remaining soil and plant mass are placed into a pan of water. The roots can then be held under a gentle stream of tap water, or they can be sprayed with water from a spray bottle, to gently dislodge as many of the remaining soil particles as possible. This also aids in separating the roots of the plants from each other. The plants are then placed onto a moistened, labelled sheet of paper towel, one for each test vessel, and covered with plastic to minimize water loss until measurements can be made and recorded. For trembling aspen, bluejoint reedgrass, Canada goldenrod, and paper birch, measurements of shoot and root lengths are made from the transition point between the *hypocotyl* and the root (Figures 5 to 8). The leaves and roots are gently straightened, and measurements made to the tip of the longest shoot or root when gently straightened. Shoot and root length for each plant in each replicate are measured with a ruler and recorded in millimetres. For the conifers (white spruce, black spruce, and jack pine), measurements of shoot and root lengths are made from the point where a discernible transition between root and shoot tissue is visible (e.g., colour, difference in stem alignment or texture, in this order of priority) (see Figures 9 to 11).

Using a scalpel, the shoots and roots are then separated from each other at the point at which there is a discernible transition between root and shoot tissue, and from the seed itself. The remaining seed or seed coat is discarded. The shoot and root structures from each replicate test vessel are weighed separately, as two groups (i.e., shoots and roots). The entire rinsed shoot *biomass* from each test vessel

[44] Surviving plants, for the purposes of this method, include live plants that are at least partly green (especially at the base) or red (in the case of aspen and paper birch), having shoots at least 3 mm long, and sufficient integrity that they can be removed from soil, washed, and measured without disintegration. Plants that are not measurable (e.g., dry/shriveled, and brown) should be noted, but must be excluded from the emergence and biomass determinations at the end of the test.

must be transferred as a group to a damp paper towel or blotting paper. Thereafter, they should be placed into a clean aluminum weighing pan (1−2.5 g) that has been previously numbered, weighed, and held in a desiccator.[45] This process is repeated with the entire rinsed root *biomass* from each test vessel. It is acceptable to cut plant tissue into segments or to divide large amounts of *phytomass* between two or more weighing pans. If wet mass is being determined, the aluminum pans containing shoots and roots are weighed immediately with an analytical balance that measures consistently to 0.1 mg. The dry mass must be determined and is done so in a similar way once the plants are dried in an oven at 90ºC until a constant weight is achieved (this usually takes a minimum of 24 h) (Aquaterra Environmental and ESG, 2000). Upon removal from the oven, the weighing pans are moved immediately to a desiccator. Once cooled, each weighing pan should be individually and randomly removed from the desiccator and weighed immediately[46] to the nearest 0.1 mg on a balance capable of measuring accurately to this limit. Mean dry weight per surviving plant (see footnote 44) is calculated for each replicate (see Section 4.8.2).

Although it is the intention of Environment Canada to use mean shoot dry weight and mean root dry weight as additional test validity criteria for definitive tests, there are insufficient data at this time on which to base minimum weight requirements for control plants. It is recommended, however, that for definitive tests:

- The mean shoot dry weight per surviving plant, for each plant species grown in negative control soil, at test end be:

  ≥ 2.1 mg for trembling aspen;
  ≥ 1.5 mg for bluejoint reedgrass;
  ≥ 1.0 mg for Canada goldenrod;
  ≥ 8.9 mg for paper birch;
  ≥ 6.6 mg for jack pine;
  ≥ 3.0 mg for white spruce; or
  ≥ 2.3 mg for black spruce.

- The mean root dry weight per surviving plant, for each plant species grown in negative control soil, at test end be:

  ≥ 0.4 mg for trembling aspen;
  ≥ 0.4 mg for bluejoint reedgrass;
  ≥ 1.1 mg for Canada goldenrod;
  ≥ 1.3 mg for paper birch;
  ≥ 2.0 mg for jack pine;
  ≥ 0.6 mg for white spruce; or
  ≥ 0.3 mg for black spruce.

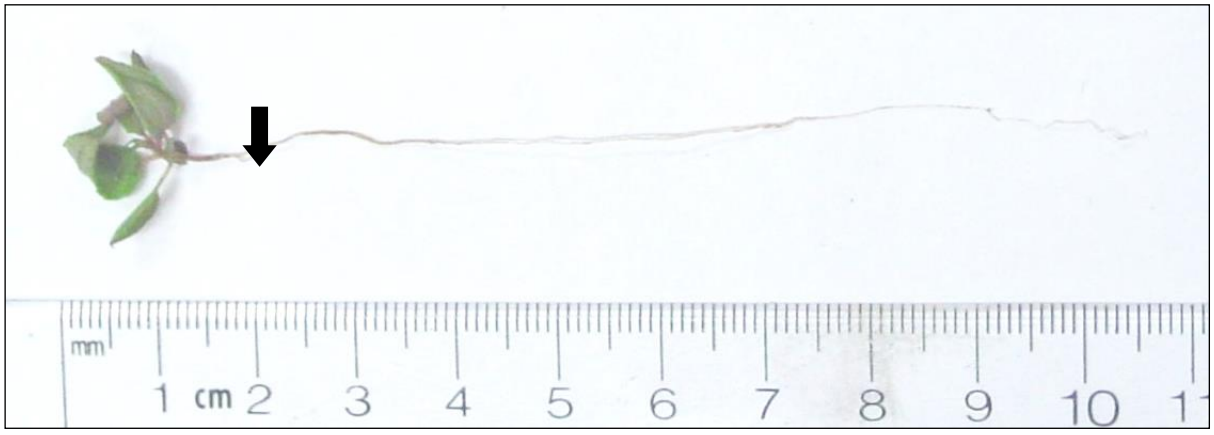


Figure 5 Aspen seedling showing transition point between shoot and root (photo: M. Moody)

[45] If any deposits (e.g., wax) associated with the weighing pans are cause for concern with respect to providing weighing errors, the weighing pans should be oven-dried for at least 48 h to achieve a constant weight (EC, 1997a, b, 2004a, 2005a, 2007a).

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

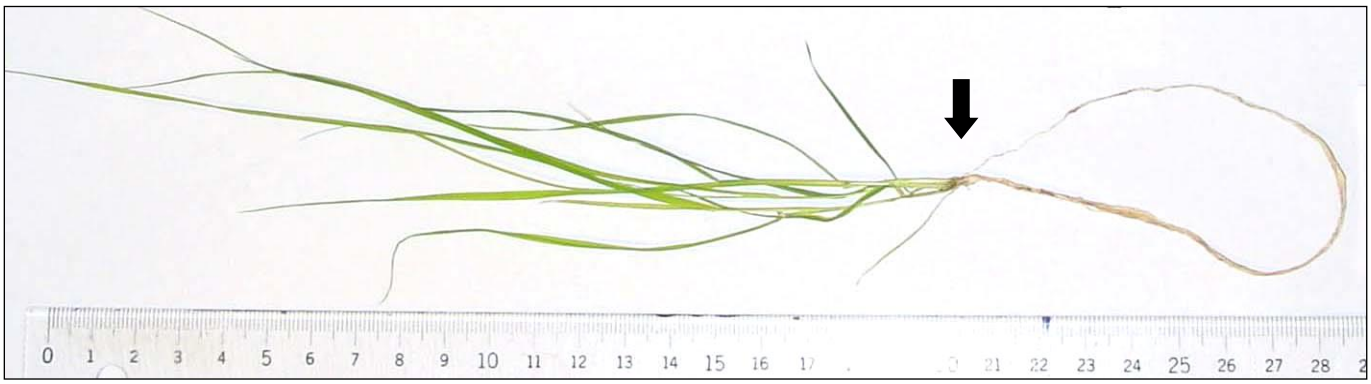


Figure 6 Bluejoint reedgrass seedling showing transition point between shoot and root (photo: M. Moody)

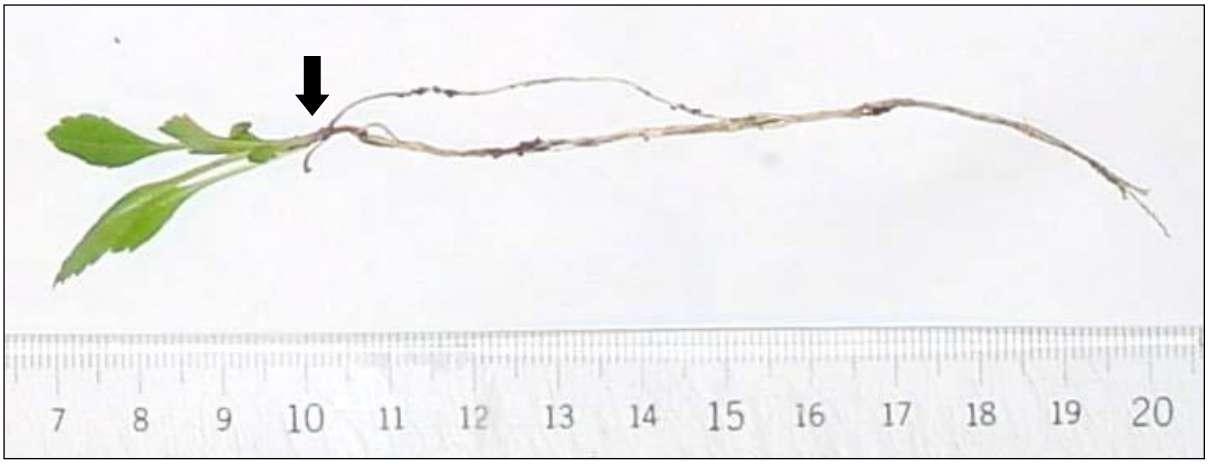


Figure 7 Transition point between shoot and root for goldenrod (photo: M. Moody)

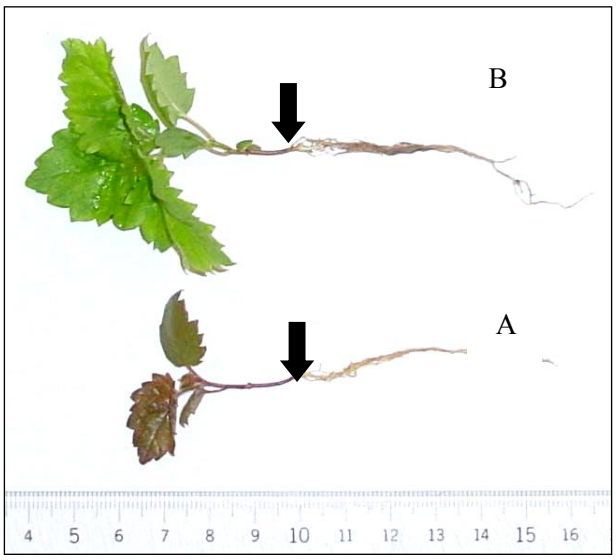


Figure 8 Paper birch seedlings grown in artificial soil (A) and New Brunswick podzol (B) showing transition points between shoot and root and colour difference (photo: M. Moody)

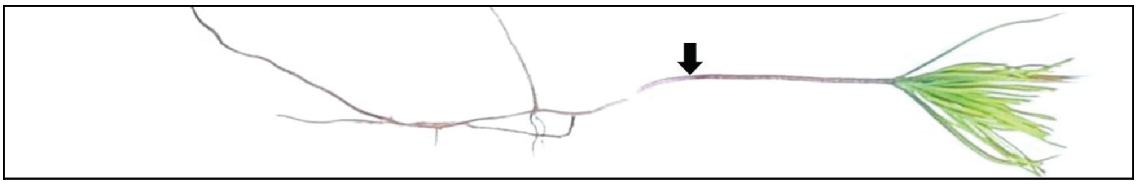


Figure 9 Jack pine seedling showing transition point between root and shoot based on colour (photo: M. Moody)

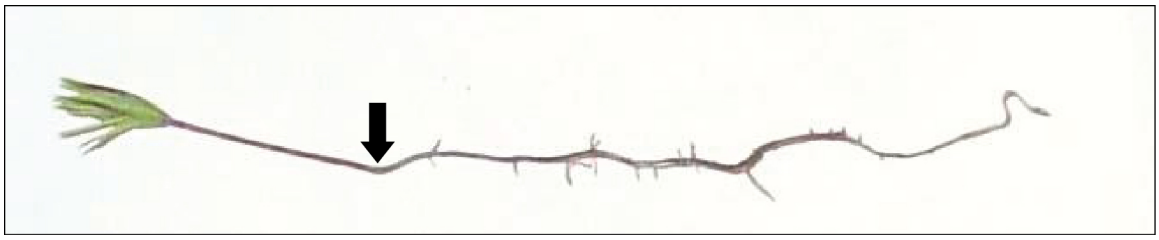


Figure 10 White spruce seedling showing transition point between shoot and root based on alignment of stem (photo: M. Moody)

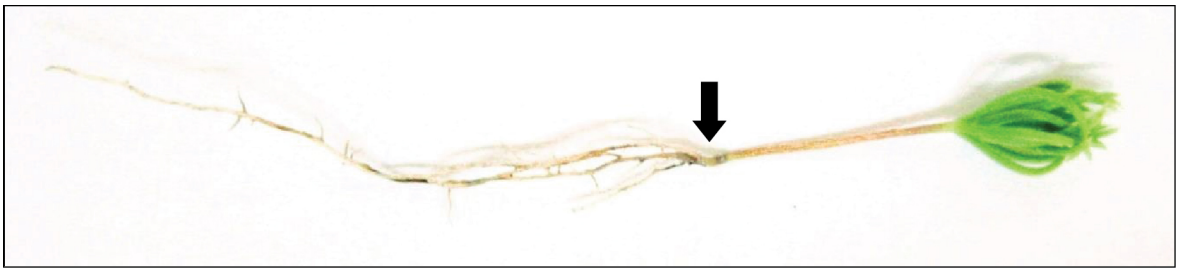


Figure 11 Black spruce seedling showing transition point between root and shoot based on texture (photo: M. Moody)

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

Following the removal of plants from each test vessel, subsamples of each horizon of *test soil* (including the *negative control soil* and, if included in the test, *reference soil*) should be taken for pH determination (Section 4.6). Analyses for other chemical constituents (i.e., concentrations of contaminants) should also be made at this time using representative subsamples of each test soil horizon collected (Section 4.6).[47]

[47] If soaking of soils is necessary to ease removal of plants (see second paragraph of Section 4.7), aliquots for analyses of chemical constituents (i.e., concentration of contaminants) should be collected before the soil is soaked and the plants are removed. Alternatively, independent replicates might be set up for this purpose alone (see footnote 40).

4.8 Test Endpoints and Calculations

Although *emergence* is not a statistical endpoint in this test (see Section 1.1), emergence data are required for the calculation of mean shoot and root dry weight per plant and for the determination of test validity (i.e., emergence in negative control soil). The number of emerged seedlings in each test vessel at the end of the test (Day 28 for trembling aspen and bluejoint reedgrass; Day 35 for Canada goldenrod, paper birch, and jack pine; and Day 42 for white spruce and black spruce) must be recorded for each test. The mean (± SD) percent *emergence* for all replicate groups of plants exposed to each treatment, including the control(s) for 28, 35, or 42 days (depending on the species used) must also be calculated and reported. Any optional observations of *emergence* taken (see Section 4.6) should also be calculated and reported as percent *emergence* in each test vessel, as well as mean (± SD) percent *emergence* for each treatment.

The *growth* endpoints for this test are based on shoot and root length, as well as shoot and root dry weight, of surviving (see footnote 44) plants in each *replicate* and each *treatment* as measured at the end of the 28-, 35-, or 42-day test period. Shoot and root wet weight are additional (optional, but recommended) endpoints. A significant reduction in the length or weight of the plants is considered indicative of an adverse *toxic* effect of the treatment on the growth of test plants.

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

(1) Multiple *sampling locations*, in which responses at one or more test site sampling location(s) are compared with those at a reference site sampling location,[48] 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.

(2) Multiple concentrations of a test soil, achieved by mixing a test soil with reference or control soil, or by spiking a soil with various concentrations of a chemical or chemical product. For a multi-concentration test, the 28-, 35-, or 42-day *ICp* for growth inhibition represented by each endpoint measurement (i.e., decreased mean length of individual plant shoots and roots, and decreased mean dry weights of individual plant shoots and roots) must be calculated and reported (data permitting).[49]

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 may 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 (1) undesirable

[48] 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.

[49] Historically, investigators have frequently analyzed *quantitative* sublethal data from multi-concentration tests by calculating the no-observed-effect concentration (NOEC) and the lowest-observed-effect concentration (LOEC). Disadvantages of these statistical endpoints include their dependence on the test concentrations chosen and the inability to provide any indication of *precision* (i.e., no 95% or other confidence limits can be derived) (NERI, 1993; EC, 2005b). Given these disadvantages, ICp is the required statistical endpoint for growth data derived in a multi-concentration test using boreal plants. Environment Canada has fully adopted regression based methods in aquatic-, sediment- and soil-based environmental toxicity testing (EC, 2004a, 2005a, b, 2007a,c, d, 2011a,b; Van der Vliet *et al*., 2012), 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).

effects on Type I and Type II error rates may occur; (2) interpretation of results is often more difficult; and (3) unwarranted focus may be given to particular comparisons after data have been collected.[50] Detailed statistical guidance on hypothesis testing for shoot and root length and shoot and root dry weight at test end is provided in Section 5.6.

Environment Canada (2005b) provides direction and advice for calculating ICps, which should be followed; Section 4.8.2 gives further guidance in this regard. Initially, regression techniques (see Section 4.8.2.1) must be applied to multi-concentration data intended for calculation of an ICp.[51] In the event that the data do not lend themselves to calculating the 28-, 35-, or 42-day ICps for the growth 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 4.8.2.2).

4.8.1 Percent Emergence

The mean and standard deviation of seedling emergence must be calculated for each test concentration. Percent *emergence* in negative control soils is measured and compared with the validity criterion for each species, to determine test validity (see Section 4.4). Emergence is also assessed in each treatment for calculation of dry weight per plant (see Section 4.8.2). Percent emergence is not a sensitive test endpoint for tests of boreal forest plants, so EC50 or other such endpoints are not calculated.[52]

4.8.2 ICp

When a multi-concentration test for effects of exposure of plants to *spiked soil* mixtures (including dilutions of test soil with a reference or control soil) is conducted, the *quantitative* continuous data representing growth inhibition (i.e., shoot and root length, and shoot and root dry mass) must be used to calculate an *ICp* (*inhibiting concentration for a specified percent effect*) for each of these four endpoints, data permitting (see introductory paragraphs of Sections 4.8 and 6.2). The ICp is a quantitative estimate of either:

(1) the concentration causing a fixed percent reduction in the mean length of individual plant shoots at test end;

(2) the concentration causing a fixed percent reduction in the mean length of individual plant roots at test end;

(3) the concentration causing a fixed percent reduction in the mean dry weight of individual plant shoots at test end; or

(4) the concentration causing a fixed percent reduction in the mean dry weight of individual plant roots at test end.

The ICp is calculated as a specified percent reduction for each endpoint (e.g., the IC25 and/or IC20, for a 25% and/or 20% reduction, 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 *growth*, the length and weight measurements of individual shoots or roots in each replicate (test vessel) are pooled for each of these measurements, and the mean of these lengths and weights are used in the analyses. For length measurements, the length of the longest shoot and the longest root is measured for each surviving plant (see footnote 44), and then the mean is calculated for each replicate. For dry weight measurements, the mean weight of individual shoots (or roots) in each replicate is calculated as the total dry weight

[50] Zajdlik & Associates Inc. (2010) made this last point in the defense 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.

[51] Regression is the method of choice for estimating ICp. It involves fitting the data mathematically to a selected model and then calculating the statistical endpoint using the model that best describes the exposure-concentration response relationship. Nonlinear regression techniques were originally recommended by Stephenson *et al.* (2000) for several reasons, including: the relationship that exists between exposure concentration and plant response is typically nonlinear; 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 nonlinear regression can fit effect distributions showing *hormesis*. By using standard mathematical techniques, a regression can be well-described in terms that convey useful information to others, effects at high and low concentrations can be predicted, and confidence intervals can be estimated. Deficiencies of the smoothing and interpolation method can be largely remedied (EC, 2005b).

[52] If the study objectives include the calculation of an EC50, the reader should consult EC (2005a, b) for guidance.

of all of the plant shoots (or roots) that survived in the test vessel divided by the number of plants that survived in that vessel and were measurable at the end of the test.[53, 54]

The mean lengths and weights from all the replicates of a given treatment (concentration) are used to calculate the average for the treatment; this is the average individual shoot and root length and shoot and root dry weight of surviving plants per concentration. These data are compared to the average individual lengths of shoots and roots and the average individual weights of shoots and roots in the *negative control*, obtained by the same procedure. If there are no emerged plants in a replicate (test vessel), that replicate does not contribute to the average for the treatment. If there are no emerged plants in all replicates at a given concentration, that concentration does not have an average length or weight of emerged plants and cannot be used in the analysis for comparison with the average length or weight in the negative control.

As indicated in the introductory paragraphs of Section 4.8, separate ICps for individual shoot and root length and shoot and root dry mass must be calculated and reported (data permitting) upon completion of the test. These calculations must be made using the appropriate linear or nonlinear regression analyses (see Section 4.8.2.1). If, however, regression analyses fail to provide meaningful ICps for shoot/root length or shoot/root dry weight, the ICPIN analyses described in Section 4.8.2.2 should be applied to the corresponding data.

4.8.2.1 Use of Regression Analysis

Upon completion of a *definitive* 28-, 35-, or 42-day multi-concentration test, separate ICps (including their respective 95% confidence limits) for the individual mean lengths and dry weights of shoots and roots must be calculated using regression analysis, provided that the assumptions below are met. A number of models are available to assess length and dry weight data (using quantitative statistical tests) via regression analysis. The proposed models for application consist of one linear model and the following four nonlinear regression models: exponential, Gompertz, logistic, and logistic adjusted to accommodate *hormesis*[55] (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.[56]

[53] Productivity can be analyzed by measuring the weight of all shoots or all roots as a group in each replicate (i.e., no division by emerged seedlings). Productivity is a population indicator that combines both effects (i.e., survival and growth) and therefore provides less opportunity for interpretation of the data. Futhermore, if there is an interest in calculating "productivity," the investigator can do so once the individual effects have been determined (i.e., growth and survival have been assessed separately) (A. Renoux, personal communication, SANEXEN Environmental Inc., Varennes, QC, 2004).

[54] To measure the dry weight of individual shoots (or roots) within each replicate, the total dry weight of plant shoots (or roots) in each replicate is divided by the number of plants that emerged in that replicate and are measurable (see footnote 44) at the end of the test. For example, if 10 white spruce seeds were planted in a given replicate, but only 7 plants were measurable at the end of the test, then the dry weight measurement for that replicate would be the dry weight of all 7 shoots, divided by 7. The same holds true for whole roots or shoots lost (e.g., washed down the drain) during the processing of a test vessel. For plants that break apart during the processing of test vessels (i.e., a portion of the shoot or root breaks off), professional judgment should be used to determine what to include in the final length and weight measurements. If the portion of the root or shoot that broke off is recovered (i.e., not lost), then it can be included in both the length and weight measurements. If, however, the portion of the shoot or root that broke off is lost, then the length measurement for that shoot or root should be excluded from the final length determination for the replicate. For dry weight measurements, the decision as to what to include (or not to include) depends on the estimated proportion of the root (or shoot) that is missing. For example, if the portion of a missing root is less than half of the root (i.e., based on the length of the other roots in the replicate), then the missing portion would not have a large overall effect on the final dry weight per root, and therefore, the remaining portion of the root may be included in the dry weight measurement for that replicate. If, however, the portion of the missing root is estimated to be more than half of the root, then that root should be excluded from the weight analysis (i.e., the root is not dried and weighed and the final "per plant" dry weight for the replicate is based only on the number of roots that were dried and weighed).

[55] A hormetic response (i.e., *hormesis*) might be found in sublethal observations at the lowest concentration(s), i.e., performance at such concentration(s) is enhanced relative to that in the negative control. For instance, the shoot and root lengths might be longer for seedlings grown in soil with low concentrations than for those grown in the control treatment, or the dry weights of shoots or roots might be substantially greater relative to those for seedlings grown in the control. 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.

[56] 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

An initial plot of the raw data (dry weight) 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).

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 12. The selection process begins with an examination of a scatter plot or line graph of the test data to determine the shape of the concentration-response curve. The shape of the curve is then compared to available models so that one or more appropriate models that best suit the data are selected for further examination (see 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.[57] 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 depends on the subjective assessment of the user. If the data are not normally distributed, then the user is advised to try another model, consult a statistician for further guidance on model selection, or to perform the less-desirable linear interpolation (using ICPIN, see Section 4.8.2.2) 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. Before choosing the weighted regression, the standard error of the ICp is compare to that derived from the unweighted regression. If there is a difference of greater than

SYSTAT is available for purchase by contacting SYSTAT Software, Inc., 225 W. Washington St., Suite 425, Chicago, IL 60606, USA, Tel. 1-877-797-8280; see website www.systat.com/products/Systat/. The latest version of CETIS™ is available for purchase by contacting Tidepool Scientific Software, P.O. Box 2203 McKinleyville, CA 95519, USA; Tel./Fax 707-839-5174; email: sales@tidepool-scientific.com.

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

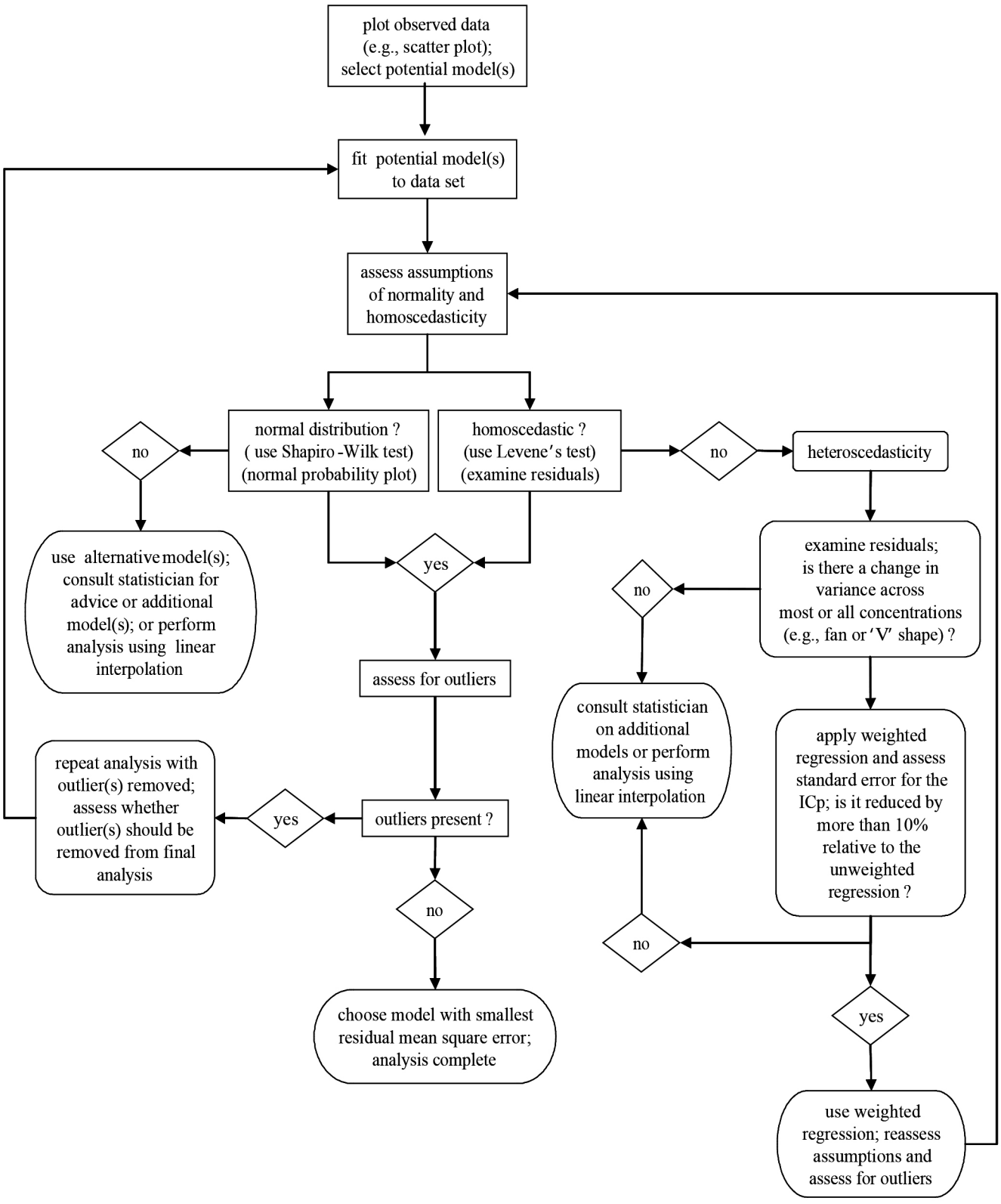


Figure 12 The general process for the statistical analysis and selection of the most appropriate model for quantitative toxicity data (adapted and modified from Stephenson *et al.*, 2000)

10% between the two standard errors,[58] then the weighted regression is selected as the regression of best choice. However, if there is less than a 10% difference in the standard error between the weighted and unweighted regressions, then the user should consult a statistician for the application of additional models, given the test data, or the data could be re-analyzed using the less-desirable linear interpolation (using ICPIN, see Section 4.8.2.2) method of analysis. This comparison between weighted and unweighted regression is completed for each of the selected models while proceeding through the process of final model selection (i.e., model and regression of best choice). Some non-divergent patterns might be indicative of an inappropriate or incorrect model (refer to Figure O.2C, Appendix O, in EC 2005b for an example), and the user is again urged to consult a 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).

4.8.2.2 Linear Interpolation Using ICPIN

If regression analyses of the endpoint data (see Section 4.8.2.1) fail to provide an acceptable ICp for growth (i.e., assumptions of normality and homoscedasticity cannot be met), linear interpolation using the computer program called *ICPIN* should be applied (see Section 6.4.3 in EC, 2005b). This program (Norberg-King, 1993; USEPA, 1995, 2002) is not proprietary, is available from the USEPA and is included in most computer software for environmental toxicology, including TOXSTAT (1996). The original instructions for ICPIN from USEPA are clearly written and make the program easy to use (Norberg-King, 1993).[59] An earlier version was called BOOTSTRP.

Analysis by ICPIN does not require equal numbers of replicates in different concentrations. The ICp is estimated by smoothing of the data as necessary, then using the two data-points adjacent to the selected ICp (USEPA, 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, and so Canadian users of the program must enter the concentrations as logarithms. Some commercial computer packages have the logarithmic transformation as a general option, but investigators should make sure that it is actually retained when proceeding to ICPIN. ICPIN estimates confidence limits by a special "bootstrap" technique since usual methods would not be valid. Bootstrapping performs many resamplings from the original measurements. The investigator must specify the number of resamplings, which can range from 80 to 1000. At least 400 is recommended here, and 1000 would be beneficial.[60]

If there are several adjacent high concentrations with no emerged plants, only the lowest of that string of concentrations should be used in the 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, since they offer nothing to the analysis (i.e., the data consist only of zero mean weights and zero mean lengths).

Besides determining and reporting the computer-derived ICps for length and weight of individual plants at test end, separate graphs of percent reduction for each of shoot and root lengths and shoot and root dry weights 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).

[58] The value of 10% is only a "rule-of-thumb" based upon experience. Objective tests for the improvement due to weighting are available but beyond the scope of this document. Weighting should be used only when necessary, as the procedure can introduce additional complications to the modeling procedure. A statistician should be consulted when weighting is necessary.

[59] The instructions in Norberg-King (1993) are sometimes misleading on the identity of "replicates." The term is used in such a way that it would apply to numbers or weights of individual organisms within the same 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.

[60] ICPIN has some deficiencies, which is why it is recommended 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 convert concentrations to a logarithm, which would introduce a slight bias towards a higher value of ICp. A modification of the bootstrap method has now remedied a problem of overly narrow confidence limits; however, regression analyses provide better methods of estimating the ICp and its 95% confidence limits (EC, 2005b) (see Section 4.8.2.1).

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 average of the replicate means) is higher ("better") than the average for the control. To apply this correction, replace the observed mean weights (or mean lengths) of the replicates in the hormetic concentration(s) with the means of replicates in the control. The geometric average for that/those concentration(s) will then be the same as that for the control.

4.9 Tests with a Reference Toxicant

Described herein are the procedures and conditions to be followed when performing reference toxicity tests in conjunction with a 28-, 35-, or 42-day test of *growth* using boreal forest plants. These procedures also apply to tests for assessing the acceptability and suitability of batches of seed for use in soil toxicity tests; and should be applied to assess intralaboratory *precision* when a laboratory is inexperienced with the biological test method defined in this document and is initially setting up to perform it (see Sections 2.5 and 3.2.1).

The routine use of a *reference toxicant* is necessary to assess, under standardized test conditions, the relative sensitivity of a lot of boreal forest plant seed being used. Tests with a *reference toxicant* also serve to demonstrate the *precision* and reliability of data produced by the laboratory personnel for that *reference toxicant*, under standardized test conditions. A reference toxicity test, conducted according to the procedures and conditions described herein, must be performed according to one of the following regimes:

(1) at least once every two months using the same lot of seed being used to provide test organisms for soil toxicity tests over an extended period (i.e., ≥ 2 months); or

(2) at the same time as the definitive soil toxicity test(s), using seed taken from the same lot number as those used for the definitive test(s) (see Section 2.5).

A laboratory that frequently performs soil toxicity tests using boreal forest plants might choose to routinely (e.g., every two months) monitor the sensitivity of their seed to one or more *reference toxicant*s, while including a reference toxicity test using a portion of the seeds used to start a definitive soil toxicity test. Alternatively, a laboratory might choose to monitor the sensitivity of their seed to a reference toxicant less frequently, and to perform a reference toxicity test at the time that each definitive soil toxicity test is performed.

Each reference toxicity test performed in conjunction with the *definitive* test for soil toxicity must be conducted as a *static* multi-concentration growth test. The duration of the reference toxicity test is 14 days if the species of organisms is bluejoint reedgrass or jack pine; 21 days if trembling aspen or Canada goldenrod; 28 days if paper birch; and 35 days if white spruce or black spruce. In each instance, the ICp for root length is determined at the end of the test. A summary checklist in Table 4 describes the conditions and procedures that 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 reference toxicity test should be performed using 1-L polypropylene containers as test vessels (Section 3.2.2) and a 500-mL aliquot of *test soil* representing each treatment (concentration) in each test vessel. The number of replicate test vessels per concentration must be ≥ 3. The number of seeds per vessel is species-specific, and are the same as those

specified for definitive tests. Reference toxicity tests with trembling aspen, birch, bluejoint reedgrass, and Canada goldenrod must include 5 seeds per vessel, whereas for white spruce, black spruce, and jack pine, 10 seeds per vessel are required (see Table 4).

Procedures for starting and ending a reference toxicity test should be consistent with those described in Sections 4.2 and 4.7. Test conditions for temperature and light, described in Section 4.3, apply. Observations and measurements should be as described in Section 4.6; however, individual root length should be determined at the end of the test. Percent emergence in the negative control soil is measured for the purposes of test validity.

To be valid, the mean percent *emergence* at the end of the test for plants held in the control treatment used in a particular reference toxicity test must be:

≥ 60% for any species.

Additionally, the mean root length for each plant species grown in negative control soil at test end must be:

≥ 30 mm for trembling aspen;
≥ 51 mm for bluejoint reedgrass;
≥ 47 mm for Canada goldenrod;
≥ 48 mm for paper birch;
≥ 38 mm for jack pine;
≥ 39 mm for white spruce; or
≥ 26 mm for black spruce.

The test endpoint to be calculated and reported must be the 14-, 21-, 28-, or 35-day ICp (including its 95% confidence limits) for root length. Results for a reference toxicity test should be expressed as mg reference chemical/kg soil, on a dry-wt basis.

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

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

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)[61] is recommended for use as the reference toxicant when performing soil toxicity tests with plants, although other chemical may be used if they prove suitable.[62] Each test concentration should be made up according to the guidance in Sections 4.1 and 6.2, using *artificial soil* (Section 3.4.2) as substrate.

Routine reference toxicity tests (e.g., those performed once every two months or in conjunction with each definitive test for soil toxicity) using boric acid spiked in *negative control soil* should consistently apply the same test conditions and procedures described herein. A series of test concentrations

[61] Boric acid has been used historically as a soil chemo-sterilant and has been found to be an effective biocide. Boric acid dissociates readily in water with neutral pH and is readily absorbed, accumulated, and translocated by the roots of plants. It is relatively persistent in soils in laboratory conditions and does not readily photodecompose or volatilize (Stephenson et al., 1997).

[62] Aquaterra Environmental (1998) initially evaluated the performance of boric acid as a candidate *reference toxicant* for use in conjunction with *acute toxicity* tests for measuring soil toxicity to terrestrial plant species. Subsequent studies by EC (2010, 2013b) and SRC (2006) using boric acid spiked in artificial soil confirmed the usefulness of boric acid as a suitable reference toxicant when performing bi-monthly reference tests with boreal forest plant species.

Table 4 Checklist of required and recommended conditions and procedures for conducting reference toxicity tests on soil using boreal forest plants

Test type	–	whole soil reference toxicity test; no renewal (static test)
Test duration	–	· 14 days for bluejoint reedgrass or jack pine · 21 days for Canada goldenrod or trembling aspen · 28 days for paper birch · 35 days for white spruce or black spruce
Approved test species	–	· gymnosperms: white spruce *(Picea glauca)*, black spruce *(Picea mariana)*, jack pine *(Pinus banksiana)* · angiosperm dicotyledons: trembling aspen *(Populus tremuloides)*, goldenrod *(Solidago canadensis)*, paper birch *(Betula papyrifera)* · angiosperm monocotyledons: bluejoint reedgrass *(Calamagrostis canadensis)*
Number of concentrations	–	minimum of 5, plus negative control
Number of replicates	–	≥ 3 replicates/concentration
Number of seeds per test vessel	–	5 seeds/vessel for trembling aspen, paper birch, Canada goldenrod or bluejoint reedgrass, and 10 seeds/vessel for black spruce, white spruce, or jack pine
Negative control soil	–	artificial soil
Test vessel	–	polypropylene cups (1 L), may be covered for full duration of the test, lids replaced by inverted test unit or other suitable container if plants reach top of cup
Amount of soil/test vessel	–	identical wet wt, equivalent to a volume of ~500 mL; ~350 g dry wt if artificial soil
Moisture content	–	for soil preparation, hydrate to 70% of water-holding capacity (WHC); during test, hydrate as needed
Air temperature	–	daily range, constant 24 ± 3°C; alternatively, day: 24 ± 3°C, night: 15 ± 3°C
Humidity	–	test area kept at ≥ 50%, optional
Lighting	–	full-spectrum fluorescent: mimic natural light spectrum (e.g, VitaLux ® by Duro-Test®); 300 ± 100 µmol/(m^2/s) adjacent to the level of the soil surface; 16 h light:8 h dark
Watering	–	hydration water sprayed conservatively over soil surface, as needed
Measurement during test	–	soil moisture content in each treatment/concentration at start; pH in each treatment/concentration at start and end; temperature in test facility, daily or continuously; light intensity once during test
Observations during test	–	number of emerged seedlings at end of test in each test vessel and root length at test end; number of emerged plants at test end showing an atypical appearance (e.g., chlorosis, lesions)
Biological endpoints	–	number of seedlings at end of test in each test vessel; length of longest root at test end
Statistical endpoints	–	mean (± SD) percent emergence in control soil (for test validity) at test end (Day 14, 21, 28, or 35); mean (± SD) length of longest roots in each treatment at test end (Day 14, 21, 28, or 35); 14-, 21-, 28-, or 35-day ICp for root length
Test validity	–	invalid if **any** of the following occurs in negative control soil at test end: · mean % emergence is < 60% · mean root length is: < 30 mm for trembling aspen < 51 mm for bluejoint reedgrass < 47 mm for Canada goldenrod < 48 mm for paper birch < 38 mm for jack pine < 39 mm for white spruce < 26 mm for black spruce

should be chosen,[63] based on preliminary tests, to bracket the ICp and enable calculation of the 14-, 21-, 28-, or 35-day ICp for root length.

Once sufficient data are available (EC, 1995), all comparable ICps for a particular *reference toxicant* derived from these toxicity tests must be plotted successively on a *warning chart*. A separate warning chart must be prepared for each plant species used in definitive toxicity tests. Each new ICp for the same reference toxicant should be examined to determine whether it falls within ± 2 SD of values obtained in previous comparable tests using the same reference toxicant and test procedure (EC, 2004a, 2005a, 2007a). A separate warning chart must be prepared and updated for each dissimilar procedure (e.g., differing plant species or differing reference toxicants). The warning chart should plot logarithm of concentration on the vertical axis against date of the test or test number on the horizontal axis. Each new ICp for the reference toxicant should be compared with established limits of the chart; the ICp is acceptable if it falls within the *warning limits*.

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

The mean of the available values of log (ICp), together with the upper and lower warning limits (± 2 SD), should be recalculated with each successive ICp for the *reference toxicant* until the statistics stabilize (EC, 1995, 2004a, 2005a, 2007a). If a particular ICp fell outside the warning limits, the sensitivity of the test organisms and the performance and precision of the test would be suspect. Since this might occur 5% of the time due to chance alone, an outlying ICp would not necessarily indicate abnormal sensitivity of the seed, or unsatisfactory precision of toxicity data. Rather, it would provide a warning that there might be a problem. A thorough check of all test conditions and procedures should be carried out. Depending on the findings, it might be necessary to repeat the reference toxicity test or purchase new seed before undertaking further soil toxicity tests.

Results that remained within the warning limits might not necessarily indicate that a laboratory was generating consistent results. Extremely variable historic data for a *reference toxicant* would produce wide warning limits; a new data point could be within the warning limits but still represent undesirable variation in test results. A CV of no more than 30%, and preferably 20% or less, has been suggested as a reasonable limit by Environment Canada (EC, 1995, 2005b) for the mean of the available values of log (ICp) (see preceding paragraph). For this biological test method, the CV for mean historic data derived for reference toxicity tests performed using boric acid should not exceed 30%.

[63] ASTM (2009) recommends a 0.5 dilution series for reference toxicity tests using boric acid. Specifically, ASTM (2009) recommends that a 640, 320, 160, 80, 40, 20, and 10 mg/kg soil dry weight dilution series will bracket the sensitivity of most plant species to boric acid. Environment Canada found that some species were not sensitive to concentrations as high as 640 mg boric acid/kg soil dry wt, and therefore recommended the inclusion of higher concentrations for reference toxicity tests using this chemical. A dilution series consisting of 2000, 1125, 630, 360, and 200 mg/kg soil dry wt should bracket the sensitivity of each of the test species recommended herein; however, the dilution series could be modified to suit each species being tested.

Environment Canada demonstrated endpoint values ranging from 212 mg boric acid/kg soil (dry wt) for paper birch to 592 mg boric acid/kg soil (dry wt) for jack pine in their results for 14-, 21-, 28-, or 35-day IC50s for root length with boric acid mixed in artificial soil using all seven plant species, and the test method for a reference toxicity test described herein (EC, 2013b). See Appendix G for guidance in selecting an appropriate series of test concentrations (assuming a log-concentration response) for use in toxicity tests with this or other chemicals to be used in a reference toxicity test.

As part of a series of interlaboratory studies performed to validate Environment Canada's reference toxicity test described in Section 4.9, six laboratories undertook concurrent five-week reference toxicity tests with black spruce and multiple concentrations of boric acid spiked in artificial soil. Each of the six participating laboratories achieved the interim validity criteria of ≥ 60% emergence and ≥ 26 mm root length established for black spruce seedlings grown in artificial soil for reference toxicity tests, and therefore data from all six laboratories were included in the final data analysis. Based on the data from all of the participating laboratories, the mean % emergence and the mean root length for the negative controls were 92% and 45 mm, respectively. The interlaboratory variability, expressed as the coefficient of variation (CV) was 7% for control emergence and 27% for root length. The mean five-week IC50 for seedling root length for boric acid in artificial soil was 555 mg H_3BO_3/kg dry wt, with values for individual laboratories ranging from 437 to 813 mg H_3BO_3/kg dry wt. The CV of 28% for these IC50s shows acceptable interlaboratory *precision* (EC, 2013a).

Concentrations of *reference toxicant* in all *stock solutions* should be measured chemically using appropriate methods (e.g., analytical methods involving AES with ICAP scan, for concentration of boron). Test concentrations of reference toxicant in soil are prepared by adding a measured quantity of the *stock solution* to negative control soil,[64] and mixing thoroughly.[65] 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.[66] Each aliquot should either be analyzed directly, or stored for future analysis (i.e., at the end of the test) if the ICp for root length, based on nominal concentrations, was found to be outside the warning limits. If stored, sample aliquots must be held in the dark at 4 ± 2ºC. Stored aliquots requiring chemical measurement should be analyzed promptly upon completion of the reference toxicity test. The 14-, 21-, 28-, or 35-day ICp for root length should be calculated based on the measured concentrations if they are appreciably (i.e., ≥ 20%) different from nominal ones and if the accuracy of the chemical analyses is satisfactory.

If boric acid is used as a *reference toxicant*, the following analytical method applies (OMEE, 1996). A 1 to 5 g subsample of soil spiked with boric acid is dried at 105ºC 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 re-adjusting the final volume to 50 mL using more extraction solution. The 50-mL extract is then filtered through a #4 Whatman filter, and diluted to a final volume of 100 mL. A blank sample is prepared in a similar manner.

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

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

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

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

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

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

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

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

Detailed guidance for the collection, handling, transport, storage, and 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 for toxicity tests with boreal plant species, described in this biological test method document.

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 terrestrial plants and/or other suitable, soil-associated test organisms (e.g., EC, 2004a, 2005a, 2007a) are frequently part of more comprehensive contaminated land assessments and *remediation* (e.g., Callahan *et al*., 1991; Menzie *et al*., 1992; Saterbak *et al*., 2000; Stephenson *et al*., 2008; EC, 2012). Such assessments often include a *test battery* 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 of the *risk* associated with soil contamination in *ecological risk assessment*s 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 taken quarterly, semi-annually, or annually from a number of contaminated or potentially contaminated sites for *monitoring* and *compliance* purposes. Samples of soil might also be collected on one or more occasions during field surveys of sites for spatial (i.e., horizontal or vertical) or temporal definition of soil quality. Increasingly, biological (toxicity) testing is being used in 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 for 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; background data collection; selection and location of sampling sites; selection of sampling strategies; *quality assurance* and *quality control* (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*) may 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), and briefly discussed in Section 4.1 herein; however, this biological test method document and the guidance provided herein applies 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 the 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 sampling location(s). These replicates samples[67] 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 can be tested for its toxicity to boreal forest plants as a single replicate (i.e., using only one test vessel per replicate sample) or as multiple replicates (i.e., using more than one test vessel per replicate sample; see

[67] *Replicate sample(s)* 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).

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). 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 3.6).[68] 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*, *conductivity*, and *fertility*. 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 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).

Field-collected soils or similar particulate material being considered for land disposal 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

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

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 horizons*. 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. Soil sampled for the assessment of effect(s) on boreal plants, 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 can 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 remove soil particles adhering to the roots. The soil sample to be collected for toxicity and evaluation of chemistry should be taken from one or more depths that represent the layer(s) of concern (e.g., a surficial layer of soil, or one or more deeper layers of soil or subsoil if there is concern about historical deposition of contaminants). 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 can vary 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 a sampling program is initiated. This calculation should take into account the quantity of soil required to prepare laboratory replicates for soil toxicity tests, as well as that required for particle size characterization, *total organic carbon* content (%), *organic matter* content (%), *moisture content* (%), and specific chemical analyses. 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 nontoxic, 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, to support the weight of the sample, and to maintain darkened conditions during sample transport (ASTM, 2009). Plastic containers or liners should not be used if there are concerns about the plastic affecting the characteristics of the soil (e.g., compounds from plastic leaching into the soil). For soils contaminated with volatile compounds, containers should be airtight and pressure resistant. 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 made at this time 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 conditions, 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 samples of soil 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 photoactivated 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 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* (i.e., 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 in airtight containers. If volatile contaminants are in the soil or of particular concern, any air "headspace" in the storage container should be purged with nitrogen gas before capping tightly. Samples 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. These storage conditions must be applied in instances in which PAHs or other light-sensitive contaminants are present, or if the samples are known to contain unstable volatiles of concern. It is also recommended that samples of soil or similar particulate material be tested as soon as possible after collection. The 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 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 collected as a distinct horizon should be thoroughly mixed (Section 5.3), and representative subsamples taken for physicochemical characterization. Each 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 (%),[69]
- *organic matter* content (%),[69]
- *pH*,
- *conductivity*,
- *moisture content* (%),
- *water-holding capacity*,
- nitrogen as total N, nitrate (NO_3^-), nitrite (NO_2^-), and ammonium (NH_4^+),
- phosphorus as total or plant-available,
- potassium as total or plant-available,
- C:N ratio, and
- *cation exchange capacity*.

Additionally, the following analyses should be performed:

- major cations, and anions (Na^+, K^+, Mg^{2+}, Ca^{2+}, Al^{3+}, S^{2-}, Cl^-).

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, and
- characteristics of the contamination (e.g., odour, staining, debris, presence of fuel or solvent).

In order to confirm that the reference soils are not contaminated, the following screening analyses are recommended:

- organophosphorous insecticide suite,
- organochlorine insecticide suite,
- herbicides suite,
- metals suite,

[69] 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.

- petroleum hydrocarbons (including PAHs),
- other site- or area-specific contaminants of concern.

Unless indicated otherwise, identical chemical, physical, and toxicological analyses should be performed with subsamples representative of each replicate sample of field-collected 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 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 pressing 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., 3 to 5 mm clumps). Soil samples comprised 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 may 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 should be documented and must be reported.

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* are collected as separate components of a soil sample and therefore must be tested as the re-assembled soil profile within a single test unit. 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 whether to conduct plant toxicity testing on this horizon alone or in the re-assembled soil horizons from this site.

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).[70] 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 pre-calculated 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 nontoxic device such as a stainless steel spoon) or mechanically (e.g., using a domestic hand-held mixer with beaters at low speed or a hand-held wire egg beater) until its texture and colour are homogeneous. 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 appears to be homogeneous, mixing should be discontinued.

For each 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.

[70] 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.

As indicated in Section 3.7, one or more samples of field-collected *test soil* might either be tested at a single concentration only (typically, 100%), or evaluated for toxicity in a multi-concentration test whereby a series of concentrations are prepared by mixing measured quantities with either negative control soil or reference soil. When performing a multi-concentration test, the following series of concentrations of test soil (mixed in negative control soil or reference soil), which spans the range of 100% to 1% test soil using nine concentrations, might prove suitable: 100%, 80%, 65%, 50%, 30%, 15%, 7.5%, 3%, 1%, and 0%. Guidance on other concentration series that might prove as or more suitable is found in Section 6.2, along with that for preparing test mixtures 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 comprised solely of negative control soil (see Section 3.4).

As described in Section 4.1, separate horizons of test and reference soils are reconstructed in layers in test vessels. The depth to which each horizon is layered in a test vessel is study-specific and, where possible, should correlate proportionally to the depths of each horizon as collected in the field; correlate to the actual field depths, if the horizon depths in the field are very shallow. To construct a test unit with multiple horizons, each horizon is prepared individually (dried, sieved, homogenized) and then placed into a test unit. Each subsequent layer is placed on the previous layer carefully so as to avoid inadvertent horizon mixing. The total volume of soil in each test vessel is still ~500 mL (i.e., a wet weight equivalent to a volume of ~500 mL), but it is composed of various horizons of test soil. A suitable wet weight of each horizon is layered in proportion to the depths of horizons in the field or according to specific study objectives (EC, 2012).

For a multi-concentration test, separate horizons of test and reference soils are used to prepare test dilutions. Each horizon of test soil is mixed with the same horizon of reference or negative control soil (see Section 4.1) at the appropriate test concentration (e.g., 25%). In some cases, it may not be possible to collect the same horizons of reference or negative control soil and test soil. For example, preliminary remedial action may have already been taken at the test site, resulting in disturbed or mixed natural soil horizons. In these scenarios, the soil sample can be tested as a mixed soil without an attempt to re-layer the different horizons. Alternatively, test concentrations can be prepared by mixing suitable weights of test soil into the available horizons of reference or negative control soils at the appropriate test concentration (e.g., 25%). These mixtures are then layered in the test vessels at depths appropriate to represent field conditions and to fulfill study objectives. 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 soils, if possible.

The *moisture content* of a given sample of a field-collected soil horizon 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 horizon must be determined before sample preparation and test initiation. In order to do so, the moisture content of each homogenized sample (i.e., each sample of test soil horizon, 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 is then hydrated to a homogeneous, crumbly consistency with clumps approximately 3 to 5 mm in diameter.[71] 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 higher WHC than mineral horizons, so they will require greater amounts of water to hydrate to a moist, crumbly texture.

Based on the initial moisture content of each separately collected horizon, the WHC of the horizon, and the amount of water added to achieve

[71] The optimal moisture content was determined for each of the diverse types of soil used while developing the biological test method described herein (see Section 3.4 and Appendix F) and is based on a percentage of each sample's WHC. These values were considered optimal since, at these levels of saturation, the soil mixed well, and formed an acceptable structure (i.e., the resulting macro-aggregation of soil particles).

the desired soil consistency, the optimal moisture content can be calculated and expressed as a percentage of the WHC for each horizon. Once this target (or optimal) percentage of the WHC has been determined, the moisture content of each sample of test soil horizon (including the negative control soil) can be standardized to the selected (sample-specific) moisture content. *Test water* (i.e., de-ionized or distilled water) should be added to each sample with a moisture content that is less than the pre-determined optimal percentage of its WHC, until this moisture content is achieved (Aquaterra Environmental, 1998).[72, 73] If a sample is too wet, it should be spread as a thin layer on a clean sheet of plastic (e.g., a new plastic garbage bag) or a clean, non-reactive (e.g., stainless steel or plastic) tray, and allowed to air-dry by evaporation at ambient (~20°C) room temperature. Rehydration to the pre-determined optimal percentage of its WHC might be necessary. Upon adjustment of a sample's moisture content to the desired percentage of its WHC, the moisture content (%) of the hydrated soil must be determined and the percent WHC and percent moisture content recorded and reported.

The WHC (and the percent WHC that is optimal for biological testing) of a particular soil horizon is generally unique to each type of soil 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).[74] This method is outlined here.

For this method, ~130 g (wet wt)[75] of sample is placed into an aluminum pan or petri dish (15 × 1 cm), and dried at 105°C until a constant weight is achieved (this usually takes a minimum of 24 h). Cool the soil 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 de-ionized water. The resulting slurry is mixed thoroughly with a glass stir rod. A folded filter paper (185-mm diameter Fisherbrand P8 coarse porosity, qualitative creped filter paper; catalogue no. 09-790-12G) is placed into a glass funnel (with a top inside diameter of 100 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 de-ionized water is slowly added to the filter paper to wet the entire surface. The funnel and hydrated filter paper are then weighed. To obtain the initial weight for the mass of the funnel plus hydrated filter paper plus dried soil (see "I" in Equation 1), the weight of the dried soil (100 g) is added to the weight of the funnel and the wet filter paper.

The funnel is then placed into a 500-mL Erlenmeyer flask and the soil slurry is slowly poured onto the hydrated filter paper held in the funnel. Any soil remaining on the beaker and stir rod is rinsed into the funnel with the least amount of water necessary to ensure that all of the solid material has been washed onto the filter. The funnel is then tightly covered

[72] 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% to 45% of its dry weight (ASTM, 2004; EC, 2000). 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% to 45% of their dry weight, whereas other *site soils* can appear considerably dryer after the same level of hydration (ASTM, 2004; EC, 2000). Accordingly, the use of this alternate approach is not recommended here.

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

[74] Certain participants at a soil toxicity testing workshop sponsored by EC in Vancouver, B.C. (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 EC's soil toxicity test methods when adjusting (if and as necessary) the moisture content of soil samples (Becker-van Slooten *et al*., 2004).

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

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 percentage of soil dry mass, is then calculated using the following equation:

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

where:

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

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

The percentage of water (i.e., P_w) that is added to a sample of soil horizon to achieve the desired hydration (i.e., the optimal percentage of the WHC) can be calculated as follows:[76]

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

where:

P_W = percentage of water to add to the soil horizon (%)
WHC = water-holding capacity (%)
MC_i = initial moisture content of the soil horizon

---

[76] The following example provides calculations that pertain to the hydration of samples of a contaminated field-collected soil horizon and a reference soil horizon, when preparing a test concentration of 25% for use in a definitive test with plants involving three replicates per treatment. The total volume of soil for preparation, in this example, is for a single soil horizon occupying a 250-mL volume of the total 500-mL volume in a test vessel [i.e., the balance of the volume in the test vessel would be comprised of another one or two soil horizon(s), calculated separately].

Assumptions:

Soil #1: Reference (r) Soil
W_r = 4.1857 g
D_r = 4.0402 g
WHC_r = 51.80%
P_{WHCr} = 55.00%
MC_r = 3.60%
P_{Wr} = 24.89%
M_{Dr} = 502.50 g dry wt
V_{Wr} = 125.07 mL
M_{Wr} = 520.60 g wet wt

Soil #2: Contaminated (c) Soil
W_c = 6.0779 g
D_c = 5.7978 g
WHC_c = 50.60%
P_{WHCc} = 42.50%
MC_c = 4.83%
P_{Wc} = 16.68%
M_{Dc} = 167.50 g dry wt
V_{Wc} = 27.94 mL
M_{Wc} = 175.59 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 reference soil:

For a definitive plant test using this example, it is assumed that a total mass of 670.00 g dry weight of one soil horizon is sufficient to satisfy the requirement for each treatment (i.e., 215.00 g dry wt per replicate × 3 replicates + 25.00 g dry wt extra soil for pH and conductivity); similar calculations would be used to add additional soil horizon(s) for that sample such that the total volume in each test vessel would be equivalent to a 500-mL aliquot of soil (see Section 4.1).

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

670.00 g dry wt × (25/100)
= 167.50 g dry wt of contaminated soil

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

670.00 g dry wt × (75/100)
[or 670.00 g dry wt – 167.50 g dry wt]
= 502.50 g dry wt of reference soil

Therefore, the final total mass of soil required, based on wet weight, is 645.67 g [520.60 g wet wt at the soil's initial moisture content (i.e., M_{Wr}) + 125.07 mL of water] for the reference soil, and 203.53 g [175.59 g wet wt at the soil's initial moisture content (i.e., M_{Wc}) + 27.94 mL of water] for the contaminated soil.

The final moisture content for each soil would be 28.49% {[(645.67 – 502.50) / 502.50] × 100} for the reference soil, and 21.51% {[(203.53 – 167.50) / 167.50] × 100} for the contaminated soil.

The final moisture content of the reference soil (i.e., 28.49%) represents 55% of that soil's water-holding capacity (28.49 ÷ 51.80 = 0.55). The final moisture content of the contaminated soil (i.e., 21.51%) represents 43% of that soil's water-holding capacity (21.51 ÷ 50.60 = 0.43).

The volume of water (i.e., V_w) that should be added to a horizon of field-collected soil to achieve the desired hydration (i.e., the optimal percentage of the sample horizon's water-holding capacity) can be calculated as follows (see footnote 76):

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

where:

V_W = volume of water to add to the soil horizon (mL)
P_W = percentage of water to add to the soil horizon (%)
M = total mass of soil horizon required for test (expressed as dry wt)[77]

Environment Canada (2012) describes various procedures that might be used for manipulating 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. 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. Studies intending to investigate the effect of a soil manipulation (e.g., pH adjustment) on sample toxicity should conduct 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 subsamples for subsequent physicochemical analyses. Any remaining portions of the homogenized sample that might be required for additional toxicity tests using plants or other test organisms (e.g., according to EC, 2004a, 2005a, and 2007a) should also be transferred to labelled containers at this time. All subsamples to be stored should be held in sealed containers with minimal air space and must be stored in darkness at 4 ± 2°C (Section 5.2) until used or analyzed. Just before it is analyzed or used in the toxicity test, each subsample must be 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 for 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, 2007a) 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

[77] In tests with samples of field-collected soil horizons, the amount of the soil horizon added to each test vessel is based on a volume of soil that will produce a layer in the test vessel correlating proportionally to the depth of that horizon as collected in the field. The amount of soil added to each test vessel is based on the wet weight of that soil that is equivalent to a proportion of the total volume of ~500 mL (see Section 4.1). However, "M" (i.e., the total mass of soil required for the horizon) is expressed as dry weight in the formula used to calculate the volume of water to be added to a horizon of field-collected soil to achieve the desired hydration (see Equation 3). To calculate the amount of soil horizon required per test vessel (by dry wt), a subsample of "wet" soil is placed into a test vessel (e.g., 1-L polypropylene cup), to determine the correct volume of soil horizon required on a wet-weight basis. For example, assume that (for a given sample) this volume is equivalent to 300 g wet wt and that the wet and dry weights of a subsample of this soil, previously determined for the purpose of calculating the sample's water-holding capacity, are 4.1857 g and 4.0402 g, respectively. The dry weight equivalent to a 300-mL volume of this sample (which has a wet weight of 300 g) can be calculated as follows:

$(300 \text{ g} \times 4.0402 \text{ g}) \div 4.1857 \text{ g} = 289 \text{ g}$

This mass of soil can be rounded up to 300 g dry weight, thereby providing a little extra soil, if necessary. Therefore, for the example provided here, the mass of this sample of soil horizon required for each replicate (expressed as dry wt) is 300 g. The total mass ("M") can then be calculated simply by multiplying the dry mass required for each replicate (in this instance, 300 g dry wt) by the number of replicates to be used in the test (i.e., for this example, three replicates).

require special procedures for sampling, handling, transport, storage, and preparation. These soils include: boreal forest soils, stony/shallow soils, organic soils, cryosolic soils, and wetland soils, and are relevant for use with the test methodologies described in this boreal plant 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 for 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 soil horizon should be made at the time the test is set up. This might include observations of sample colour, texture, and homogeneity, and the presence of plants or macroinvertebrates. Any changes in the appearance of the test soil 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 at the beginning, during, or at the end of each test. These observations and measurements apply and must be made when performing the soil toxicity test described herein using one or more samples of field-collected (site) soil.

Depending on the test objectives and experimental design, additional test vessels might be set up at the beginning of the test (Section 4.1) to monitor soil chemistry. These could 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 horizons within these vessels can 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 soil sample should be used for comparative purposes whenever possible or appropriate, because this provides a site-specific evaluation of toxicity (EC, 1997a, b, 2004a, 2005a, 2007a). 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 noncontaminant effects caused by soil physicochemical properties such as particle size, *total organic carbon* content (%), and *organic matter* content (%).[78] Regardless of whether the *reference soil* or *negative control soil* is used for the statistical comparisons, the results from *negative control soil* must be used to judge the validity and acceptability of the test (see Section 4.4).

Analyses of the results will differ according to the purposes and particular designs of the test. This section covers the analytical procedures, starting with the simplest design and proceeding to the more complex designs. Standard statistical procedures are generally all that are needed for analyzing the results. Investigators should consult EC (2005b) for guidance on the appropriate statistical endpoints and their calculation. As always, the advice of a statistician familiar with *toxicology* should be sought for the design and analyses of tests.

Analysis of variance (ANOVA) involving multiple comparisons of endpoint data derived for single-concentration tests involving replicate samples of field-collected soil from more than one *sampling location* is commonly used for statistical interpretation of the significance of findings from soil toxicity tests. This hypothesis-testing approach is

[78] Note that this may result in a test design that mixes *replicate vessels* (control soil) with *replicate samples* (from test sampling locations) and may oblige the investigator to treat replicate vessels (laboratory replicates) as equivalent to replicate samples (field replicates). While this is not appropriate statistically, it may need to be considered acceptable, given the lack of reasonable alternates. If inferences drawn from the analysis are deemed to be of high-impact (e.g., cleanup criteria), a statistician should be consulted.

subject to appreciable weaknesses. Notably, any increased variability within the test will weaken its power to distinguish *toxic* effects (i.e., less toxicity is concluded). Similarly, use of only a few replicates instead of many replicates will weaken the discrimination of a test and will lead to a conclusion of less apparent toxicity, other things being equal (see Section 5.6.2). There is no alternative to hypothesis testing, when comparing toxicity data for multiple samples of field-collected soil (i.e., field replicates of soil from more than one sampling location) that use only one concentration (usually full strength, i.e., 100% sample). There are alternatives for comparing point estimates of toxicity if various concentrations of each sample of field-collected soil are tested and multiple endpoint values for ICp are determined (see Section 6.4). Section 9 in EC (2005b) should be consulted for guidance when comparing multiple ICps.

The parametric analyses involving ANOVA for comparative data from single-concentration tests with multiple samples of field-collected soil assume that the data are normally distributed, that the treatments are independent, and that the variance is homogeneous among the different treatments. As the first step in analysis, these assumptions should be tested using the *Shapiro-Wilk's Test* for *normality* and *Bartlett's Test for Homogeneity of Variance* (Eisenhart *et al.*, 1947; Sokal and Rohlf, 1969). If the data satisfy these assumptions, analysis may proceed. If not, data could be transformed [e.g., as square roots, logarithms (Mearns *et al.*, 1986)]. The tests for normality and homogeneity might then show conformance with normality and homogeneity, and in fact that is a likely outcome of a transformation. Assumptions should be re-tested following any transformation of data. Parametric tests are reasonably robust in the face of moderate deviations from normality and equality of variance; therefore, parametric analysis (e.g., ANOVA and multiple comparison) should proceed, even if moderate nonconformity continues after transformation. Excluding a data set for minor irregularities might lose a satisfactory and sensitive analysis and forgo the detection of real effects of toxicity.[79] Analysis by nonparametric statistical procedures should also proceed in parallel, with the more sensitive (lower endpoint) of the two analyses providing the final estimates of toxicity. Section 3 in EC (2005b) should be consulted for guidance when comparing the findings for single-concentration tests involving field replicates of samples from multiple locations, using parametric or non-parametric tests.

Guidance in Section 6 (including that in Section 6.2 for performing range-finding tests, and that in Section 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

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

If test results at a single test *sampling location* are to be compared with test results at a reference sampling location, a *t-test*[80] 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 (and non-parameteric equivalents) exist (Section 3.3 in EC, 2005b). Choice of a specific test depends on:

(1) the type of comparison that is sought (e.g., complete series of pairwise comparisons

[79] Tests for normality and homogeneity become less meaningful with the small number of samples from individual *sampling locations* typically collected in studies of *environmental toxicology*. Plotting and examining the general nature of the distribution of toxicity and its apparent deviations can be more revealing and is recommended (EC, 2005b). Equality in sample sizes and the magnitude of variation are probably more important factors for the outcome of parametric analysis, but they have received scant attention in toxicology. The robustness of ANOVA is shown by its ability to produce realistic probabilities if the distribution of data is reasonably symmetrical, and if treatment variances are within threefold of each other (Newman, 1995).

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

between all sampling locations, or compare the response from each sampling location only with that of the reference site);

(2) if a chemical and/or biological response gradient is expected;[81] and

(3) if the assumptions of *normality* and *homoscedasticity* are met.

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. Opportunities for statistical analysis would be limited. The nonreplicated test data could be compared with the reference data using outlier detection methods (EC, 2005b). A sample would be considered *toxic* if its result was rejected as an extreme value when considered as part of the data for the *reference soil* and/or the *negative control soil*.

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

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

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

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

[81] 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.

which locations were particularly deserving of cleanup. An ANOVA would first be conducted to test for overall differences. A *post-hoc* test such as *Tukey's test* could then be used for such an analysis; this test is commonly found in statistical packages and can deal with unequal sample sizes.[82] Sections 3.3 and 7.5 in EC (2005b) provide further details, alternate tests and non-parametric options, and the guidance therein should be followed.

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

5.6.2 Power Analysis

An important factor to consider in the analysis of the results for toxicity tests with soil is the potential for declaring false positives (i.e., calling a *clean* site *contaminated*; Type I error) or false negatives (i.e., calling a *contaminated* site *clean*; Type II error).

Scientists are usually cautious in choosing the level of significance (α) for tolerating false positive results (Type I error), and usually set it at P = 0.05 or 0.01. Commonly, scientists following a specified test design will never consider the relationship between power, variability, and effect size, leaving the Type II error completely unspecified. There are several factors that influence statistical power, including:

- variability of replicate samples representing the same treatment;
- α (i.e., the probability of making a Type I error);
- *effect size (ES)* (i.e., the magnitude of the true effect being tested); and
- *n* (i.e., the number of samples or replicates used in a test, and in some cases, the allocation of those replicates[83]).

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

Power analysis can be used *a priori* to determine the magnitude of the Type II error and the probability of false positive results. It can also be used to ascertain the appropriate number of field and laboratory replicates for subsequent surveys involving this test, or to assist in the selection of future sampling sites. It is always prudent to include as many replicates in the test design as is economically and logistically warranted (see Section 5.1); power analysis will assist in this determination. A good explanation of the power of a test, and how to assess it, can be found in USEPA (2000). Guidance on power analysis is also provided in EC (2005b).

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

[82] An alternative approach is available (EC, 1997a, b, 2004a, 2005a, b 2007a). For equal replicates, *Fisher's Least Significant Difference (LSD)* is recommended. It is based on a smaller "pairwise error rate" for α in comparing data for samples from any given location with those for samples from another location, but holds the overall value of α to the pre-selected value (usually 0.05). *LSD* is seldom included in software packages for toxicity, but it is described in some textbooks (e.g., Steel and Torrie, 1980). Instead, *Tukey's test* is recommended here, partly because *LSD* might declare significant differences too readily. *LSD* is also intended for only a few of all the possible comparisons in a set of data, and those comparisons would have to be specified in advance.

[83] If the experimental design requires the comparison of test sampling locations with the reference sampling location only (e.g., using *Dunnett's* test or *Williams'* test), optimal power for the final length or dry weight endpoints is achieved by allocating a higher number of *replicate samples* at the reference sampling location (Dunnett, 1955; Williams, 1972; OECD, 2006b). As a general rule, the number of replicate samples at the reference sampling location (n_o) can be related to the number of test sampling locations (k) and the number of replicate samples at each test sampling location (n) using: $n_o = n\sqrt{k}$ for *Dunnett's* test (OECD, 2006b). 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 replicate samples (i.e., field replicates). If the investigator was interested in increasing the number of replicate samples beyond the minimum, extra replicate samples should be allocated to the reference sampling location to maximize power and minimize Type II error. As an example using *Dunnett's* formula, consider an experiment with reference sampling location and four test sampling locations, and each test sampling location with five replicate samples. To maximize power, the optimal number of replicate samples at the reference sampling location would be $n_o = n\sqrt{k} = 5x\sqrt{4} = 10$ replicates.

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).[84] 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 may specify the expected effect size (e.g., 25%) to be detected (AE, 2007c).

[84] 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. We include it in discussions 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).

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) for *spiking* negative control sediment with chemical(s) and conducting toxicity tests with chemical/sediment mixtures is also relevant here, for *chemical-spiked soil*. Further evaluation and standardization of procedures for preparing chemical-spiked soil provided herein (Section 6.2) might be required before soil toxicity tests with plants 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 3.4) with these substances. The *spiking* might be done with one or more chemicals or chemical products. Other options for toxicity tests with boreal plants, performed using the procedures described herein, include the *spiking* of chemical(s) or chemical product(s) in *reference soil* (Section 3.6) or *test soil* (Section 3.7). 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 re-stratified in each test vessel (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 estimate statistical endpoints based on threshold concentrations causing specific *sublethal effects* (see Section 4.8).

Procedures are described in Section 6.2 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, and Sections 4.8 and 6.4 provide 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 dredged material intended for land disposal) in negative control soil or reference soil, and to performing multi-concentration tests and determining statistical endpoints for these mixtures (see Section 5, and especially 5.6). Multi-concentration tests with *positive control soil* (Section 3.5) or one or more *reference toxicants* 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.[85] Information should also be obtained for individual chemicals or chemical products (e.g., pesticides or other commercial formulations) on their concentration of major or "active" ingredients and impurities, water solubility, vapour pressure, chemical stability, dissociation constants, adsorption coefficients, toxicity to humans and terrestrial organisms, and biodegradability. Where aqueous solubility is in doubt or problematic, acceptable procedures previously used for preparing aqueous

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

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 of 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 Material Safety Data Sheets (MSDSs) should be obtained and reviewed.

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

6.2 Preparing Test Mixtures

On the day of the start of the toxicity test (i.e., Day 0), the mixture(s) of chemical(s) or chemical product(s) spiked in *negative control soil* should be prepared and transferred to test vessels (see Sections 4.1 and 4.2). Each *batch* of *test soil* horizon representing a particular treatment (concentration) should be prepared in a quantity sufficient to enable all test replicates of that treatment (concentration) to be set up along with any additional replicates or quantities required for physicochemical analyses (Section 6.3) or the performance of other soil toxicity tests using earthworms or other soil organisms (e.g., those performed according to EC, 2004a, 2005a, or 2007a).

Natural control soil (Section 3.4.1) is recommended for use 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 a definitive test using boreal plants. An investigation should include spiking of separate soil horizons of a natural control soil or reference soil (if collected) followed by re-layering of the horizons in test vessels prior to testing (Sections 4.1, 5.1, and 5.3). The depth to which each horizon is layered in a test vessel is study-specific and, where possible, should correlate proportionally to the depths of each horizon as collected in the field; correlate to the actual field depths, if the horizon depths in the field are very shallow. To construct a test unit with multiple horizons, each horizon is treated as a separate soil and prepared individually (dried, sieved, homogenized). The moisture content, WHC and optimal percentage of the WHC of each soil horizon must be determined separately. The final moisture content of each horizon of chemical-spiked soil (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) and control soil, prepared using field-collected soil, should be adjusted to the optimal percentage of its WHC using guidance in Section 5.3. Each soil horizon is then spiked separately to the appropriate test concentration for a given treatment before being layered into a test unit. All horizons must be spiked to the same test concentration. Each subsequent layer is placed on the previous layer carefully so as to avoid inadvertent horizon mixing; however, depending on the study objectives, the horizons might be tested separately. The total volume of soil in each test vessel is still ~500 mL (i.e., a wet weight equivalent to a volume of ~500 mL), but it is composed of various horizons of soil.

The volume of soil in each test vessel might differ, due to differences in bulk density of the various soils that might be used.

Artificial soil is recommended for use in reference toxicity tests (see Section 4.9). The quantity of artificial soil required for the reference toxicity test(s) should be prepared, hydrated to ~20% moisture content, adjusted if and as necessary to a pH within the range of 6.5 to 7.5, aged for a minimum three-day period, and stored at 4 ± 2°C until required (see Section 3.4.2). The final *moisture content* [including that due to the addition of a measured aliquot of a reference toxicant (e.g., boric acid) dissolved in *test water*] of *chemical-spiked soil* prepared for reference toxicity test using artificial soil should be ~70% of the water-holding capacity of

the final mixture (Section 3.4.2), for each treatment (concentration).[86] The final moisture content of each mixture (treatment) included in a reference toxicity test should be as similar as possible.

The procedure to be used for experimentally *spiking* soil is contingent on the study objectives and the nature of the test substance to be mixed with negative control soil or other soil. In many instances, a chemical/soil mixture is prepared by making up a *stock solution* of the test chemical(s) or chemical product(s) and then mixing one or more measured volumes into negative control soil (Section 3.4). The preferred solvent for preparing stock solutions is *test water* (i.e., de-ionized or distilled water); use of a solvent other than 100% test water should be avoided unless it is absolutely necessary. For test chemical(s) or chemical product(s) that do not dissolve readily in test water, a suitable water-miscible organic solvent of relatively low toxicity (e.g., acetone, methanol, or ethanol) may be used in small quantities to help disperse the test substance(s) in water. Surfactants should not be used.

If an organic solvent is used, the test must be conducted using a series of replicate test vessels containing only negative control soil (i.e., *clean* soil containing no solvent and no test substance), as well as a series of replicate test vessels containing only *solvent control soil* (OECD, 2006a; ASTM, 2009; EC, 2000). For this purpose, a *batch* of solvent control soil must be prepared containing the concentration of the solubilizing agent that is present in the highest concentration of the test chemical(s) or chemical product(s) in soil. Solvent from the same *batch* used to make the stock solution of test substance(s) must be used. Solvents should be used sparingly because 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 *emergence* or *growth* of plants during the test. If this information is unknown, a preliminary *solvent only* test, using various concentrations of solvent in negative control soil, should be conducted to determine the threshold-effect concentration of the particular solvent being considered for use in the definitive test.

[86] The following example provides calculations that show the volume of both water (de-ionized or distilled) and a stock solution of a reference toxicant (boric acid) to be added to a sample of artificial soil with an existing moisture content, to create a treatment with a moisture content that is 70% of the WHC for the artificial soil. The calculations take into account the volume of a stock solution of boric acid added when preparing the treatment, as part of the overall adjustment for soil moisture content. To simplify the calculations, this example assumes that 400 g (dry wt) of artificial soil (AS) is sufficient to provide the 500-mL aliquot of soil to be added to each test vessel when performing a reference toxicity test involving three 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) = 3.2486 g
Dry mass of AS = 2.6924 g
Moisture content (MC) of AS
= [(3.2486 – 2.6924)/2.6924] × 100
= 20.66% (initial moisture content)

Water-holding capacity (WHC) of AS = 72.10%
Percentage of WHC desired (P_{WHC}) = 70.00%

Dry mass of AS required for test (M_D)
= [400.00 g per rep × 3 reps] + 25.00 g extra
= 1225.00 g dry wt

Wet mass of AS required for test (M_W)
= (1225.00 × 3.2486)/2.6924
= 1478.06 g wet wt

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

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

The amount of boric acid required, on a dry-mass basis is:

H_3BO_3 = (2 g H_3BO_3/1000 g soil dry wt) × 1225.00 g dry wt
= 2.45 g H_3BO_3

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

H_3BO_3 = 2.45 g H_3BO_3/(25 g H_3BO_3/1000 mL of water)
= 98.00 mL stock solution

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

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

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

V_W = (P_W × M_D)/100
= (29.81 × 1225.00 g dry wt)/100
= 365.17 mL of water required

However, as part of this required volume, 98.00 mL of the stock solution is to be added for dosing; therefore, an additional volume of water of only 267.17 mL will be required (365.17 mL of water – 98.00 mL of stock solution).

Accordingly, the final total mass of soil required, based on wet weight, would be 1843.23 g [1478.06 g wet wt at the soil's initial moisture content (i.e., M_W) + 267.17 mL of water + 98.00 mL of stock solution], and the final moisture content of the soil, based on dry weight, would be 50.47% {[(1843.23 – 1225.00)/1225.00] × 100}.

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

Concentrations of chemical(s) or chemical product(s) in soil are usually calculated, measured, and expressed as mg test substance/kg soil (or µg substance/g soil) on a dry-weight basis (OECD, 2006a; ISO, 2012a, b). The assessment endpoints (e.g., ICps) are similarly expressed on a dry-weight basis (Section 4.8).

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

For some studies, it might be necessary to prepare only one concentration of a particular mixture of negative control (or other) soil and chemical(s) or chemical product(s), or a mixture of only one concentration of *contaminated* soil or particulate waste in negative control or other soil. For instance, a single-concentration test might be conducted to determine whether a specific concentration of chemical or chemical product in *clean* soil is *toxic* to the test organisms. Such an application could be used for research or regulatory purposes (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(s) (i.e., ICp; see Section 4.8) 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 nine test concentrations plus the appropriate control treatment(s) must be prepared for each multi-concentration test, and more (i.e., ≥ 11 plus controls) are recommended (see Sections 4.1 and 4.8). An appropriate geometric dilution series may be used when selecting test concentrations, in which each successive concentration of chemical(s) or chemical product(s) in soil is at least 50% of the previous one (e.g., 160, 80, 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 G) or may be derived based on the findings of preliminary "range-finding" toxicity tests.

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 toxicity assays or research studies.

Based on the objectives of the test, it might be desirable to determine the influence of substrate characteristics (e.g., particle size or organic matter content) on the toxicity of chemical/soil mixtures. For instance, the influence of soil particle size on chemical toxicity could be measured by conducting concurrent multi-concentration tests with a series of mixtures comprised of the test chemical(s) or chemical product(s) mixed in differing fractions (i.e., segregated particle sizes) or types of negative *control soil* (Section 3.4). 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 negative control soil used to prepare these mixtures should be included as a separate control in the test.

6.3 Test Observations and Measurements

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

Section 4.6 provides guidance and requirements for the observations and measurements to be made at the beginning, during, and at the end of the test.

These observations and measurements apply and must be made when performing the soil toxicity test described herein using one or more samples of chemical-spiked soil. 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 at the beginning of the test (see Section 4.1) to monitor soil chemistry. These would be destructively sampled during or at the end of the test. Test organisms might or might not be added to these extra test vessels, depending on study objectives. Measurements of chemical concentrations in the soil horizons 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 horizon of *spiked soil* being tested (including the *negative control soil*) must be made and recorded at the beginning and end of the test for pH and at the beginning of the test only for moisture content, 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 horizon 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;[87] in which instance, endpoint values calculated (Sections 4.8 and 6.4) would be based on nominal ones. Any 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 for calculating an ICp (based on data showing growth inhibition; see Section 4.8) is presented in Section 4.8.2. Section 5.6 provides guidance for calculating and comparing endpoints for single-concentration tests using samples of field-collected soil, which applies equally to single-concentration tests performed with mixtures of spiked soil. For further information on these or other appropriate parametric (or nonparametric) statistics to apply to the endpoint data, the investigator should consult the Environment Canada report on statistics for the determination of toxicity endpoints (EC, 2005b).

For any test that includes *solvent control soil* (see Section 6.2), the test results for plants held in that soil and in the *negative control soil* must be examined to determine if they independently meet the test validity criteria (see Section 4.7). 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 for the two controls must be statistically compared to each other using a *Student's t-test*. If the results for the two controls are not statistically different from each other, then only the data from the *negative control soil* should be used to calculate the test results.[88] If, however the final shoot/root length or dry weight in the solvent control differs significantly from the results of the clean control soil, this might be indicative of a potential solvent interference which would then require additional evaluation to determine the impact on the validity of the study. The USEPA (2008) provides guidance on what might be included in such an evaluation:

[87] 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.

[88] The solvent control is not favoured for the calculation of test results by the USEPA because it requires the assumption that the effects of the solvent and toxicant are independent of one another, and the current experimental designs do not allow this assumption to be tested (K. Sappington, Office of Pesticide Programs, USEPA, Washington, DC, written communication, 2012).

(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 (i.e., are the responses different for both endpoints or just one?); (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. If a solvent interference is identified, then the solvent control should be used as the basis for calculating results.

Reporting Requirements

Each test-specific report must indicate if there has been any deviation from any of the *must* requirements delineated in Sections 2 to 6 and, if so, provide details of the deviation(s). The reader must be able to establish from the test-specific report whether the conditions and procedures preceding and during the test rendered the results valid and acceptable for the use intended.

Section 7.1 provides a list of items that must be included in each test-specific report. A list of items that must either be included in the test-specific report, provided separately in a general report, or held on file for a minimum of five years is found in Section 7.2. Specific *monitoring* programs, related test *protocols*, or regulations might require selected test-specific items listed in Section 7.2 (e.g., details about the test material and/or explicit procedures and conditions during sample collection, handling, transport, and storage) to be included in the test-specific report, or might relegate certain test-specific information as *data to be held on file*.

Procedures and conditions common to a series of ongoing tests (e.g., routine toxicity tests for *monitoring* or *compliance* purposes) and consistent with specifications in this document may be referred to by citation or by attachment of a general report that outlines standard laboratory practice.

Details on the procedures, conditions, and findings of the test, which are not conveyed by the test-specific report or general report, must be kept on file by the laboratory for a minimum of five years so that the appropriate information can be provided if an audit of the test is required. Filed information might include:

- a record of the chain-of-continuity for field-collected or other samples tested for regulatory or *monitoring* purposes;
- a copy of the record of acquisition for the sample(s);
- chemical analytical data on the sample(s) not included in the test-specific report;
- bench sheets for the observations and measurements recorded during the test;
- bench sheets and warning chart(s) for the reference toxicity tests;
- detailed records of the source of the test organisms, their taxonomic confirmation, and all pertinent information regarding their sorting, preparation, and storage; and
- information on the calibration of equipment and instruments.

Original data sheets must be signed or initialled, and dated by the laboratory personnel conducting the tests.

7.1 Minimum Requirements for a Test-Specific Report

The following items must be included in each test-specific report.

7.1.1 Test Substance or Material

- brief description of sample type (e.g., reference or contaminated field-collected soil, negative control soil) or coding, as provided to the laboratory personnel;
- information on labelling or coding of each sample;
- information on sample horizons as they were collected (i.e., number, relative depth, and classification of each soil horizon), for test, reference, and negative control soils, if applicable; and
- date of sample collection; date and time sample(s) received at test facility.

7.1.2 Test Organisms

- species and source of test seeds;
- scientific name and lot number;
- duration and method of seed stratification and ethanol separation, if used; and
- any unusual appearance or treatment of the seeds, before their use in the test.

7.1.3 Test Facilities

- name and address of test laboratory; and
- name of person(s) performing the test (or each component of the test).

7.1.4 Test Method

- citation of biological test method used (i.e., as per this document);
- design and description if specialized procedure(s) (e.g., preparation of mixtures of *spiked soil*; preparation and use of solvent and, if so, solvent control) or modification(s) of the standard test method described herein;
- brief description of soil layering in test vessels (e.g., wet weights and/or depths of each soil), if applicable;
- brief description of frequency and type of all measurements and all observations made during test; and
- name and citation of program(s) and methods used for calculating statistical endpoints.

7.1.5 Test Conditions and Procedures

- design and description of any deviation(s) from, or exclusion of, any of the procedures and conditions specified in this document;
- number of discrete samples per treatment; number of replicate test vessels for each treatment; number and description of treatments in each test including the control(s); test concentrations (if applicable);
- volume and/or mass of soil in each test vessel;
- number of seeds per test vessel and treatment;
- dates when test was started and ended;
- measurements of light intensity adjacent to surface of soil in test vessels;
- for each soil sample—any measurements of soil particle size, moisture content, water-holding capacity, pH, and conductivity; and
- for each composite sample of subsamples taken at the same time from all replicates of each treatment—all measurements of temperature, pH, moisture content, and water-holding capacity.

7.1.6 Test Results

- percent emergence of plants in each test vessel at test end (Day 28, 35, or 42; depending on species of test organism);
- mean (± SD) percent emergence in control(s) at test end (Day 28, 35, or 42; depending on species of test organism), related to test validity criteria;
- mean (± SD) shoot length of individual plants surviving[44] in each treatment [including the control(s) and reference soil(s)] at test end; mean (± SD) root length of individual plants surviving in each treatment at test end; mean (± SD) shoot dry weight of individual plants surviving in each treatment at test end; mean (± SD) root dry weight of individual plants surviving in each treatment at test end;
- mean (± SD) shoot and root wet weight of individual plants surviving in each treatment [including the control(s) and reference soil(s)] at test end, if determined;
- any ICp (together with its 95% confidence limits) determined for the data on growth (i.e., shoot and root lengths and shoot and root wet and dry weights of individual plants surviving at test end); details regarding any transformation of data, and indication of quantitative statistical method used or procedures applied to the data;
- for a multi-concentration test with *chemical-spiked soil*, indication as to whether results are based on nominal or measured concentrations of chemical(s) or chemical product(s); all values for measured concentrations;
- results for any 14-, 21-, 28-, or 35-day (depending on test species) ICp (including its 95% confidence limits) performed with the *reference toxicant* in conjunction with the definitive soil toxicity test, using the same lot of test seed; *geometric mean* value (± 2 SD) for the same reference toxicant and test species, as derived at the test facility in previous 14-, 21-, 28-, or 35-day ICp tests using the procedures and conditions for reference toxicity tests described herein; and
- anything unusual about the test, any problems encountered, any remedial measures taken.

7.2 Additional Reporting Requirements

The following items must be either included in the test-specific report or the general report, or held on file for a minimum of five years.

7.2.1 Test Substance or Material

- identification of person(s) who collected and/or provided the sample;
- records of sample chain-of-continuity and log-entry sheets; and
- conditions (e.g., temperature, in darkness, in sealed container) of sample upon receipt and during storage.

7.2.2 Test Organisms

- name and address of seed supplier;
- year of collection (if applicable), packet size, lot number, percent germination rating, date of germination rating, date of purchase, name of supplier, and date seed package was opened;
- description of procedures used to sort and stratify seeds;
- description of storage conditions and procedures, including temperature and duration of seed-lot storage;
- results of any assessment of seed germination prior to testing.

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 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 analyses 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

- procedure for adding seed to test vessels;
- relative humidity of test facility (optional);
- appearance of each sample (or mixture thereof) in test vessels; changes in appearance noted during test;
- records of hydration of *test soils* in each test vessel throughout duration of test and qualitative description of drying of soils during the test (e.g., rate, appearance);
- any other physicochemical measurements (e.g., analyses of aliquots from the same *batch* to determine conductivity, homogeneity, contaminant concentration, total volatile solids, biochemical oxygen demand, chemical oxygen demand, total inorganic carbon, *cation exchange capacity*, *redox potential*, total nitrogen) 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 pore water;
- any other observations or analyses made on the test material (including samples of *negative control soil* or *reference soil*); e.g., qualitative and/or quantitative data regarding indigenous macrofauna or detritus, or results of geochemical analyses; and

- any chemical analyses of the concentration of chemical in *stock solution*(s) of *reference toxicant* and, if measured, in test concentrations.

7.2.7 Test Results

- results for any range-finding test(s) conducted;
- number of emerged seedlings and observations on seedling condition in each test vessel, as noted during each observation period over the test duration;
- number of surviving plants in each test vessel at test end (Day 28, 35, or 42); mean individual shoot/root length; and replicate dry weight and mean individual dry weights of shoots and roots (and the same for shoot and root wet weight, if performed) of plants surviving in each test vessel at test end; for regression analyses, hold on file information indicating sample size (e.g., number of replicates per treatment), parameter estimates with variance, any ANOVA table(s) generated, plots of fitted and observed values of any models used, and the output provided by the statistical program (e.g., SYSTAT, CETIS);
- warning charts (for ICps causing reduced root lengths) showing the most recent and historic results for toxicity tests with the reference toxicant and the selected species of test organism used in these tests;
- graphical presentation of data; and
- original bench sheets and other data sheets, signed and dated by the laboratory personnel performing the test and related analyses.

References

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

———, "Ecological Framework of Canada – Ecozone and Ecoregion Descriptions," URL: http://ecozones.ca/english/zone/index.html (no date).

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

AE (Alberta Environment), "Alberta Tier 1 Soil and Groundwater Remediation Guidelines," June 2007, ISBN: 978-0-7785-6306-8 [online version] (2007a).

———, "Alberta Tier 2 Soil and Groundwater Remediation Guidelines," June 2007, ISBN: 978-0-7785-6308-2 [online version] (2007b).

———, "Tier 2 Ecocontact Guideline Derivation Protocol," July 13, 2007, draft, ISBN: 978-0-7785-6753-0 [online version] (2007c).

Ancient Forest Exploration and Research, "Ontario's Oldest Trees," URL: www.ancientforest.org/oldtrees.htm. Last accessed January 6, 2013 (no date).

Angell, R.A., S. Kullman, E. Shrive, G.L. Stephenson, and M. Tindal, "Direct Soil Contact Values for Ecological Receptors Exposed to Weathered Petroleum Hydrocarbon (PHC) Fraction 2," *Environ. Toxicol. Chem.,* 31 (11):2536-2544 (2012).

Aquaterra Environmental, "Development of Plant Toxicity Tests for Assessment of Contaminated Soils," Report prepared by Aquaterra Environmental, Ltd. (Orton, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (1998).

Aquaterra Environmental and ESG (ESG International Inc.), "Assessment of the Biological Test Methods for Terrestrial Plants and Soil Invertebrates: Metals," August 2000 Final Report, prepared by Aquaterra Environmental, Ltd. (Orton, ON) and ESG International Inc. (Guelph, ON) for the Method Development and Applications Section, Environment Canada, Ottawa, ON (2000).

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

———, "Standard Guide for Use of Lighting in Laboratory Testing," E-1733-95, American Society for Testing and Materials, West Conshohocken, PA (2008).

———, "Standard Guide for Conducting Terrestrial Plant Toxicity Tests," E1963-09, American Society for Testing and Materials, West Conshohocken, PA (2009).

Baskin, C.C., and J.M. Baskin, *Propagation Protocol for Production of Container* Picea mariana *(Miller)*, *Britton plants*, University of Kentucky, Lexington, Kentucky, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2002).

Becker-van Slooten, K., S. Campiche, and J. Tarradellas, "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 the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004).

Best, K.F., J. Looman, and J.B. Campbell, "Prairie Grasses Identified and Described by Vegetative Characters," Can. Dept. Ag. Pub. 1413 (1971).

Bevington, J.M and M.C. Hoyle, "Phytochrome Action during Prechilling Induced Germination of *Betula papyrifera* Marsh," *Plant Physiol,* 67:705-710 (1981).

Bjorkbom, J.C., "Production and Germination of Paper Birch Seed and its Dispersal into a Forest Opening," Res. Pap. NE-209, Upper Darby, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, 14 pp. (1971).

Bjorkroth, G., "Separation of Birch Seed in Liquids According to Density," Proc. Int. Symp. Seed Processing. Vol 1, IUFRO Work Part. S2.01.06, Bergen, Norway, 9 pp. (1973).

Bonnell Environmental Consulting, "Assessment of Soil Toxicity Test Species for Canadian Representativeness," Technical Report TS-28, Prepared for the Method Development and Applications Section, Environment Canada, by Bonnell Environmental Consulting, Ottawa, ON (1994).

Boutin, C., and C.A. Rogers, "Pattern of Sensitivity of Plant Species to Various Herbicides – An Analysis with Two Databases," *Ecotoxicol.,* 9:255-271 (2000).

Boutin, C., H.B. Lee, E.T. Peart, P.S. Betchelor, and R.J. Maguire, "Effects of the Sulfonylurea Herbicide Metsulfuron Methyl on Growth and Reproduction of Five Wetland and Terrestrial Plant Species," *Environ. Toxicol. Chem.,* 19(10):2532-2541 (2000).

Boutin, C., N. Elmegaard, and C. Kjær, "Toxicity Testing of Fifteen Non-Crop Plant Species with Six Herbicides in a Greenhouse Experiment: Implications for Risk Assessment," *Ecotoxicol.,* 13 (4):349−369 (2004).

Boutin C, A.L. White, and D. Carpenter, "Measuring Variability in Phytotoxicity Testing using Crop and Wild Plant Species," *Environ. Toxicol. Chem.* 29(2):327-337 (2010).

Butler, J., and C. Frieswyk, *Propagation Protocol for Production of* Calamagrostis canadensis *Seeds*, USDI NPS – Rocky Mountain National Park, Estes Park, Colorado, in: "Native Plant Network." URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2001).

CCME (Canadian Council of Ministers of the Environment), "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, Ottawa, ON, 108 pp. (1994).

———, "A Framework for Ecological Risk Assessment: General Guide," CCME Subcommittee on Environmental Quality Criteria for Contaminated Sites, The National Contaminated Sites Program, Winnipeg, MB, 32 pp. (1996).

———, "A Framework for Ecological Risk Assessment: Technical Appendices," CCME Subcommittee on Environmental Quality Criteria for Contaminated Sites, The National Contaminated Sites Program, Winnipeg, MB, 67 pp. (1997).

———, "Canada-Wide Standard for Petroleum Hydrocarbons (PHC) in Soil: Scientific Rationale," Supporting Technical Document PN 1399, Winnipeg, Manitoba, (2008).

Callahan, C.A., C.A. Menzie, D.E. Burmaster, D.C. Wilborn, and T. Ernst, "On-Site Methods for Assessing Chemical Impact on the Soil Environment Using Earthworms: A Case Study at the Baird and McGuire Superfund Site, Holbrook, Massachusetts," *Environ. Toxicol. Chem.,* 10:817-826 (1991).

Carey, J.H., *Pinus banksiana*, in: "Fire Effects Information System" [online], URL: www.fs.fed.us/database/feis/. Last accessed January 7, 2013 (1993).

Carter, M.R. (ed.), "Soil Sampling and Methods of Analysis," Lewis Publishers, CRC Press Inc., Boca Raton, FL (1993).

Cermak, J.H., G.L. Stephenson, D. Birkholz, Z. Wang, and D.G. Dixon, "Toxicity of Petroleum Hydrocarbon Distillates to Soil Organisms," *Environ. Toxicol. Chem.,* 29(12):2685-2694 (2010).

Chaineau, C.H., J.L. Morel, and J. Oudot, "Phytotoxicity and Plant Uptake of Fuel Oil Hydrocarbons," *J. Environ. Qual.,* 26:1478-1483 (1997).

Coladonato, M., *Solidago Canadensis*, in: "Fire Effects Information System" [online], U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer), URL: www.fs.fed.us/database/feis/. Last accessed January 7, 2013 (1993).

Courchesne, F., S. Savoie, and A. Dufresne, "Effects of Air-Drying on the Measurement of Soil pH in Acidic Forest Soils of Quebec, Canada," *Soil Science*, 160(1):56-68 (1995).

Crépin, J., and R.L. Johnson, *Soil Sampling for Environmental Assessment*, pp. 5-18, in: "Soil Sampling and Methods of Analysis," M.R. Carter (ed.), Lewis Publishers, CRC Press Inc., Boca Raton, FL (1993).

Croser, C., S. Renault, J. Franklin, J. Zwiazek, "The Effect of Salinity on the Emergence and Seedling Growth of *Picea mariana*, *Picea glauca*, and *Pinus banksians*," *Environmental Pollution*, 115:9-16 (2001).

Dalton, R.C., and C. Boutin, "Comparison of the Effects of Glyphosate and Atrazine Herbicides on Nontarget Plants Grown Singly and in Microcosms," *Environ. Toxicol. Chem.,* 29(10):2304-2315 (2010).

Day, R.A., R.P. Walter, J.J. Kozar, S.J. Bricker, J.G. Bowers, "Propagation Protocol for Bareroot Bigtooth and Quaking Aspen Using Seeds," *Native Plants Journal*, 4(2):125-128 (2003)

Deitzer, G., "Spectral Comparisons of Sunlight and Different Lamps," *c74*, pp. 197-199, in: *Proceedings of International Lighting in Controlled Environments Workshop*, T. W. Tibbits (ed.), 1 Mar. 1994, Madison, WI (1994).

Denton, D., J. Diamond, and L. Zheng, "Test of Significant Toxicity: A Statistical Application for Assessing Whether an Effluent or Site Water is Truly Toxic," *Environ. Toxicol. Chem.,* 30:1117-1126 (2011).

Dunnett, C.W., "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *J. Amer. Stat. Assoc.,* 50:1096-1121 (1955).

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

———, "Biological Test Method: Test for Survival and Growth in Sediment Using the Freshwater Amphipod *Hyalella azteca,*" Environmental Protection Service, Ottawa, Ontario, Report EPS 1/RM/33, 123 pp. (1997a).

———, "Biological Test Method: Test for Survival and Growth in Sediment Using Larvae of Freshwater Midges *(Chironomus tentans* or *Chironomus riparius)*," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/32, 131 pp. (1997b).

———, "Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/34, 203 pp. (1999a).

———, "Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/37, 98 pp. (1999b).

———, "Standard Operating Procedure for Conducting Early Seedling Growth Bioassays," Standard Operating Procedure EC50V1.1, 21 pp., April 2000, Environmental Toxicology Section, Pacific Environmental Science Centre, North Vancouver, BC (2000).

———, "Biological Test Method: Test for Toxicity of Contaminated Soil to Earthworms *(Eisenia andrei, Eisenia fetida,* or *Lumbricus terrestris)*," including 2007 amendments, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/43, 156 pp. (2004a).

———, "Proceedings to the Workshop on Toxicity Test methodologies for Assessing the Impacts of Contaminant Mixtures in Soil Using Terrestrial Species of Ecological Relevance to Canadian Soil Systems," Workshop held February 19-21, 2003 at the Pacific Environmental Science Centre, North Vancouver, BC, report published by the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004b).

———, "Biological Test Method: Test for Measuring Emergence and Growth of Terrestrial Plants exposed to Contaminants in Soil," including 2007 amendments, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/45, 131 pp. (2005a).

———, "Guidance Document on Statistical Methods for Environmental Toxicity Tests," including 2007 amendments, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/46, 241 pp. (2005b).

———, "Biological Test Method: Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/47, 146 pp. (2007a).

———, "Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*," 2nd edition, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/21, 100 pp. (2007b).

———, "Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*," 2nd edition, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/37, 141 pp. (2007c).

———, "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," Biological Assessment and Standardization Section, Ottawa, ON, 167 pp. (2010).

———, "Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows," 2nd edition, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/22, 100 pp. (2011a).

———, "Biological Test Method: Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars)," 2nd edition, Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/27, 137 pp. (2011b).

———, "Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing," Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/53, 222 pp. (2012).

———, "Inter-laboratory Validation of Environment Canada's New Test Method for Measuring Soil Toxicity Using Plants Native to the Canadian Boreal Region," Biological Assessment and Standardization Section, Ottawa, ON (2013a).

———, "Development and Standardization of Toxicity Test Methodologies for Assessment of Impacts from Contamination using Species and Microbial Communities Representative of Canadian Boreal and Taiga Regions. Technical Report: 2010–2013," Biological Assessment and Standardization Section, Ottawa, ON (2013b).

EC and SRC (Environment Canada and Saskatchewan Research Council), "Validation of Toxicology Test Methods for Assessing Petroleum Hydrocarbon and Brine Spills in Boreal Forest Soils," Prepared for Environmental Research Advisory Council, Canadian Association of Petroleum Producers; prepared by: Biological Methods Division, Environment Canada, Ottawa, ON and Environment and Forestry Division Saskatchewan Research Council, October 2007.

Earle, C.J., *Picea glauca*, in: "The Gymnosperm database," URL: www.conifers.org/pi/pic/glauca.php. Last accessed January 7, 2013 (2009a).

———, *Pinus banksiana*, in: "The Gynmnosperm database," URL: www.conifers.org/pi/pinus_banksiana.php. Last accessed January 7, 2013 (2009b).

EcoDynamics Consulting Ltd., "Sampling in the Swan Hills and Peace River Areas of Alberta and the Torch River Area of Saskatchewan," Prepared for Environment Canada, Ottawa, Ontario, Canada (2007).

———, "Sampling of a Reference Soil in the Bathurst Region of New Brunswick – Ecological Site Characterization," Prepared for Environment Canada, Ottawa, Ontario, Canada (2008a).

———, "Site SK09 Robins Lake, Saskatchewan – Soil Sampling and Ecological Description," Prepared for Environment Canada, Ottawa, Ontario, Canada (2008b).

———, "Ecological Description of the Ontario Reference Site," Prepared for Environment Canada, Ottawa, Ontario, Canada (2011).

———, "Biological Test Method: Test for Survival and Growth in Sediment Using Spionid Polychaete Worms (*Polydora cornuta*)", Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/41, 108 pp. (2001).

———, "Guidance Document on Statistical Methods for Environmental Toxicity Tests", Method Development and Applications Section, Environment Canada, Ottawa, ON, Report EPS 1/RM/46 (2004a).

———, "Proceedings to the Workshop on Toxicity Test Methodologies for Assessing the Impacts of Contaminant Mixtures in Soil Using Terrestrial Species of Ecological Relevance to Canadian Soil Systems," Workshop held February 19-21, 2003 at the Pacific Environmental Science Centre, North Vancouver, BC, report published by the Method Development and Applications Section, Environment Canada, Ottawa, ON (2004b).

———, "Biological Test Method: Tests for Toxicity of Contaminated Soil to Earthworms (*Eisenia andrei*, *Eisenia fetida*, or *Lumbricus terrestris*)", Method Development and Applications Section, Environment Canada, Ottawa, ON, Report EPS 1/RM/43, 156 pp. (2004c).

———, "Inter-laboratory Validation of Environment Canada's New Test Method for Measuring Soil Toxicity Using Plants," June, 2005, report prepared by J. Princz for the Biological Methods Division, Environment Canada, Ottawa, ON (2005a).

———, "Development of Performance Data for the Determination of Validity Criteria for Environment Canada's Biological Test Method for Assessing Soil Toxicity Using Plants," Biological Methods Division, Ottawa, ON, 50 pp. (2005b).

———, "Biological Test Method: Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil," Biological Methods Division, Environment Canada, Ottawa, ON, Report EPS 1/RM/47, 176 pp. (2007).

Eisenhart, C., M.W. Hastay, and W.A. Wallace, "Techniques of Statistical Analysis," Chapter 15, McGraw-Hill Book Co., New York, NY (1947).

Emery, S.M., and K.L. Gross, "Dominant Species Identity, not Community Evenness, Regulates Invasion in Experimental Grassland Plant Communities," *Ecology,* 88:954-964 (2007).

ESG (ESG International Inc.), "Impacts of Metal-Contaminated Forest Soils From the Canadian Shield on Terrestrial Organisms," report prepared by ESG International Inc. (Guelph, ON) for the Metals in the Environment Research Network Canadian Network of Toxicology Centers, Guelph, ON and the Method Development and Applications Section, Environment Canada, Ottawa, ON (2002).

Feisthauer, N.A., G.L. Stephenson, J.I. Princz, and R.P. Scroggins, "Effects of Metal-Contaminated Forest Soils from the Canadian Shield to Terrestrial Organisms," *Environ. Toxicol. Chem., 25*(3):823-835 (2006).

Fletcher, J.S., M.J. Muhitch, D.R. Vann, J.C. McFarlane, and F.E. Benenati, "PHYTOTOX Database Evaluation of Surrogate Plant Species Recommended by the U.S. Environmental Protection Agency and the Organization for Economic Cooperation and Development," *Environ. Toxicol. Chem.,* 4:523-532 (1985).

Fletcher, J.S., F.L. Johnson, and J.C. McFarlane, "Database Assessment of Phytotoxicity Data Published on Terrestrial Vascular Plants," *Environ. Toxicol. Chem.,* 7:615-622 (1988).

Fletcher, J.S., T.G. Pfleeger, and H.C. Ratsch, "Chlorsulfuron Influence on Garden Pea Reproduction," *Physiol. Plant,* 94:261-267 (1995).

Fletcher, J.S., T.G. Pfleeger, H.C. Ratsch, and R. Hayes, "Potential Impact of Low levels of Chlorsulfuron and Other Herbicides on Growth and Yield of Non-Target Plants," *Environ. Toxicol. Chem.,* 15(7)*:*1189-1196 (1996).

Flora of North America, Vol. 3, "Betulaceae," URL: www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=10101. Last accessed January 7, 2013 (2009a).

———, Vol. 2. "Pinaceae," URL: www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=200005326. Last accessed January 7, 2013 (2009b).

Fowells, H.A., *White Spruce (*Picea glauca *(Moench) Voss)*, Agriculture handbook 271, "Silvics of Forest Trees of the United States, USDA-Forest Service," Washington, DC (1965).

Franklin J.A., S. Renault, C. Croser, J.J. Zwiazek, and M. MacKinnon, "Jack Pine Growth and Elemental Composition are Affected by Saline Tailings Water," *Journal of Environmental Quality* 31:648-653 (2002).

Frankton, C., and G.A. Mulligan, "*Weeds of Canada*," Canada Department of Agriculture, Ottawa (1970).

Godbold, D.L., and A. Hüttermann, "Effect of Zinc, Cadmium and Mercury on Root Elongation of *Picea abies* (Karst.) Seedlings and the Significance of These Metals to Forest Die-Back," *Environ. Pollut. Ser. A* (38):375-381 (1985).

Harrington, J.T., *Propagation protocol for production of container* Populus tremuloides *(Michx.) plants (164-mL containers)*, New Mexico State University-Mora Research Center, Mora, New Mexico, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2010).

Hausenbuiller, R.L., "Soil Science – Principles & Practices," 3rd edition, W.C. Brown Publisher, Dubuque, IA (1985).

Hendershot, W.H., H. Lalonde, and M. Duquette, *Soil Reaction and Exchangeable Acidity*," pp. 141-145, in: "Soil Sampling and Methods of Analysis," M.R. Carter (ed.), Canadian Society of Soil Science, Lewis Publishers, Boca Raton, FL (1993).

Hettinger, L.R., "Vegetation Response to Brine Spill Reclamation Measures: Boreal Forest, Alberta, Canada [Crude oil production]," Information Systems Division, National Agricultural Library (United States of America) NAL/USDA, Beltsville, Maryland (1982).

Hoenig, J.M., and D.M. Heisey, "The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis," *Amer. Statistician*, 55:19-24 (2001).

Hosie, R.C., "Native Trees of Canada," Don Mills, ON: Fitzhenry & Whiteside, 380 pp. (1979).

Howard, J.L., *Populus tremuloides*, in: "Fire Effects Information System" [online], U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer), URL: www.fs.fed.us/database/feis/. Accessed November 25, 2010 (1996).

Howat D.R., "Acceptable Salinity, Sodicity and pH Values for Boreal Forest Reclamation," Alberta Environment, Environmental Sciences Division Rep. ESD/LM/00-2 (2000).

ISO (International Organization for Standardization), "Soil – Determination of the Effect of Chemical Substances on the Reproduction of Earthworms," Proposal from the Netherlands, 9 pp., April 1991, Geneva, Switzerland (1991).

———, "Soil Quality – Effects of Pollutants on Earthworms *(Eisenia fetida)*. Part 1: Determination of Acute Toxicity Using Artificial Soil Substrate," ISO 11268-1, Geneva, Switzerland (1993).

———, "Soil Quality – Effects of Pollutants on Earthworms *(Eisenia fetida)*. Part 2: Determination of Effects on Reproduction," ISO 11268-2, Geneva, Switzerland (1998).

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

———, "Soil quality – Vocabulary," ISO 11074, Geneva, Switzerland (2005).

———, "Soil Quality – Determination of the Effects of Pollutants on Soil Flora- Part 1: Method for the Measurement of Inhibition of Root Growth," ISO 11269-1:2012, Geneva, Switzerland (2012a).

———, "Soil Quality – Determination of the Effects of Pollutants on Soil Flora-Part 2: Effects of Contaminated Soil on the Emergence and Early Growth of Higher Plants," ISO 11269-2:2012, Geneva, Switzerland (2012b).

Kabzems, A., "Growth and Yield of Well Stocked White Spruce in the Mixedwood Section," Saskatchewan Tech. Bull. No. 5., Dept. Nat. Res., SK. Forestry Branch, 75 pp. (1971).

Kabzems, A., and C.L. Kirby, "The Growth and Yield of Jack Pine in Saskatchewan," Saskatchewan Dept of Nat. Res. Tech. Bull. 2, 66 pp. (1956).

Kalsch, W., and J. Römbke, *Zur chronischen Wirkung von TNT auf die Stoppelrübe* Brassica rapa *im Labortest*," pp. 100-105, in: "Ökotoxikologie – Ökosystemare Ansätze und Methoden," J. Oehlmann and B. Barkert (eds.), Verlag Urban & Fischer, Stuttgart, Jena (1999).

Kaputska, L.A., *Selection of Phytotoxicity Tests for Use in Ecological Risk Assessments*," pp. 515-548, in: "Plants for Environmental Studies," W. Wang, J.W. Gorsuch, and J.S. Hughes (eds.), CRC Press, New York, NY (1997).

Kaputska, L.A., J. Lipton, H. Galbraith, D. Cacela, and K. LeJeune, "Metal and Arsenic Impacts to Soils, Vegetation Communities and Wildlife Habitat in Southwest Montana Uplands Contaminated by Smelter Emissions: II. Laboratory Phytotoxicity Studies," *Environ. Toxicol. Chem.,* 14 (11):1905-1912 (1995).

Kenaga, E.E., "Comparative Toxicity of 131,596 Chemicals to Plant Species," *Ecotoxicol. Environ. Safety,* 5:469-475 (1981).

Khasa P.D., B. Hambling, G. Kernaghan, M. Fung, and E. Ngimbi, "Genetic Variability in Salt Tolerance of Selected Boreal Woody Seedlings," *Forest Ecol Manage* 165:257-269 (2002).

Khasa, D.P., M. Fung, B. Logan, "Early Growth Response of Container-grown Selected Woody Boreal Seedlings in Amended Composite Tailings and Tailings Sand," *Bioresource Technology,* 96:857-864 (2004).

Kjær, C., and N. Elmegaard, "Effects of Copper Sulfate on Black Bindweed," *Ecotoxicol. Environ. Safety*, 33:110-119 (1996).

Kjær, C., M.B. Pedersen, and N. Elmegaard, "Effects of Soil Copper on Black Bindweed *(Fallopia convolvulus)* in the Laboratory and in the Field," *Arch. Environ. Contam. Toxicol.,* 35:14-19 (1998).

Lakehead University, Pinus banksiana*: Jack pine*, in: "Borealforest.org," URL: http://bolt.lakeheadu.ca/~borfor/trees/tree8.htm. Last accessed January 7, 2013 (2009).

Lawrence, D., C.T. Lane, M.G. Willoughby, C. Hincz, D. Moisey, and C. Stone, "Range Plant Community Types and Carrying Capacity for the Lower Foothills Subregion of Alberta: Fourth Approximation," Alberta Sustainable Resource Development, Public Lands and Forests Pub. No. T/083, Edmonton, ISBN: [online edition] 0-7785-4098-7 (2005).

Lenth, R.V., "Statistical Power Calculations," *J. Anim. Sci. 85(E. Suppl.)*: E24-E29 (2007).

Lock, K., and C.R. Janssen, "Influence of Aging on Copper Bioavailability in Soils," *Environ. Toxicol. Chem.,* 22 (5):1162-1166 (2003).

Luna, T., J. Evans, J. Hosokawa, and D. Wick, *Propagation protocol for production of container* Populus tremuloides *Michx. plants (3 L Containers)*, USDI NPS – Glacier National Park, West Glacier, Montana, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2001).

Marquis, D.A., "Silvical Requirements for Natural Birch Regeneration," in: *The Birch Symposium: Proceedings; 1969 August 19-21; Durham, NH*, Res. Pap. NE-146, Upper Darby, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, pp. 40-49 (1969).

Mearns, A.J., R.C. Swartz, J.M. Cummins, P.A. Dinnel, P. Plesha, and P.M. Chapman, "Interlaboratory Comparison of a Sediment Toxicity Test Using the Marine Amphipod, Rhepoxynius abronius," *Marine Environ. Res.*, 19:13-37 (1986).

Meier, J.R., L.W. Chang, S. Jacobs, J. Torsella, M.C. Meckes, and M.K. Smith, "Use of Plant and Earthworm Bioassays to Evaluate Remediation of Soil from a Site Contaminated with Polychlorinated Biphenyls," *Environ. Toxicol. Chem.*, 16(5):928-938 (1997).

Menzie, C.A., D.E. Burmaster, J.S. Freshman, and C.A. Callahan, "Assessment of Methods for Estimating Ecological Risk in the Terrestrial Component: A Case Study at the Baird & McGuire Superfund Site in Holbrook, Massachusetts," *Environ. Toxicol. Chem., 11:*245-260 (1992).

Menzie, C.A., B. Southworth, G. Stephenson, and N. Feisthauer, "The Importance of Understanding the Chemical Form of a Metal in the Environment: The Case of Barium Sulfate (Barite)," *Human and Ecological Risk Assessment: An International Journal,*14(5):974-991 (2008).

Miller, W.E., S.A. Peterson, J.C. Greene, and C.A. Callahan, "Comparative Toxicology of Laboratory Organisms for Assessing Hazardous Waste Sites," *J. Environ. Qual.,* 14:569-574 (1985).

Moench, R., *Propagation protocol for production of container* Populus tremuloides *(Michx.) plants (1+0 container)*, Colorado State Forest Service Nursery, Fort Collins, Colorado, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Accessed December 3, 2010, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2001).

Moss, E.H., and J.G. Packer, "Flora of Alberta," 2nd edition, University of Toronto Press, Toronto, ON, 687 pp. (1994).

NERI (National Environmental Research Institute), "Manual of SECOFASE – Development, Improvement and Standardization of Test Systems for Assessing Sublethal Effects on Chemicals on Fauna in the Soil Ecosystem," H. Lokke and C.A.M. van Gestel (eds.), Report from a Workshop held in Silkeborg January 18-19, 1993, Silkeborg, Denmark, 41 pp. (1993).

Newman, M.C., "Quantitative Methods in Aquatic Ecotoxicology," Lewis Pub., Boca Raton, FL, 426 pp. (1995).

———, "What Exactly are you Inferring? A Closer Look at Hypothesis Testing," *Environ. Toxicol. Chem.,* 27:1013-1126 (2008).

Nienstaedt, H., and J.C. Zasada, *White Spruce (*Picea glauca *(Moench) Voss)*, in: Burns, R.M., and B.H. Honkala, (technical coordinators), "Silvics of North America: 1. Conifers; 2. Hardwoods," Agriculture Handbook 654, U.S. Department of Agriculture, Forest Service, Washington, DC., vol. 2 877 pp. (1990).

Noller, G.L., *Propagation protocol for production of* Calamagrostis canadensis *(Michx.) Beauv. Plants,* Upper Colorado Environmental Plant Center, Meeker, Colorado, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2001).

Norberg-King, T.J., "A Linear Interpolation Method for Sublethal Toxicity: The Inhibition Concentration (ICp) Approach (Version 2.0)," USEPA, Environ. Res. Lab.-Duluth, Duluth, Minnesota, Tech. Report 03-93 of National Effluent Toxicity Assessment Center (1993).

OECD (Organization for Economic Cooperation and Development), "OECD Guideline for Testing of Chemicals: Earthworm, Acute Toxicity Tests," No. 207, Paris, France, 9 pp. (1984).

———, "OECD Guideline for the Testing of Chemicals: Earthworm Reproduction Test *(Eisenia fetida/ Eisenia andrei)*," No. 222, Paris, France, 18 pp. (2004).

———, "OECD Guideline for the Testing of Chemicals: Terrestrial Plant Test: Seedling Emergence and Seedling Growth Test," No. 208, Paris, France, 21 pp. (2006a).

———, "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 (2006b).

OMEE (Ontario Ministry of Environment and Energy), "Guidance on Sampling and Analytical Methods for Use at Contaminated Sites in Ontario – Section 8.3.7 Inorganics," Version 1.1, ISBN-0-7778-4056-1, December 1996, Queen's Printer, Toronto, ON (1996).

Perala, D.A., Populus tremuloides *Michx.: Quaking Aspen*, in: Burns, R.M., and B.H. Honkala, (technical coordinators), "Silvics of North America: 1. Conifers; 2. Hardwoods," Agriculture Handbook 654, U.S. Department of Agriculture, Forest Service, Washington, DC., vol. 2, 877 pp. (1990).

Pfleeger, T., C. McFarlane, R. Sherman, and G. Volk, *A Short-Term Bioassay for Whole Plant Toxicity*, pp. 355-364, in: "Plants for Toxicity Assessment: 2nd Volume," ASTM STP 1115, J.W. Gorsuch, W.R. Lower, W. Wang, and M.S. Lewis (eds.), American Society for Testing and Materials, Philadelphia, PA (1991).

Powell, G.W., and E.W. Bork, "Competition and Facilitation in Mixtures of Aspen Seedlings, alfalfa, and marsh reedgrass," *Can. J. For. Res.* 34:1858-1869 (2004).

Princz, J.I., M. Moody, C. Fraser, L. Van der Vliet, H. Lemieux, R. Scroggins, and S.D. Siciliano, "Evaluation of a New Battery of Toxicity Tests for Boreal Forest Soils: Assessment of the Impact of Hydrocarbons and Salts," *Environ. Toxicol. Chem.,* 31(4):766-777 (2012).

Qualtiere, E.J., "Variation in Germination Response to Temperature Among Collections of Three Conifers from the Mixed Wood Forest," MSc Thesis, Univ. Sask., Saskatoon (2008).

Rader, B.R., D.W.R. Nimmo, and P.L. Chapman, "Phytotoxicity of Floodplain Soils Contaminated with Trace Metals Along the Clark Fork River, Grant-Kohrs Ranch National Historic Site, Deer Lodge, Montana, USA," *Environ. Toxicol. Chem.,* 16(7):1422-1432 (1997).

Redente, E.F., H. Zadeh, and M.W. Paschke, "Phytotoxicity of Smelter-Impacted Soils in Southwestern Montana, USA," *Environ. Toxicol. Chem.,* 21(2):269-274 (2002).

Redfield, E.B., "Tolerance Mechanisms of Black Spruce *(Picea mariana)* Seedlings Exposed to Saline Oil Sands Tailings," MSc. Thesis, in Land Reclamation and Remediation, Department of Renewable Resources, University of Alberta, Edmonton, Alberta, 89 pp. (2001).

Renault S, J.J. Zwiazek, M. Fung, and S. Tuttle, "Germination, Growth and Gas Exchange of Selected Boreal Forest Seedlings in Soil Containing Oil Sands Tailings," *Environmental Pollution* 107:357-365 (2000).

Rocchini, R.J., M.J.R. Clark, A.J. Jordan, S. Horvath, D.J. McLeay, J.A. Servizi, A. Sholund, H.J. Singleton, R.G. Watts, and R.H. Young, "Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish," B.C. Ministry of Environment, Victoria, BC (1982).

Rook, J.S., "*Calamagrostis canadensis*: Bluejoint reedgrass," URL: www.rook.org/earl/bwca/nature/grass/calamagrostis.html. Last accessed January 7, 2013 (1998).

———, "*Picea mariana*: Black Spruce," URL: www.rook.org/earl/bwca/nature/trees/piceamar.html. Last accessed January 7, 2013 (2002).

Rudolph, T.D., and P.R. Laidly, Pinus banksiana *Lamb.: Jack Pine*, in: Burns, R.M., and B.H. Honkala, tech. coords., "Silvics of North America: 1. Conifers; 2. Hardwoods," Agriculture Handbook 654, U.S. Department of Agriculture, Forest Service, Washington, DC, Vol.2, 877 pp. (1990).

Sager, J.C., and C. McFarlane, *Radiation*, pp. 1-30, in: "Plant Growth Chamber Handbook," R.W. Langhans and T.W. Tibbits (eds.), North Central Regional Research Publication No. 340, Iowa Agriculture and Home Economics Experiment Station Special Report No. 99, Iowa State University of Science and Technology, Ames, IA (1997).

SPAC (Soil and Plant Analysis Council Inc.), "Soil Analysis-Handbook of Reference Methods," Soil and Plant Analysis Council Inc., Georgia University Station, Athens, GA, 202 pp. (1992).

SRC (Saskatchewan Research Council), "Boreal Forest Plant Species for Testing of Contaminated Soil," SRC Publication No. 11627-2E03, 49 pp. (2003).

———, "Development of Terrestrial Plant Test Methods using Boreal Forest Species," SRC Publication No. 11627-1C04, 62 pp. (2004).

———, "Further Development of Boreal Forest Plant Toxicity Test," Saskatchewan Research Council (SRC) Publication No. 11627-1E06, 56 pp. (2006).

———, "Investigations to Refine Toxicity Test Methods using Boreal Forest Plants: Soil pH Tolerance, Seed Source Variability, Test Design Differences and Testing of Soils with Brine and PHC Contamination," SRC Publication No.11627-3E07, 54 pp. (2007).

———, "Investigations to Refine Toxicity test using Boreal Forest Plants: Testing of Brine and PHC Contaminated Soil, an Ontario Podzol Soil and Two Wetland Plants," SRC Publication No. 11627-2E08, 48 pp. (2008).

———, "Research in Support of a Standardized Toxicity Test Method using Boreal Forest Plants and Testing of a podzol from New Brunswick," SRC Publication No. 11627-1C09, 92 pp. (2009).

———, "Investigation of Paper Birch Germination," SRC Publication No. 11627-5E12, 14 pp. (2012).

Saterbak, A., R.J. Toy, D.C.L. Wong, B.J. McMain, M.P. Williams, P.B. Dorn, L.P. Brzuzy, E.Y. Chai, and J.P. Salanitro, "Ecotoxicological and Analytical Assessment of Hydrocarbon-contaminated Soils and Application to Ecological Risk Assessment," *Environ. Toxicol. Chem.,* 18 (7):1591-1607 (1999).

Saterbak, A., R.J. Toy, B.J. McMain, M.P. Williams, and P.B. Dorn, "Ecotoxicological and Analytical Assessment of Effects of Bioremediation on Hydrocarbon-Containing Soils," *Environ. Toxicol. Chem.,* 19:2643-2652 (2000).

Scoggan, H.J., "*Flora of Canada*," National Museum of Natural Sciences, Publications in Botany, No. 7, Ottawa, ON (1978).

Sheppard, S.C., "Development of Bioassay Protocols for Toxicants in Soil," Report PIBS 2838, ISBN 0-7778-1440-X, Environmental Research, Research and Technology Branch, Ontario Environment, Queen's Printer, Toronto, ON, February 1994 (1994).

Sheppard, S.C., and W.G. Evenden, *An Approach to Defining a Control or Diluent Soil for Ecotoxicity Assays*," pp. 215-226, in: "Environmental Toxicology and Risk Assessment, 7th Volume," ASTM STP 1333, E.E. Little, A.J. DeLonay, and B.M. Greenberg (eds.), American Society for Testing and Materials, Philadelphia, PA (1998).

Sheppard, S.C., and G.L Stephenson, "Ecotoxicity of Aged Uranium in Soil Using Plant, Earthworm and Microarthropod Toxicity Tests," *Bull Environ Contam Toxicol* 88:43-47 (2012).

Siciliano, S.D., J.J. Germida, and J.V. Headley, "Evaluation of Prairie Grass Species as Bioindicators of Halogenated Aromatics in Soil," *Environ. Toxicol. Chem.,* 16(3):521-527 (1997).

Simpson, D., and B. Daigle, "Improving the Quality of White Birch *(Betula papyrifera)* Seed," Can. Tree Imp. Assoc., Tree Seed Work. Grp. Newsbull. 54:15-16 (2011).

Sokal, R.R., and F.J. Rohlf, "Biometry," W.H. Freement and Co., San Francisco, CA (1969).

Steel, R.G.D., and J.H. Torrie, "Principles and Procedures of Statistics," McGraw-Hill Book Co., New York, NY (1980).

Stephenson, G.L., "Terrestrial Test Methods for Plants and Soil Invertebrates," Ph.D. Thesis, Department of Environmental Biology, University of Guelph, Guelph, ON, 230 pp. (2003a).

———, Unpublished Data, Stantec Consulting, Ltd., Guelph, ON (2003b).

Stephenson, G.L., K.R. Solomon, B. Hale, B.M. Greenberg, and R.P. Scroggins, *Development of Suitable Test Methods for Evaluating the Toxicity of Contaminated Soils to a Battery of Plant Species Relevant to Soil Environments in Canada*, pp. 474-489, in: "Environmental Toxicology and Risk Assessment: Modeling and Risk Assessment, Sixth Volume," ASTM STP 1317, J. Dwyer, T.R. Doane, and M.L. Hinman (eds.), American Society for Testing and Materials, Philadelphia, PA (1997).

Stephenson, G.L., N. Koper, G.F. Atkinson, K.R. Solomon, and R.P. Scroggins, "Use of Nonlinear Regression Techniques for Describing Concentration-Response Relationships of Plant Species Exposed to Contaminated Site Soils," *Environ. Toxicol. Chem.,* 19:2968-2981 (2000).

Stephenson, G.L., J.I. Princz, N. Koper, and P.G. Miasek, *Toxicity Testing with Volatile Substances*," pp. 236-252, in: "Environmental Toxicology and Risk Assessment: Science, Policy, and Standardization – Implications for Environmental Decisions: 10th Volume," ASTM STP 1403, B.M. Greenberg, R.N. Hall, M.R. Roberts Jr., and R.W. Gensemer (eds.), American Society for Testing and Materials, West Conshohocken, PA (2001).

Stephenson, G.L., R.G. Kuperman, G.L. Linder, and S. Visser, *Toxicity Tests for Assessing Contaminated Soils and Ground Water*, pp. 25-43, in: "Environmental Analysis of Contaminated Sites." G.I. Sunahara, A.Y. Renoux, C. Thellen, C.L. Gaudet, and A. Pilon (eds.), Ecological and Environmental Toxicology Series, John Wiley & Sons Ltd., Chichester, United Kingdom (2002).

Stephenson, G, N. Feisthauer, K. Bessie, and R. Scroggins, "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, BC, April 28-May 2, 2008, hosted by the Real Property Institute of Canada, URL: www.rpic-ibic.ca. Last accessed January 7, 2013 (2008).

Sun, B., J. Tan, Z. Wan., F. Gu, and M. Zhu, "Allelopathic Effects of Extracts from *Solidago canadensis* L. Against Seed Germination and Seedling Growth of Some Plants," *J. Env. Sc.* 18(2):304-309 (2006).

Tesky, J.L., *Calamagrostis canadensis,* in: "Fire Effects Information System," U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer), URL: www.fs.fed.us/database/feis/. Last accessed January 7, 2013 (1992).

Thorpe, J., and B.C. Godwin, "Saskatchewan Rangeland Ecosystems: Ecosite and Communities of Forested Rangelands," Saskatchewan Research Council Pub. No. 11969-1E08, Saskatoon, SK (2008).

Thorpe, J.P., and B.C. Godwin, "Succession and Old-growth Forests in Prince Albert National Park," Saskatchewan Research Council Publication E-2500-3-E-92, Saskatoon, SK (1992).

Thursby, G.B., J. Heltshe, K.J., and K.J. Scott, "Revised Approach to Toxicity Test Acceptability Criteria Using a Statistical Performance Assessment," *Environ. Toxicol. Chem.,* 16:1322-1329 (1997).

Tilley, D.J., *Propagation Protocol for Production of Container* Calamagrostis canadensis *(Michx.) P. Beauv. plants (10-Cubic-Inch Container)*, USDA NRCS – Aberdeen Plant Materials Center, Aberdeen, Idaho, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2010).

TOXSTAT, Version 3.5, Lincoln Research Associates, Inc., PO Box 4276, Bisbee, Arizona 85603, USA, phone 520-432-4092, e-mail danlra@msn.com [Programs on disk, with printed user's manual] (1996).

Uchytil, R.J., *Betula papyrifera*, in: "Fire Effects Information System," U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer) [online], URL: www.fs.fed.us/database/feis/. Last accessed January 7, 2013 (1991a).

———, *Picea mariana*, in: "Fire Effects Information System," U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer) [online], URL: www.fs.fed.us/database/feis/. Last accessed January 7, 2013 (1991b).

USDA-NRCS (United States Department of Agriculture – Natural Resources Conservation Services), "The PLANTS Database," URL: http://plants.usda.gov. Last accessed January 7, 2013, Plants Profile: *Solidago Canadensis*, National Plant Data Center, Baton Rouge, LA 70874-4490 USA (2010).

USDI (United States Department of the Interior), *Digital Representations of Tree Species Range Maps from "Atlas of United States Trees" by Elbert L. Little, Jr. (and other publications)*, URL: http://esp.cr.usgs.gov/data/atlas/little/. Last accessed January 7, 2013 (2006).

USEPA (United States Environmental Protection Agency), "Protocols for Short Term Toxicity Screening of Hazardous Waste Sites," Report EPA/600/3-88/029, prepared by J.C. Greene, C.L. Bartels, W.J. Warren-Hicks, B.R. Parkhurst, G.L. Linder, S.A. Peterson, and W.E. Miller, February 1989, Corvallis Environ. Research Lab., Corvallis, OR (1989).

———, "Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms," report EPA/600/R-95/136, Office of Research and Development, USEPA, Washington, DC, 661 pp. (1995).

———, "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates," 2nd ed. (prepared by the Office of Research and Development, Mid-Continent Ecology Division, USEPA, Duluth, Minnesota, and the Office of Science and Technology, Office of Water, USEPA, Washington, DC), Report EPA 600/R-99/064, 192 pp. (2000).

———, "Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms," 4th ed., Office of Water, Report EPA 821-R-02-013, 350 pp., USEPA, Washington, DC (2002).

———, "Guidance for the Use of Dilution-Water (negative) and Solvent Controls in Statistical Data Analysis for Guideline Aquatic Toxicology Studies," September 25, 2008, 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 (2008).

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

USGS (United States Geological Survey), "Native Wildflowers of the North Dakota Grasslands: Canada Goldenrod *(Solidago canadensis)*," Northern Prairie Wildlife Research Centre, URL: www.npwrc.usgs.gov/resource/plants/wildflwr/species/solicana.htm. Accessed November 2010 (2006).

University of Michigan, "Native American Ethnobotany, A Database of Foods, Drugs, Dyes and Fibers of Native American Peoples, Derived from Plants," http://herb.umd.umich.edu/. Last accessed January 7, 2013 (2003).

van Dam, R.A., A.J. Harford, and M. Warne, "Time to Get Off the Fence: The Need for Definitive International Guidance on Statistical Analysis of Ecotoxicity Data," Debate and Commentary, *Integr Environ Assess Manag* 8(2):242-245 (2012).

van der Grinten, M., *Propagation Protocol for Production of Container* Solidago canadensis *L. Plants*,USDA NRCS – Big Flats Plant Materials Center, Corning, New York, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2002).

van der Hoeven, N., "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 (1998).

Van der Vliet, L., L.N. Taylor, and R. Scroggins, "NOEC: Notable Oversight of Enlightened Canadians: A Response to van Dam et al. (2012)," Letter to the Editor, *Integr. Environ. Assess. Manag.*, 8(3):397-398 (2012).

Velicogna, J., E. Ritchie, J. Princz, M. Lessard, and R. Scroggins, "Ecotoxicity of Siloxane D5 in Soil," *Chemosphere*, 87:77-83 (2012).

Viereck, L.A., and E.L. Little, Jr., "Alaska Trees and Shrubs," Agric. Handb. 410, U.S. Department of Agriculture, Forest Service, Washington, DC, 265 pp. (1972).

Viereck, L.A., and W.F. Johnston, Picea mariana *(Mill.) B.S.P.: Black spruce*, in: Burns, R.M., and B.H. Honkala (technical coordinators), "Silvics of North America: 1. Conifers; 2. Hardwoods," Agriculture Handbook 654, U.S. Department of Agriculture, Forest Service, Washington, DC., vol. 2 877 pp. (1990).

Wang, W., "Literature Review on Higher Plants for Toxicity Testing," *Water, Air and Soil Pollution*, 59:381-400 (1991).

———, "Use of Plants for the Assessment of Environmental Contaminants," *Rev. of Environ. Contam. Toxicol.*, 126:87-127 (1992).

Wang, W., and K. Freemark, "The Use of Plants for Environmental Monitoring and Assessment," *Ecotoxicol. Environ. Safety*, 30:289-301 (1995).

Watson, L.E., R.W. Parker, and D.F. Polster, "Manual of Plant Species Suitability for Reclamation in Alberta Volume II – Forbs, Shrubs and Trees," Volume II, Alberta Land Conservation and Reclamation Council, Government of Alberta Report #RRTAC 80-5, pp. 365-369 (1980).

White A.L., C. Boutin, R.L. Dalton, B. Henkelman, and D. Carpenter, "Germination Requirements for 29 Terrestrial and Wetland Wild Plant Species Appropriate for Phytotoxicity Testing," *Pest Manag. Sci.*, 65:19-26 (2009).

Wick, D., T. Luna, J. Evans, and J. Hosokawa, *Propagation protocol for production of container* Solidago canadensis *L.* salabrosa *(Piper) Jones plants (160-mL containers)*, USDI NPS – Glacier National Park, West Glacier, Montana, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2008).

Williams, D.A., "The Comparison of Several Dose Levels with a Zero Dose Control," *Biometrics,* 28:519-532 (1972).

Winder, R.S., "Evaluation of *Colletotrichum* sp. and *Fusarium* spp. as Potential Biological Control Agents for Marsh Reed Grass *(Calamagrostis canadensis)*," *Can. J. Plant Pathol.*, 21:8-15 (1999).

Wong, D.C.L., E.Y. Chai, K.K. Chu, and P.B. Dorn, "Prediction of Ecotoxicity of Hydrocarbon-Contaminated Soils Using Physicochemical Parameters," *Environ. Toxicol. Chem.,* 18(11):2611-2621 (1999).

Wren, C.D., M.K. Gilbertson, G. Spiers, P. Beckett, D. Stanbury, and M. Kershaw, *ERA Objective #1; Are the Chemicals of Concern Preventing the Recovery of Terrestrial Plant Communities*, in: Wren, C.D. (ed), "Risk Assessment and Environmental Management; A case study in Sudbury, Ontario, Canada," Maralte Publishing, The Netherlands, 454 pp. (2012).

Wynia, R.L., *Propagation Protocol for Production of* Solidago canadensis *Seeds,* USDA NRCS – Manhattan Plant Materials Center, Manhattan, Kansas, in: "Native Plant Network," URL: www.nativeplantnetwork.org. Last accessed January 7, 2013, Moscow (ID): University of Idaho, College of Natural Resources, Forest Research Nursery (2002).

———, "Plant Guide: Bluejoint Reedgrass: *Calamagrostis canadensis* (Michx.) Beaux.," USDA-NRCS Plant Materials Center, Kansas, URL: http://plants.usda.gov/plantguide/pdf/pg_caca4.pdf. Last accessed January 7, 2013 (2006).

Zajdlik and Associates Inc., "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's Biological Methods Section (Ottawa, ON), November (2010).

Zasada, J.C., K. Van Cleve, R.A. Werner, J.A. McQueen, and E. Nyland, *Forest Biology and Management in High Latitude American Forests*, in: "Proceedings of North American Forest Lands at Latitudes North of 60 Degrees: Forest Research and Management at High Latitudes," Proceedings of a symposium, Fairbanks, AK, September 19-22, 1977 (University of Alaska, Fairbanks, Alaska, USA), pp. 137-195 (1978).

Biological Test Methods and Supporting Guidance Documents Published by Environment 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 Publication Catalogue, Environment Canada, Ottawa, Ontario, K1A 0H3, Canada. Printed copies can also be requested by email at: enviroinfo@ec.gc.ca. These documents are freely available in PDF 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 Canada, Ottawa, Ontario, K1A 0H3

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods (*continued*)			
Test for Survival and Growth in Sediment and Water Using the Freshwater Amphipod Hyalella azteca	EPS 1/RM/33 2nd edition	June 2012	—
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	September 2007	—
B. Reference Methods[b]			
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 2nd edition	December 2000	May 2007
Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*	EPS 1/RM/14 2nd edition	December 2000	—
Reference Method for Determining Acute Lethality of Sediment to Marine or Estuarine Amphipods	EPS 1/RM/35	December 1998	—
Reference Method for Determining the Toxicity of Sediment Using Luminescent Bacteria in a Solid-Phase Test	EPS 1/RM/42	April 2002	—

[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 which are described precisely in a written document. Unlike other generic (multi-purpose or "universal") biological test methods published by Environment Canada, the use of a *reference method* is frequently restricted to testing requirements associated with specific regulations.

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
C. Supporting Guidance Documents			
Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants	EPS 1/RM/12	August 1990	—
Guidance Document on Collection and Preparation of Sediment for Physicochemical Characterization and Biological Testing	EPS 1/RM/29	December 1994	—
Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant	EPS 1/RM/30	September 1995	—
Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology	EPS 1/RM/34	December 1999	—
Guidance Document for Testing the Pathogenicity and Toxicity of New Microbial Substances to Aquatic and Terrestrial Organisms	EPS 1/RM/44	March 2004	—
Guidance Document on Statistical Methods 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	—

Environment Canada, Environmental Protection Service, Regional and Headquarters Offices

Headquarters
351 St. Joseph Boulevard
Place Vincent Massey
Gatineau, Quebec
K1A 0H3

Atlantic Region
45 Alderney Drive
Dartmouth, Nova Scotia
B2Y 2N6

Manitoba Office
150–123 Main Street
Winnipeg, Manitoba
R3C 4W2

Quebec Region
1550 d'Estimauville Avenue
Québec, Quebec
G1J 0C3

Pacific and Yukon Region
Vancouver Office
401 Burrard Street
Vancouver, British Columbia
V6C 3S5

Ontario Region
4905 Dufferin St.
Downsview, Ontario
M3H 5T4

Western and Northern Region
Alberta Office
4999 – 98th Avenue
Edmonton, Alberta
T6B 2X3

Yukon Office
91782 Alaskan Highway
Whitehorse, Yukon
Y1A 5B7

Members of the Inter-Governmental Ecotoxicological Testing Group (as of May 2012)

Federal, Environment Canada

Suzanne Agius
Marine Protection Programs Section
Gatineau, Quebec

Fabiola Akaishi
Atlantic Laboratory
for Environmental Testing
Moncton, New Brunswick

Adrienne Bartlett
National Water Research Institute
Burlington, Ontario

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

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

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

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

Garth Elliott
Prairie and Northern Laboratory
for Environmental Testing
Edmonton, Alberta

Chris Fraser
Science and Technology Laboratories
Ottawa, Ontario

François Gagné
Fluvial Ecosystem Research
Montréal, Quebec

Patricia Gillis
Aquatic Ecosystem Protection
Research Division
Burlington, Ontario

Manon Harwood
Quebec Laboratory
for Environmental Testing
Montréal, Quebec

Ryan Hennessy
Science and Technology Laboratories
Ottawa, Ontario

Dale Hughes
Atlantic Laboratory
for Environmental Testing
Moncton, New Brunswick

Paula Jackman
Atlantic Laboratory
for Environmental Testing
Moncton, New Brunswick

Nancy Kruper
Prairie and Northern Laboratory
for Environmental Testing
Edmonton, Alberta

Heather Lemieux
Science and Technology Laboratories
Ottawa, Ontario

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

Danielle Milani
Aquatic Ecosystem Impacts Research Division
Burlington, Ontario

Warren Norwood
Aquatic Ecosystem Protection Research Division
Burlington, Ontario

Heather Osachoff
Pacific and 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
Science and Technology Laboratories
Ottawa, Ontario

Jessica Rahn
Science and Technology Laboratories
Ottawa, Ontario

Ellyn Ritchie
Science and Technology Laboratories
Ottawa, Ontario

Grant Schroeder
Pacific and Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Rick Scroggins
Science and Technology Laboratories
Ottawa, Ontario

Rachel Skirrow
Pacific and Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Troy Steeves
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

David Taillefer
Marine Environmental Protection
Gatineau, Québec

Lisa Taylor (Chairperson)
Science and Technology Laboratories
Ottawa, Ontario

Sylvain Trottier
Quebec Laboratory for Environmental Testing
Montréal, Quebec

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

Leana Van der Vliet
Science and Technology Laboratories
Ottawa, Ontario

Brian Walker
Quebec Laboratory for Environmental Testing
Montréal, Quebec

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

Federal, Fisheries and Oceans Canada

Robert Roy
Institut Maurice Lamontagne
Mont-Joli, Quebec

Federal, Natural Resources Canada

Melissa Desforges
Ecosystem Risk Management Program
Mining and Mineral Sciences Laboratory, CANMET, NRCan
Ottawa, Ontario

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

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

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

Provincial

Richard Chong-Kit
Ontario Ministry of Environment
Etobicoke, Ontario

Kim Hunter
Ontario Ministry of Environment
Etobicoke, Ontario

Mary Moody
Saskatchewan Research Council
Saskatoon, Saskatchewan

David Poirier
Ontario Ministry of Environment
Etobicoke, Ontario

Julie Schroeder
Ontario Ministry of Environment
Toronto, Ontario

Trudy Watson-Leung
Ontario Ministry of Environment
Etobicoke, Ontario

Private Research Facilities/Others

Christian Bastien
Centre d'expertise en analyse
environnementale du Québec
Sainte-Foy, Quebec

Barbara Bayer
ALS Environmental
Winnipeg, Manitoba
Bozena Glowacka
ALS Environmental
Winnipeg, Manitoba

Members of the Scientific Advisory Team

Participants of Boreal Species Priority Ranking

Bob Godwin
Associate Research Scientist
Ecosystems and Forestry
Saskatchewan Research Council
125-15 Innovation Blvd.
Saskatoon SK S7N 2X8
Tel.: 306-933-8190
Email: godwin@src.skc.a

Dick Jiricka, R.P.F.T.
Forest Service (retired)
Forest Practices and Accountability
Saskatchewan Environment
P.O. Box 3003
Prince Albert SK S6V 6G1

Mark Johnston
Distinguished Scientist
Ecosystems and Forestry
Saskatchewan Research Council
125-15 Innovation Blvd.
Saskatoon SK S7N 2X8
Tel.: 306-933-8175
Email: johnston@src.skc.a

Michael McLaughlan, R.P.F.
Forest Ecologist
Forest Service
Forest Practices and Accountability
Saskatchewan Environment
P.O. Box 3003
Prince Albert SK S6V 6G1
Tel.: 306-953-2436
Email: michael.mclaughlan@gov.sk.ca

Jeff Thorpe
Principal Scientist
Ecosystems and Forestry
Saskatchewan Research Council
125-15 Innovation Blvd.
Saskatoon SK S7N 2X8
Tel.: 306-933-8172
Email: thorpe@src.skc.a

Inter-laboratory Testing Participants

Heather Lemieux
Environment Canada
Soil Toxicology Laboratory
Science & Technology Centre
335 River Road South
Ottawa ON K1A 0H3
Tel.: 613-949-0177
Fax: 613-949-1347
Email: heather.lemieux@ec.gc.ca

Darlene Lintott
Exova
7217 Roper Road NW
Edmonton AB T6B 3J4
Tel.: 780-438-5525 Ext 281
Fax: 780-438-8586
Email: darlene.lintott@exova.com

Mary Moody
Saskatchewan Research Council
125-15 Innovation Blvd,
Saskatoon SK S7N 2X8
Tel.: 306-933-5469
Fax: 306-933-7817
Email: moody@src.skc.a

Bryon Shore
HydroQual Laboratories Ltd.
6125-12th Street SE, Unit 4
Calgary AB T2H 2K1
Tel.: 403-253-7121
Fax: 403-252-9363
Email: bryon_shore@golder.com

Gladys L. Stephenson
Stantec
Suite 1, 70 Southgate Dr.
Guelph ON N1G 4P5
Tel.: 519-836-6050 Ext. 219
Fax: 519-836-2493
Email: gladys.stephenson@stantec.com

Tam Vo
Maxxam Analytics
4606 Canada Way,
Burnaby BC V5G 1K5
Tel.: 604-639-8420
Fax: 604-731-2386
Email: tvo@maxxam.ca

Consultants

Jennifer Miller
Miller Environmental Sciences Inc.
48 Curtis Crescent,
King City ON L7B 1C3
Tel.: 905-833-4221
Fax: 905-833-4200
Email: miller.smith@sympatico.ca

Mary Moody
Saskatchewan Research Council
125-15 Innovation Blvd,
Saskatoon SK S7N 2X8
Tel.: 306-933-5469
Fax: 306-933-7817
Email: moody@src.skc.a

Scientific Authorities

Juliska Princz
Environment Canada
Ecotoxicology and Wildlife Health Division
Science & Technology Centre
335 River Road South
Ottawa ON K1A 0H3
Tel.: 613-990-9544
Fax: 613-949-1347
Email: juliska.princz@ec.gc.ca

Richard Scroggins
Environment Canada
Ecotoxicology and Wildlife Health Division
Science & Technology Centre
335 River Road South
Ottawa ON K1A 0H3
Tel.: 613-990-8569
Fax: 613-990-0173
Email: rick.scroggins@ec.gc.ca

Preparation of Paper Birch *(Betula papyrifera)* Seed for Toxicity Testing

Background

During the development of this method, low germination and emergence for paper birch *(Betula papyrifera)* seed was regularly observed by both the Saskatchewan Research Council (SRC) and Environment Canada's Soil Toxicology Laboratory (STL). The SRC and STL undertook an investigation in order to improve emergence of paper birch seed in soil (SRC, 2012; EC, 2013b). Various procedures were investigated, and it was determined that ethanol flotation, followed by hand-selection using a microscope lit from below were the most effective methods of separating debris and empty or defective seeds from the full healthy ones. Removal of the seed wings prior to the ethanol separation further improved germination rates. These procedures are described herein and should be used to prepare paper birch seed for toxicity testing.

De-Winging Seed

Seeds are placed onto cheese cloth and gently rubbed to detach the wings from the seeds (see Figures E-1 and E-2). The debris is then separated from the seeds by hand-sorting.

Figure E-1 Paper birch seed being de-winged in cheese cloth. (photos: H. Lemieux)

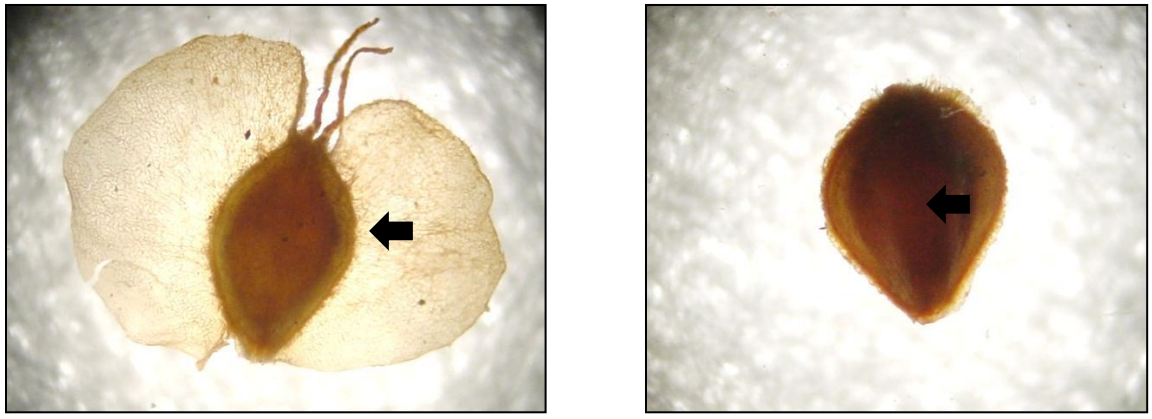


Figure E-2 Paper birch seed with and without wings both fertile, unstratified (25× magnification). Arrows indicate the edge of the embryo. (photos: M. Moody)

Ethanol Floatation

Following de-winging, the seeds are separated using the ethanol-floatation method. The apparatus consists of a funnel, clear tubing, two clamps, a strainer, and a beaker (see Figure E-3). Initially, the top clamp is left open, and the bottom clamp is closed; and ethanol is poured into the apparatus until the level of ethanol reaches halfway up the funnel. The seeds are then immersed into the ethanol that was previously poured into the funnel and stirred to promote separation of the denser (full) seeds.

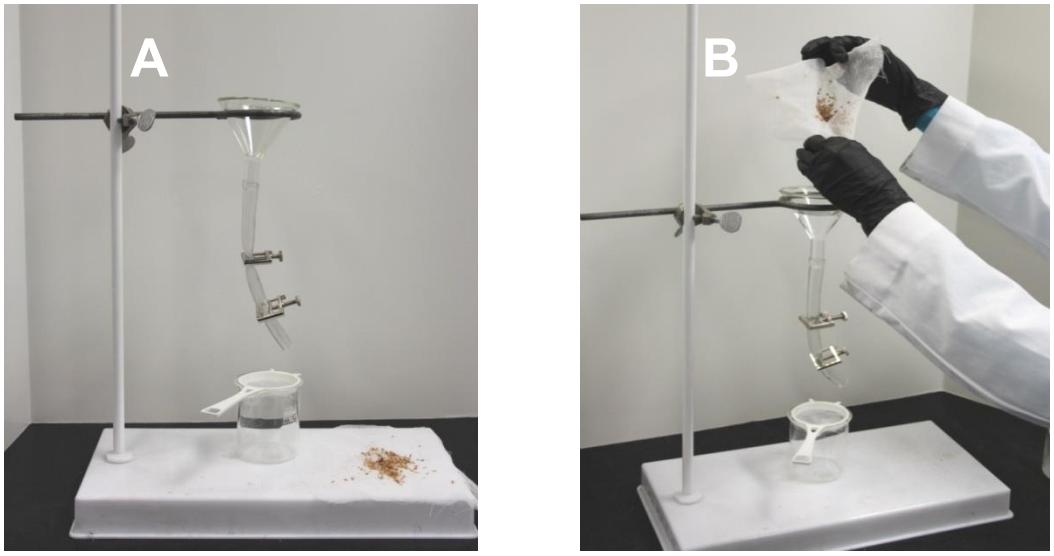


Figure E-3 **A – Ethanol seed separation apparatus. The top clamp is left open, and the bottom clamp closed. B – Ethanol is poured into the funnel, the seeds are then added and the mixture stirred.** (photos: H. Lemieux)

The mixture is then left for approximately 30 seconds for the denser seeds to accumulate at the bottom clamp; the top clamp is then closed. The bottom clamp is then opened to allow the ethanol to flow through the strainer, leaving the seeds on the strainer's surface (see Figure E-4).

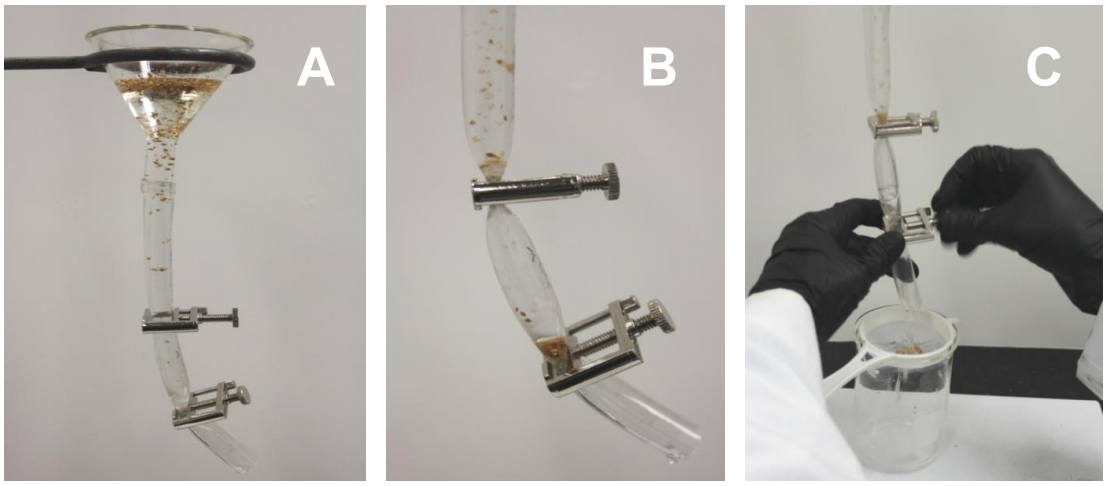


Figure E-4 **A – Denser seeds begin to sink into the tube. B – After 30 seconds, the top clamp is closed. C – The bottom clamp is then released to allow the denser seeds to pour into the strainer.** (photos: H. Lemieux)

The seeds are then immediately rinsed with de-ionized water for approximately 30 seconds and spread out to dry on paper towel (see Figure E-5). The seeds are then ready for stratification.

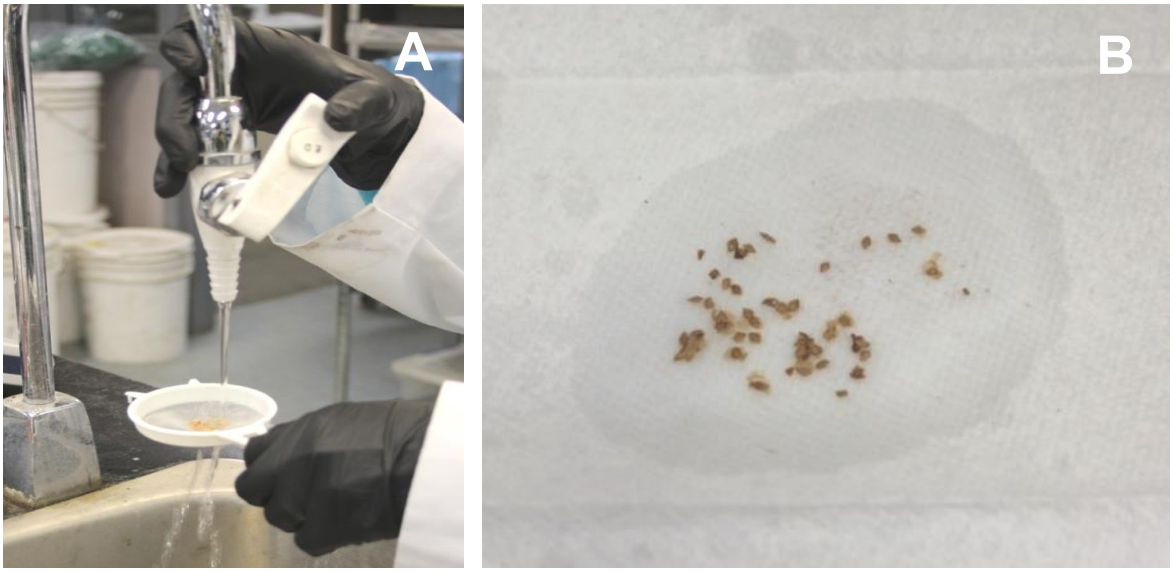


Figure E-5 A– The separated seeds are rinsed with de-ionized water. B – After rinsing seeds in de-ionized water, they are left to dry on paper towel. (photos: H. Lemieux)

Sorting using Microscopic Observation

After the stratification period (see Section 2.3), the seeds are examined and selected using a microscope (illuminated from below) to ensure that only full healthy seeds are selected for testing. Microscopic selection allows the selection of healthy, plump embryos that are more likely to be fertile (Figure E-2). In a fertile seed, the embryo fills the seed coat and the seed appears plump when examined from the side. Figure E-6 illustrates three different seed conditions: fertile, questionable, and infertile seeds. Questionable and infertile seeds possess a shrunken embryo, may be thinner in profile and may also have a dark line visible lengthwise on the seed. These seeds should not be used in testing.

Figure E-6 Visual selection of fertile (left), questionable (centre), and infertile (right) stratified paper birch seed (25× magnification). Arrows indicate the edge of the embryo. (photos: M. Moody)

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 essentially free of any contaminants that could adversely affect the performance of plants during the test (see Section 3.4). Before applying the test method described in this document as a standardized test to be conducted according to Environment Canada, it was necessary to first assess the performance of test organisms in different types of negative control soil representative of an array of *clean* boreal forest soils found within Canada. Nine types of negative control soils were used to develop the biological test method described herein and to further assess the robustness of the test method with samples of soil that varied considerably in their physical and chemical characteristics. These soils were also used to establish reasonable criteria for valid test results, based on control performance. The nine soils tested included an *artificial soil* (see Section 3.4.2) and eight natural soils and soil horizons (see Section 3.4.1) (SRC, 2004, 2006, 2007, 2008, 2009; EC, 2010, 2013b). The *artificial soil* was formulated in the laboratory from natural ingredients. The eight natural boreal forest soils included one from Newfoundland, one from New Brunswick, one from Ontario, three from Saskatchewan, and two from Alberta. The physicochemical characteristics of the artificial soil and eight (including horizons) forest soils are summarized in Table F-1.

The artificial control soil (AS) used in this series of performance evaluation studies with diverse soil types was the same as that recommended for use herein (see Section 3.4.2). It consists of 70% silica sand, 20% kaolin clay, 10% *Sphagnum* sp. peat, and calcium carbonate (10 to 30 g $CaCO_3$/kg peat). The soil was formulated by mixing the ingredients in their dry form thoroughly, then gradually hydrating with de-ionized water, and mixing further until the soil was visibly uniform in colour, texture, and degree of wetness.

The eight natural soils used as negative control soil while developing this biological test method and establishing the test validity criteria herein (see Section 4.4) do not represent all Canadian soil types. However, they do vary greatly in their physicochemical characteristics and include boreal forest soils with diverse textures (see Table F-1). 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 was classified as a Gleyed Humo-ferric Podzol, developed on a stony, loamy-to-sandy, non-calcareous glacial till. 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 *(Gaultheria 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 feather mosses [e.g., Shreber's moss *(Pleurozium schreberi)*, stair-step moss *(Hyloconium splendens)*, and knight's plume *(Ptilium crista-castrensis)*]. Prior to sampling, woody debris and leaf litter were removed, and the under-lying 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.

The New Brunswick soil (NB Podzol) was classified as an imperfectly drained Gleyed Humo-ferric Podzol, developed in non-calcareous, medium to moderately fine-textured basal or lodgement till (EcoDynamics Consulting Ltd., 2008a). The main canopy consisted of a mixed-wood forest, consisting of beech *(Fagus grandifolia)*, red maple *(Acer rubrum)*, yellow birch *(Betula alleghaniensis)* and sugar maple *(Acer saccharum)*, underlain by balsam fir (*Abies balsamea)*, with an understory of hazel *(Corylus cornuta)* and regenerating maple

and balsam fir (EcoDynamics Consulting Ltd., 2008a). The forest litter (L horizon) was removed, and the underlying FH and Ahe-Aegj horizons were collected separately and placed into 25-L pails. The underlying Bf horizon was then collected; however, given the variation and wavy nature of the soil horizon boundaries, the collection of some BCgj material was unavoidable.

The Ontario soil (ON Podzol) was classified as a Gleyed Humo-ferric Podzol developed within a non-calcareous fluvial-lacustrine deposit (EcoDynamics Consulting Ltd., 2011). 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)*, with a lower canopy consisting of a mixture of paper birch *(Betula papyrifera)*, eastern white cedar *(Thuja occidentalis)*, black spruce *(Picea mariana)*, white spruce *(Picea glauca)*, red maple *(Acer rubrum)*, 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 cassinoides)*, velvet blueberry *(Vaccinium myrtilloides),* and twinflower *(Linnaea borealis)*. The ground surface was dominated by bunch berry *(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).

Three soils were collected from Saskatchewan. The first soil (SK01 Luvisol) was classified as a well- to moderately well-drained Dark Grey Luvisol, developed on stone-free, loamy-to-clayey glaciolacustrine materials (EcoDynamics Consulting Ltd., 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).

The second soil (SK02 Brunisol) was classified as a rapidly drained Orthic Eutric Brunisol, developed in a stone-free, sandy glaciofluvial materials (EcoDynamics Consulting Ltd., 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 F and H horizons were 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).

The third soil (SK03 Brunisol) was representative of the Taiga Shield Ecozone and the Selwyn Lake Upland Ecoregion, and was classified as an Eluviated Dystric Brunisol (EcoDynamics Consulting Ltd., 2008b). The upland vegetation was dominated by a black spruce *(Picea mariana)* and jack pine *(Pinus banksiana)*, with an understory of reindeer lichens (mostly *Cladina mitis)* and feather mosses *(*mostly *Pleurozium schreberi)*, Labrador tea *(Ledum groenlandicum)*, bog cranberry *(Vaccinium vitis-idaea)*, blueberry *(Vaccinium myrtilloides)*, bog bilberry *(Vaccinium uliginosum)*, and crowberry *(Empetrum nigrum)*. The surface woody debris and leaf litter were removed to expose the F and H horizons, which were then collected and placed into 25-L pails. Subsequently, the underlying A (Ae) and B (Bfj and Bm) mineral horizons were collected together, as their combined depth was approximately 10 cm thick.

Two soils were collected from Alberta. The first 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 Ltd., 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).

The second soil (AB02 Chernozem), was collected on a river floodplain terrace, and was characterized as a well- to moderately well-drained Rego Dark Gray Chernozem (EcoDynamics Consulting Ltd., 2007). The texture of the

organic-rich Ah horizon was classified as a silt loam, with a very-fine-sand/loamy-very-fine-sand to very-sandy loam texture occurring with depth. The dominant vegetation consisted of smooth brome *(Bromus inermis* Leyss.*)* interspersed with small amounts of rose *(Rosa* sp.*)*, northern bedstraw *(Galium boreale* L.*)* and fireweed *(Chamerion angustifolium)*. Forested areas close to the river valley slopes contained an aspen over-story, with scattered white spruce. Two horizons were also collected, the Ah horizon to a depth of 11 cm, and the Ckgj horizon to a depth of approximately 25 to 30 cm; there was no defined B horizon.

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

		Soil type:	**Artificial soil**	**NFLD01 podzol**			
		Source:	In-house	Newfoundland			
		Soil classification:	n/a	Gleyed humo-ferric podzol			
		Horizon:	n/a	FH	Ahe	Ae	Bf
Parameter	**Units**	**Analytical method**					
Soil Texture[b]		n/a[c]	SL	-	-	-	-
Sand	%	Particle size distribution (filter candle system)	76	-	-	-	-
Silt	%		12	-	-	-	-
Clay	%		12	-	-	-	-
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
Conductivity	mS/cm	Saturated paste method	-[4]	-	-	-	-
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	-	-	-	-
NO_3-N	mg/kg		5	-	-	-	-
NO_2-N	mg/kg		< 1	-	-	-	-
Phosphorous (total)	%		0.03	-	-	-	-
Phosphorous	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:	**NB Podzol**		**ON Podzol**		
		Source:	New Brunswick		Ontario		
		Soil classification:	Gleyed Humo-ferric Podzol		Gleyed Humo-ferric Podzol		
		Horizon:	A	B	A	Ae	B
Parameter	**Units**	**Analytical method**					
Soil Texture[b]		n/a[c]	SCL	SL	LS	LS	LS
Sand	%	Particle size distribution (filter candle system)	79	62	82	88	86
Silt	%		1	28	12	6	6
Clay	%		20	10	6	6	8
Water-holding capacity	%	EC (2005a)	67.6	80.6	41.0	181.9	40.9
Optimal moisture content	%		65.0	65.0	65.0	52.5	47.5
pH	units	1:1 water method	4.7	4.6	4.6	4.6	5.8
Conductivity	mS/cm	Saturated paste method	0.23	0.06	-	-	-
Organic carbon	%	Leco furnace method	41.1	3.7	32.1	1.6	1.0
Organic matter	%	Loss on ignition	77.1	10.9	58.1	2.1	2.2
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method			26	9	12
Total nitrogen	%	Kjeldahl method	1.72	0.23	0.96	0.06	0.05
NH_3	mg/kg	2N KCL extractable	783	19	128	4	2
NO_3-N	mg/kg		3	9	< 1	< 1	< 1
NO_2-N	mg/kg		-	-	< 1	< 1	< 1
Phosphorous (total)	%		-	-	-	-	-
Phosphorous	mg/kg	$NaHCO_3$ extractable	99	18	16	2	< 2
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis	917	1030	143	23	16
Magnesium	mg/kg		784	6560	151	31	40
Calcium	mg/kg		4190	608	765	184	191
Sodium	mg/kg		128	< 100	57	35	21
C/N			23.9	16	33.4	26	20.6
Sodium adsorption ratio		Saturated paste method	1.8	1.2	2.0	2.8	2.4

		Soil type:	SK01 Luvisol			SK02 Brunisol		SK03 Brunisol	
		Source:	Saskatchewan			Saskatchewan		Saskatchewan	
		Soil classification:	Dark grey luvisol			Orthic eutric brunisol		Eluviated dystric brunisol	
		Horizon:	LFH	Ahe	Bt	FH	AB	FH	AeB
Parameter	**Units**	**Analytical method**							
Soil Texture[b]		n/a[c]	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	70.0	41.7
Optimal moisture content	%		55.0	52.5	42.5	55.0	45.0	63.5	55.0
pH	units	1:1 water method	6.6	6.4	6.6	6.9	6.8	4.5	5.6
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	8.2	2.5
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method	43	22	11	22	6	19	7
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	-	-
Phosphorous (total)	%		0.18	0.14	0.06	0.05	0.02	-	-
Phosphorous	mg/kg	$NaHCO_3$ extractable	56	62	9	24	16	17	5
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis	411	363	170	200	83	70	20
Magnesium	mg/kg		586	315	198	785	196	20	10
Calcium	mg/kg		7260	3540	1780	2860	795	< 100	< 100
Sodium	mg/kg		93	100	67	64	50	30	20
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	-	-

		Soil type:	AB01 Gleysol		AB02 Chernozem	
		Source:	Alberta		Alberta	
		Soil classification:	Rego humic gleysol		Rego dark grey chernozem	
		horizon:	Of/Oh	Ahg	Ah	Ck
Parameter	**Units**	**Analytical method**				
Soil Texture[b]		n/a[c]	Peat	SL	SL	SL
Sand	%	Particle size distribution (filter candle system)	n/a	59	51	71
Silt	%		n/a	33	43	24
Clay	%		n/a	8	6	6
Water-holding capacity	%	EC (2005a)	248.1	73.9	68.3	51.4
Optimal moisture content	%		100.0	70.0	55.0	47.5
pH	units	1:1 water method	3.9	4.3	7.1	7.7
Conductivity	mS/cm	Saturated paste method	0.38	0.1	0.34	0.2
Organic carbon	%	Leco furnace method	34.6	11.3	6.3	1.5
Organic matter	%	Loss on ignition	67.8	21.5	9.5	2.6
Cation exchange capacity	$Cmol^+/kg$	Barium chloride method	27	39	25	16
Total nitrogen	%	Kjeldahl method	2	0.63	0.43	0.09
NH_3	mg/kg	2N KCL extractable	114	9	2	1
NO_3-N	mg/kg		3	9	15	1
NO_2-N	mg/kg		< 1	< 1	< 1	< 1
Phosphorous (total)	%		-	-	-	-
Phosphorous	mg/kg	$NaHCO_3$ extractable	28	33	17	8
Potassium	mg/kg	NH_4 acetate extraction, colourimetric analysis	53	81	430	203
Magnesium	mg/kg		66	108	431	235
Calcium	mg/kg		462	570	3380	2400
Sodium	mg/kg		57	28	-	12
C/N			17.3	-	14.6	16.2
Sodium adsorption ratio		Saturated paste method	0.9	1.3	1.2	1.2

[a] 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 in this test method document (SRC 2004, 2006, 2007, 2008, 2009; EC and SRC 2007; EC 2010, 2013b).
[b] SL = sandy loam; LS = loam sand; SCL = sandy clay loam; L = loam.
[c] Not applicable.

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 percentage by weight (e.g., mg/kg) or weight-to-volume (e.g., mg/L). 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.

www.**ec.gc.ca**

Additional information can be obtained at:

Environment Canada
Inquiry Centre
10 Wellington Street, 23rd Floor
Gatineau QC K1A 0H3
Telephone: 1-800-668-6767 (in Canada only) or 819-997-2800
Fax: 819-994-1412
TTY: 819-994-0736
Email: enviroinfo@ec.gc.ca