Technical Manual: FEMMtool Version 1.0

Hourglass Climate

October 2026

Disclaimer: FEMMtool assessments are generated automatically from user inputs, third-party data and modelling, and is provided for general information only. It is not professional environmental, engineering, legal or financial advice and may contain errors or omissions. Have a qualified professional verify all results before relying on them. Hourglass Climate accepts no liability for any loss or damage arising from use of or reliance on FEMMtool assessments.

DRAFT

Table of Contents

Glossary 4
1. Introduction 7
1.1 Scope and purpose 7
1.2 Design philosophy 8
1.3 Workflow summary 10
1.4 Relation to US EPA and CCME methodology 11
2. Prerequisites for using FEMMtool 15
2.1 Project 15
2.2 Compartments 15
2.3 Stressors 16
2.4 Predicted and Measured Environmental Concentrations (PECs and MECs) 16
2.5 Study climate zone 17
3. Data sourcing and requirements 18
3.1 ECOTOX Database 18
3.2 FEMMtool database requirements 18
3.2.1 Stressor-CAS mapping 18
3.2.2 Species identification 22
3.2.3 Inclusion criteria 22
3.2.4 Units 23
4. Data extrapolation 24
4.1 Extrapolation factors 24
4.2 Non-negligible percent effects 25
4.3 Acute-to-chronic extrapolation 25
4.3.1 Framework-specific and default criteria 25
4.3.2 Default criteria 26
4.3.3 CCME criteria 27
4.3.4 Unreported life stage 28
4.3.5 Acute-to-chronic ratio 29
4.4 Cross-medium extrapolation 29
5. Data selection 30
5.1 Endpoint preference 31
5.2 Default (US EPA): Saltwater 31
5.3 Default (US EPA): Freshwater 32
5.4 Canada: Saltwater 33
5.5 Canada: Freshwater 35
6. Risk assessment outputs 36

DRAFT

6.1 Assessment factor **36**
6.1.1 Taxonomic completeness scoring **36**
6.1.2 Datasets that cannot be assessed **39**
6.2 PNEC **39**
6.3 Risk quotient and risk level **39**
6.4 Confidence **40**
6.4.1 Definition **40**
6.4.2 Interpreting risk level and confidence together **41**
6.4.3 Acceptable outcomes **41**
6.5 Mixtures **41**
7. SSD fitting procedure **43**
7.1 Minimum dataset **43**
7.2 Candidate distributions and maximum likelihood **43**
7.3 Model averaging **44**
7.4 Goodness of fit and diagnostics **44**
7.5 Uncertainty **45**
8. Software architecture **45**
8.1 Components **45**
8.2 Provenance and reproducibility **46**
8.3 Administrative configuration **47**
8.4 Report structure **47**
9. Future directions **48**
Literature cited **49**

DRAFT

Glossary

Term	Definition
ACR	Acute-to-chronic ratio. The extrapolation factor applied to an acute effect concentration to estimate a chronic value (default 10)
AF	Assessment factor. The divisor applied to the hazard concentration for 5% of species to derive the predicted no-effect concentration. In FEMMtool it reflects gaps in taxonomic coverage of the species sensitivity distribution
AIC / AICc	Akaike Information Criterion / small-sample corrected Akaike Information Criterion. Measures of relative model fit used to weight candidate distributions
AD, CvM, KS	Anderson–Darling, Cramér–von Mises and Kolmogorov–Smirnov goodness-of-fit statistics
CAS	Chemical Abstracts Service registry number. Identifies a specific chemical substance
CCME	Canadian Council of Ministers of the Environment
Confidence Level	A rating (High, Moderate, Low or Unknown) of how robust a risk level is to statistical uncertainty in the species sensitivity distribution, based on where the predicted environmental concentration falls relative to the confidence bounds of the threshold (§6.4)
Compartment	A part of a project site with its own exposure dynamics (e.g. water column, porewater). Each compartment has its own predicted environmental concentration
ECOTOX	US EPA ECOTOXicology Knowledgebase. The toxicity data source for FEMMtool

DRAFT

Term	Definition
ECx / LCx	Concentration expected to cause an effect (EC) or lethality (LC) in x% of test organisms
EF	Extrapolation factor. A divisor applied to an individual effect concentration to standardize it to a chronic, negligible-effect, same-medium basis. FEMMtool uses three extrapolation factors: the percent-effect factor, the acute-to-chronic ratio and the cross-medium factor. The term is analogous to “uncertainty factor”, “safety factor” and “conversion factor” in other frameworks. FEMMtool uses “extrapolation factor” to keep these record-level factors distinct from the assessment factor applied to the hazard concentration
Total EF	The product of all extrapolation factors applied to one record. Unextrapolated records have a total extrapolation factor of 1
FEMM	Framework for Ecotoxicological Modeling of mCDR (marine carbon dioxide removal)
HC5	Hazard concentration for 5% of species (5th percentile of the species sensitivity distribution). Dividing it by the assessment factor gives the predicted no-effect concentration
HC10	Hazard concentration for 10% of species. Used in FEMMtool only to grade the size of an exceedance of the predicted no-effect concentration (§6.3). It is not an acceptability threshold
LOEC / NOEC	Lowest observed effect concentration / no observed effect concentration

DRAFT

Term	Definition
MATC	Maximum acceptable toxicant concentration (geometric mean of the no observed and lowest observed effect concentrations)
mCDR	Marine carbon dioxide removal
MEC	Measured environmental concentration. A monitored concentration that can replace the predicted environmental concentration for post-deployment impact assessment
MRV	Monitoring, reporting and verification
PEC	Predicted environmental concentration. The user-supplied exposure estimate for a stressor in a compartment
PNEC	Predicted no-effect concentration. The hazard concentration for 5% of species divided by the assessment factor
RQ	Risk quotient. The predicted environmental concentration divided by the predicted no-effect concentration. A value above 1 indicates potential for ecological harm
SSD	Species sensitivity distribution. A statistical distribution fit to one effect concentration per species, describing the range of species sensitivity to a stressor

DRAFT

1. Introduction

1.1 Scope and purpose

Marine carbon dioxide removal (mCDR) is a new and rapidly evolving field. Approaches such as ocean alkalinity enhancement (OAE) alter seawater carbonate chemistry, and the mineral feedstocks used in some OAE approaches release trace metals as they dissolve. Deploying these technologies raises ecological questions that call for rigorous, science-based environmental risk assessment (ERA), yet no mCDR-specific regulatory process for ERA exists.

Ecotoxicology already offers mature, regulator-accepted methods for exactly this kind of question. The **Framework for Ecotoxicological Modeling of mCDR** (**FEMM**; Weighman et al. 2026) rests on the premise that these existing ERA methods should be adopted, and where necessary adapted, for mCDR rather than invented anew. Assessments built this way draw on decades of established practice, fit within current permitting and **Measurement, Reporting and Verification (MRV)** processes, and can be understood by regulators who already apply the same methods to other stressors.

Central to FEMM's ERA guidance are **species sensitivity distributions (SSDs)**. SSDs are used to derive water quality guidelines in Canada (CCME 2007), Australia and New Zealand (Warne et al. 2025), and China (MEE 2022), and in the European Union as a higher-tier alternative to **assessment factors (AFs)** (European Commission 2018). The US EPA uses a related distribution-based approach based on genus mean values (Stephan et al. 1985). For an overview of the SSD method, see Posthuma et al. (2002). In SSDs, toxicity test results for a stressor are compiled across several species, a statistical distribution is fit to them, and a low percentile of that distribution, typically the 5th percentile (**hazard concentration, HC5**), is taken as the concentration expected to protect the remaining 95% of species. FEMM's screening-level risk assessment (Weighman et al. 2026, Section 6) uses the SSD to derive a **predicted no-effect concentration (PNEC)** for each stressor and compares it with the **predicted environmental concentration (PEC)**. Where risk is indicated, FEMM describes an advanced risk assessment tier in which mechanistic models refine the exposure and effect estimates (Weighman et al. 2026, Section 7).

FEMMtool (femmtool.org) currently covers the screening-level tier only and was developed to make FEMM guidance accessible to, and implementable for, non-specialists. Constructing an SSD involves many technical decisions, from which studies to include to how to fit and interpret the distribution. FEMMtool makes those decisions according to the framework, so that a project planner can complete a screening-level assessment and be confident that risk has been evaluated to an accepted standard, and a regulator can review the result knowing exactly how it was produced. A unified tool also establishes consistent risk assessment metrics for mCDR projects: results can be compared across projects and sites, regulators can evaluate submissions more readily, and there is less incentive to seek out jurisdictions with less demanding regulatory requirements.

Version 1.0 of FEMMtool assesses trace metals released by mCDR feedstocks, the stressors for which regulatory ERA methods are most directly applicable. For each metal a project

DRAFT

may release, the tool builds a chronic SSD from published toxicity data, derives an HC5, applies an AF to obtain a PNEC, and compares each user-supplied PEC with the PNEC. The result is classified as Low, Moderate or High Risk, with a **confidence level**, and documented in a report suitable for regulatory review during permitting and for MRV. The same pathway can be run after deployment with **measured environmental concentrations (MECs)** in place of PECs (Weighman et al. 2026, Section 5).

FEMMtool's outputs never supersede regulatory criteria. Where a water-quality criterion for a stressor exists in the project's jurisdiction, a predicted exceedance of that criterion requires project redesign regardless of the FEMMtool result (Weighman et al. 2026, Section 5). FEMMtool is a screening tool for the many stressors and settings where no such criterion exists or where a criterion does not address the project's specific exposure.

This manual describes the scientific and statistical methods behind FEMMtool and the rationale for the choices made where the framework, the literature, or regulatory guidance allow more than one approach. The User Guide gives an overview of these methods alongside instructions for using the web interface; this manual is the exhaustive companion to that overview. It is intended to give regulators, risk assessors and verifiers the detail they need to evaluate a FEMMtool report, and to give project planners a fuller account of how their results were produced.

1.2 Design philosophy

The SSD procedure rests on many judgement calls: which data are admissible, how to treat acute or lethal endpoints, which distribution to fit, how large an AF to apply. Expert tools such as the US EPA SSD Toolbox (Etterson 2020) leave those choices to the analyst. FEMMtool takes the opposite approach. It is designed for users who need a dependable screening result rather than a statistical workbench, and it is built on the following principles.

Fixed methodology, minimal user input. The user supplies a project (name, location, project type, exposure medium), one or more stressors, one or more **compartments**, and a PEC for each stressor in each compartment (§2). Everything else follows rules encoded in FEMMtool v1.0 and described in this manual: data inclusion, extrapolation, endpoint preference, data selection, distribution fitting, AFs and risk classification. None of these can be changed by the user. A small number of methodological settings, such as the SSD fitting algorithm, exist as administrative configuration that is accessible only to administrators (§8.3). In this manual, "default" means the value the tool uses, not a user-selectable option.

Conservative by construction. FEMM v1.0 holds that assumptions in mCDR risk assessment should default toward conservatism and be clearly documented (Weighman et al. 2026, Sections 5 and 6.1). FEMMtool applies this at every step where a judgement would otherwise be required. Only chronic, low-effect, laboratory data on accepted endpoints enter the SSD, and any record extrapolated toward that standard is divided by an **extrapolation factor (EF)** (§4). The most protective endpoint is selected within each species (§5.1). A graduated AF penalises taxonomic gaps rather than declining to answer (§6.1), provided the dataset contains at least one primary producer, one invertebrate and one chordate; without all three, no SSD is fitted (§6.1.2). Only a PEC at or below the PNEC is

DRAFT

reported as acceptable (Low Risk); a small exceedance is graded as Moderate Risk. While a Moderate Risk rating may be acceptable in degraded project sites under certain regulatory frameworks, a Moderate Risk grading should generally be interpreted as indicating that a project may achieve a Low Risk rating with minimal redesign. (§6.3). The confidence level requires a Low Risk result to hold at the lower confidence bound of the PNEC (§6.4). Taken together, these choices mean FEMMtool will tend to overstate risk relative to a bespoke expert assessment. That is intentional: a false Low is the outcome the tool is built to avoid.

Transparency in place of interactivity. What the user cannot adjust, the report must show. Every candidate record carries its inclusion outcome and reason code; every selected value carries its associated EFs (§6.4); every SSD carries its candidate weights, goodness-of-fit (GoF) statistics, parameter uncertainty and leave-one-out sensitivity; every AF carries the completeness scores that produced it. A reviewer with ecotoxicological expertise can therefore audit, and if necessary contest, any step in the assessment.

Based on established regulatory methods. . FEMM's overall structure follows the architecture common to regulatory ERA schemes, and its technical components draw on guidance from several regulatory bodies (Weighman et al. 2026, Section 2.5); FEMMtool implements that workflow. FEMM, and FEMMtool, do not follow a single regulatory framework because no one framework was designed for the combination of sparse marine data, multiple stressors, and multi-jurisdictional project sites that are relevant for holistic mCDR risk screening. Elements common to most frameworks are adopted as they are generally practiced: SSDs, for example, are the standard basis for water-quality thresholds in the United States, Canada, the European Union and Australia/New Zealand. Where an element is more fully developed in one framework than in others, FEMMtool adopts that framework's treatment: AFs appear in both CCME guidance, as a safety factor applied to the lowest endpoint when data are insufficient for an SSD, and ECHA guidance, but ECHA sets out explicit criteria for choosing the factor in relation to the SSD dataset (§6.1). The aim is a single, coherent workflow that meets the most comprehensive and rigorous standard available for each step. Where a step follows or adapts a particular source, the relevant section of this manual cites it.

Location-specific criteria where protection goals differ. Some differences between frameworks reflect not rigor but what is being protected. The taxonomic requirements for a complete dataset, in particular, reflect the species assemblages and ecological priorities of the region a framework serves; CCME, for example, gives preference to temperate species. Because species sensitivity can differ between regions, and SSDs built from regionally relevant species better protect local biodiversity (Weighman et al. 2026, Section 6.3), the risk posed by the same release may genuinely differ between locations, and FEMMtool's results should reflect that. For these criteria FEMMtool therefore defers to the regulatory framework of the project's location, while the structure around them (the number of species required and how completeness is scored and converted to an AF) is the same everywhere. Where a location's framework has not yet been implemented in the tool, FEMMtool applies default criteria drawn from the framework that addresses that element most comprehensively. The frameworks currently implemented, and the defaults, are described in the relevant sections of this manual.

DRAFT

FEMMtool does not determine whether a project is permissible, and it does not produce jurisdiction-specific regulatory criteria. It provides a standardized, reproducible, conservative screening result and the diagnostics needed to judge how much weight that result can bear.

1.3 Workflow summary

The user supplies a project (its location, which determines the criteria that apply (§2.1); project type; and exposure medium), one or more compartments (water column, porewater, sediment, etc.), the stressors to assess, and a PEC (or MEC) for each stressor in each compartment (§2). FEMMtool then performs the following eight steps automatically for each stressor. The step numbers match Figure 1.

1. **Data compilation and inclusion screening.** Candidate toxicity records for the stressor are drawn from the FEMMtool database, which is built from US EPA **ECOTOX** using the stressor's **CAS** identifiers, and each record is tested against the FEMMtool inclusion criteria; failures are recorded with a reason code (§3).
2. **Extrapolation to a chronic, negligible-effect, same-medium basis.** Each retained record is standardised by dividing by the applicable EFs; records with a **total EF** above 100 are dropped (§4).
3. **One value per species.** Within each species the most protective permissible endpoint is selected, and where several values are equally preferred the lowest is used; unextrapolated values are preferred over extrapolated ones (§5.1).
4. **Minimum data and taxonomic requirements.** The species set is assembled against the eight taxonomic requirements that apply to the project's location and medium; extrapolated species are added only to fill requirements the unextrapolated data leave unmet, and SSDs with fewer than ten species are flagged (§5, §7.1).
5. **Distribution fitting.** Six candidate distributions are fitted by maximum likelihood and combined by **AICc**-weighted model averaging, with goodness-of-fit and leave-one-out diagnostics (§7.2 to §7.4).
6. **Hazard concentration.** The HC5 is read from the model-averaged distribution with 95% and 90% bootstrap confidence intervals (§7.5).
7. **Assessment factor.** Taxonomic completeness of the final species set is scored and converted to an AF, and PNEC = HC5 / AF (§6.1, §6.2).
8. **Comparison and decision.** Each compartment's PEC is compared with the PNEC to determine the risk level, and an RQ and, for Low Risk results, a confidence level are reported (§6.3, §6.4).

A PDF report and CSV data tables documenting every step are generated (§8).

1.4 Relation to US EPA and CCME methodology

FEMMtool's workflow follows the same broad sequence that US EPA (Stephan et al. 1985) and CCME (2007) prescribe for deriving water-quality criteria and guidelines from toxicity data. Figure 1 places the two regulatory methods on either side of the FEMMtool workflow, so that a reviewer familiar with one of them can see where FEMMtool does the same thing,

DRAFT

where it automates a step the guidance leaves to expert judgement, and where it has no direct counterpart.

Two FEMMtool steps stand alone. Extrapolation (Figure 1, step 2) exists because data availability is the main barrier to fitting SSDs for mCDR stressors; FEMMtool's position is that a standardised SSD built with conservative, fully tracked extrapolation is preferable to abandoning the SSD for a lower-tier method, and neither regulator has an equivalent record-level step. The graduated AF (Figure 1, step 7) has a partial counterpart only in CCME's Type B derivation, since US EPA derives no criterion when its minimum data requirements are unmet.

Distribution fitting (Figure 1, step 5) is where FEMMtool departs most from both frameworks. US EPA fits a regression through the four lowest-ranked Genus Mean Acute Values and takes the 5th percentile (Stephan et al. 1985, Section IV), and CCME fits candidate distributions to the SSD and selects the best-fitting model by statistical fit, visual inspection and professional judgment (CCME 2007, Part II, Section 3.1), whereas FEMMtool uses model averaging over a fixed candidate set, which removes the choice of distribution from the analyst. The final step differs in purpose rather than method: the regulatory documents produce a criterion for others to apply, while FEMMtool applies its PNEC to the project's own PEC and reports the result.

DRAFT

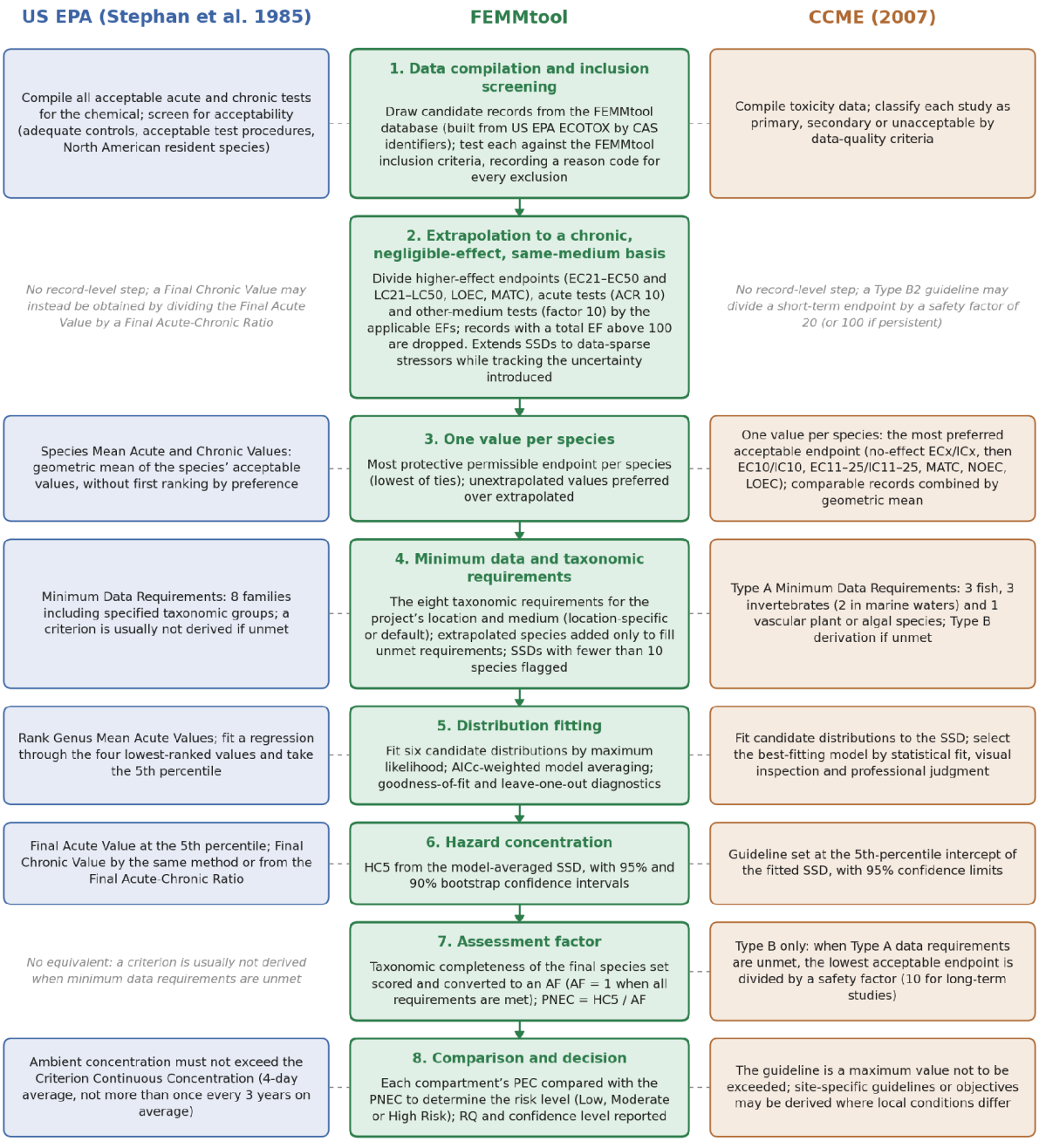


Figure 1. The FEMMtool screening workflow (centre; steps 1 to 8 as described in §1.3) alongside the corresponding steps of the US EPA guidelines for deriving national water-quality criteria (Stephan et al. 1985; left) and the CCME protocol for deriving water-quality guidelines (CCME 2007; right). Dashed lines link corresponding steps. Where a framework has no direct counterpart to a FEMMtool step, an italic note gives its nearest equivalent. The regulatory summaries are deliberately brief and omit each method's special cases; §1.4 discusses the main points of difference.

2. Prerequisites for using FEMMtool

FEMMtool asks the user for four inputs: a project, its **compartments**, the stressors to assess, and a **PEC** (or, after deployment, a **MEC**) for each stressor in each compartment. One optional setting, the study climate zone, is also available in some jurisdictions. This section describes each input, what it is used for, and what the user is expected to bring from their own project planning.

2.1 Project

A name, an optional description, and a location given as latitude and longitude and resolved to a country. The project location determines which taxonomic requirements (§5) and acute/chronic exposure durations (§4.3) apply: those of the regulatory framework for that location, where it has been implemented in the tool, and FEMMtool's defaults otherwise (§1.2). This is not chosen by the user. Projects must currently be located in the United States, Canada, the European Union or Australia/New Zealand; other locations cannot be assessed. The project type (e.g. OAE) is recorded for the report. The exposure medium (marine or freshwater) determines which toxicity data are treated as same-medium (§4.4) and which taxonomic completeness criteria apply (§5).

2.2 Compartments

A compartment is a part of the project site, such as the water column, sediment or porewater, in which exposure to a stressor is expected to differ from elsewhere on the site. Concentrations depend on where the stressor enters, how it disperses and how long it resides, so a single project can produce quite different exposures in different compartments, and each compartment therefore has its own PEC.

FEMM asks planners to identify the compartments relevant to their site and, in particular, the one in which concentrations are likely to be highest (Weighman et al. 2026, Section 5). This is often the compartment into which the stressor is released because dilution is lowest. FEMM treats the PEC for the highest-risk compartment as the project PEC for assessing risk at the community level, since a Low Risk result there is generally protective of the site as a whole. Other compartments may also be assessed, for example where a lower-exposure compartment is the one most relevant to a species of special concern, and the compartments identified before deployment provide natural points of comparison for post-deployment monitoring (§2.4).

In FEMMtool, the user defines one or more compartments and enters a PEC for each stressor in each compartment. Each compartment receives its own risk level and, where risk is Low, a confidence level (§6.3, §6.4), while the **SSD** and **PNEC** for a stressor are shared across compartments.

2.3 Stressors

A stressor is a substance for which FEMMtool can build an SSD and derive a PNEC. Version 1.0 supports thirteen trace metals: aluminum, arsenic, cadmium, chromium(III), chromium(VI), cobalt, copper, iron, lead, manganese, nickel, vanadium and zinc. Users

DRAFT

select stressors from this list; a stressor is listed only once its **ECOTOX** data mapping has been configured and validated (§3.2.1), and users cannot define new stressors or change how an existing one is mapped to toxicity data.

Identifying which stressors a project may release is the planner's responsibility. FEMM v1.0 (Weighman et al. 2026, Section 3 and Appendix A) gives an overview of stressors associated with some current **mCDR** approaches, which can serve as a starting point but is not exhaustive. For mineral-based OAE, the stressors of interest are the dissolution products of the feedstock, which vary with the mineral and may include iron, magnesium, manganese, nickel, chromium, cobalt, vanadium and selenium. A chemical characterisation of the specific feedstock and batch is therefore the appropriate starting point. FEMMtool assesses only the stressors listed above; a stressor that is not listed cannot be assessed by substituting a different one. The stressor list is maintained by administrators and is expected to grow as toxicity data for further stressors are compiled and validated (§9); the availability of further relevant trace metals and carbonate-chemistry parameters as stressors is planned for a later version.

For some of these metals, marine toxicity data are sparse. In such cases the **AF** will tend to be large (§6.1) and the confidence level low (§6.4), and the report will show which values rest on extrapolation (§4.1). This is the intended behaviour: the result reflects what the data can support.

2.4 Predicted and Measured Environmental Concentrations (PECs and MECs)

The PEC is the concentration of a stressor predicted to occur in a compartment (§2.2) as a result of the project. FEMMtool does not model dispersion or fate, so the PEC must come from the project's own exposure modelling and is entered as an input. For each stressor in each compartment, the user enters a concentration and its unit (mg/L, µg/L, ng/L, g/L, ppm or ppb), which the tool converts to mg/L.

FEMM v1.0 (Weighman et al. 2026, Section 5) gives a conceptual overview of how a screening PEC should be established and refers readers to established exposure assessment guidance for the detail: ECHA's guidance on environmental exposure assessment (ECHA 2016) and US EPA's guidelines for exposure assessment and ecological risk assessment (US EPA 1992, 1998). Three points from FEMM's overview bear directly on how FEMMtool results should be read. First, the screening PEC is intended to be a maximum plausible, time-averaged concentration: a plausible worst case for the project design and site, averaged over a period consistent with the chronic exposures the SSD represents rather than an instantaneous peak. Second, FEMM recommends that background concentrations and other sources of the stressor should be added to the project's contribution, since acclimation of local organisms to elevated background cannot be assumed. Third, the methodology used to derive PEC values should default towards conservatism, and this process and its inherent assumptions should be thoroughly documented. Because the protective margins built into FEMMtooltool do not compensate for under-estimated exposure, the veracity of user-input PEC values has direct implications for the reliability of FEMMtool assessment outcomes. As such, it is recommended that this documentation be presented together with associated FEMMtool reports.

DRAFT

After deployment, FEMMtool can be run with MECs from monitoring in place of PECs. FEMMtool accepts a MEC time series for a compartment and summarises it as the maximum four-day rolling mean, consistent with the averaging period used for chronic criteria under US EPA guidance; the same four-day window is applied in all jurisdictions. Where MECs remain at or below the corresponding PECs, a pre-deployment finding of Low Risk is supported; where a MEC exceeds its PEC, re-running the assessment with the MEC provides an assessment of realised impact. This comparison is most useful when the compartments and averaging periods used for the PEC match those used in monitoring.

2.5 Study climate zone

For jurisdictions that span both temperate and tropical regions (currently the United States), the user may select a study climate zone for the assessment. If one is selected, the tool prioritises toxicity data from temperate or from tropical species when selecting the dataset (§5; Weighman et al. 2026, Section 6.3), in the same way that the Canadian framework prioritises temperate species by default. If none is selected, no climate preference is applied. The selection does not change the inclusion criteria or the AF criteria.

DRAFT

3. Data sourcing and requirements

The SSD for each stressor is built from published toxicity test results. This section describes where those results come from, how they are organised in FEMMtool, and which records are eligible for use.

3.1 ECOTOX Database

FEMMtool's toxicity data come from the US EPA ECOTOXicology Knowledgebase (ECOTOX, US EPA 2026), a curated compilation of single-chemical toxicity test results for aquatic and terrestrial species drawn from the peer-reviewed literature. ECOTOX was chosen because it is public, large and regularly updated, and because it records the test-design details needed to determine whether a result is suitable for a chronic SSD: exposure duration, whether concentrations were measured, the type of control, the number of test concentrations and the water chemistry (Olker et al. 2022).

Bulk ECOTOX exports are loaded into the FEMMtool database by administrators. Of the many fields ECOTOX provides, FEMMtool uses those that identify the test and its subject (**CAS** number and chemical name; species scientific name and ECOTOX's taxonomy from kingdom to genus; life stage; habitat), describe the exposure (media type, exposure type, exposure duration, test location), report the result (endpoint such as EC50 or **NOEC**, the effect measured such as mortality or growth, the effect concentration as reported by the authors and its unit, whether concentrations were measured or nominal), and describe the test design and conditions (number of test concentrations, control type, chemical purity, pH, salinity, temperature). To these FEMMtool adds fields of its own, described in §3.2, that ECOTOX does not supply: the stressor a record belongs to, an accepted species name, climate zone and taxonomic completeness groups, and the outcome of each inclusion criterion.

ECOTOX has limitations that the rest of the pipeline is designed around. The bulk ECOTOX files that FEMMtool loads carry effect concentrations in the units used by the original authors, so FEMMtool standardises units itself (§3.2.4). Species names are recorded as the original authors reported them, so a single species can appear under several names and FEMMtool resolves them to one accepted name (§3.2.2). And the species tested are dominated by a small set of standard laboratory organisms, which the completeness scoring in §6.1 addresses by increasing the AF when required taxonomic groups are missing.

3.2 FEMMtool database requirements

Not every ECOTOX record for a stressor is suitable for an SSD. This section describes how records are linked to stressors, how species are identified, and the criteria a record must meet to be eligible.

3.2.1 Stressor-CAS mapping

Each stressor is linked to a curated set of CAS numbers under which relevant tests appear in ECOTOX. The mapping excludes compounds whose dissolution would introduce a

DRAFT

confounding stressor (for example organic ligands, dithiocarbamate fungicides, permanganate or fluorides) and entries that do not identify a discrete test substance. The nickel stressor, for example, draws on the metal (7440-02-0) and its chloride, sulfate, nitrate and acetate salts, but not nickel citrate, benzoate or tungstate. Chromium is handled as two stressors, Cr(III) and Cr(VI), because their toxicity differs. The CAS set used for each stressor is listed in the report so that the data pull can be reproduced against ECOTOX. The list of CAS identifiers (as reported in ECOTOX) that are currently available for inclusion in FEMMtool assessments can be found below.

Table 3.2a. *CAS numbers used to retrieve aquatic toxicity records from ECOTOX for each v1.0 stressor: soluble inorganic forms of each metal used in standard aquatic tests, plus the elemental entry. Biocide and fungicide formulations, chelates, cyanide complexes, organometallic compounds, insoluble particulates and salts with toxic counter-ions are not included. Arsenic and iron pool their oxidation states; chromium is split into Cr(III) and Cr(VI).*

Stressor	CAS	Chemical
Aluminum	**10043-01-3**	**Aluminum sulfate (alum)**
	7446-70-0	**Aluminum chloride**
	7429-90-5	**Aluminum (elemental)**
	13473-90-0	**Aluminum nitrate**
	10043-67-1	**Aluminum potassium sulfate (potash alum)**
	28037-84-5	**Aluminum potassium sulfate (3:1:3)**
Arsenic	**7784-46-5**	**Sodium arsenite**
	7778-43-0	**Disodium arsenate**
	1327-53-3	**Arsenic trioxide**
	13464-37-4	**Sodium arsenite (1:3)**
	7631-89-2	**Sodium arsenate (unspecified)**
	7440-38-2	**Arsenic (elemental)**
	7778-39-4	**Arsenic acid**
	10103-60-3	**Monosodium arsenate**
	1303-28-2	**Arsenic pentoxide**
	13464-38-5	**Trisodium arsenate**
	69162-99-8	**Sodium arsenate (2:1)**
	15584-04-0	**Arsenate ion**
	15120-17-9	**Sodium arsenate**
	17428-41-0	**Arsenic(5+) ion**
	22541-54-4	**Arsenic(III)**
	7784-41-0	**Potassium arsenate**
	15502-74-6	**Arsenite**
	13464-58-9	**Arsenous acid**

DRAFT

	76080-77-8	Potassium arsenate
Cadmium	10108-64-2	Cadmium chloride
	10124-36-4	Cadmium sulfate
	7440-43-9	Cadmium (elemental)
	10325-94-7	Cadmium nitrate
	543-90-8	Cadmium acetate
Chromium(III)	13548-38-4	Chromium(III) nitrate
	10101-53-8	Chromium(III) sulfate
	10141-00-1	Chromium(III) potassium sulfate (chrome alum)
	1066-30-4	Chromium(III) acetate
Chromium(VI)	7778-50-9	Potassium dichromate
	7789-00-6	Potassium chromate
	7775-11-3	Sodium chromate
	10588-01-9	Sodium dichromate
	1333-82-0	Chromium trioxide
	12680-48-7	Sodium chromate (alternate CAS)
	7738-94-5	Chromic acid
	13530-68-2	Dichromic acid
	13907-47-6	Chromate
	13907-45-4	Chromate ion
Cobalt	7646-79-9	Cobalt chloride
	7440-48-4	Cobalt (elemental)
	10124-43-3	Cobalt sulfate
	10141-05-6	Cobalt nitrate
	71-48-7	Cobalt acetate
	544-18-3	Cobalt(II) formate
Copper	7758-98-7	Copper sulfate
	7447-39-4	Cupric chloride
	7440-50-8	Copper (elemental)
	1344-67-8	Copper chloride (generic)
	3251-23-8	Copper nitrate
	7758-99-8	Copper sulfate pentahydrate
	142-71-2	Cupric acetate
	4180-12-5	Copper acetate
	6046-93-1	Cupric acetate monohydrate
Iron	7439-89-6	Iron (elemental)
	7720-78-7	Ferrous sulfate

DRAFT

	7705-08-0	Ferric chloride
	12040-57-2	Iron chloride (generic)
	7758-94-3	Ferrous chloride
	10028-22-5	Ferric sulfate
	10124-49-9	Iron sulfate (generic)
	10421-48-4	Ferric nitrate
Lead	10099-74-8	Lead nitrate
	7758-95-4	Lead chloride
	7439-92-1	Lead (elemental)
	301-04-2	Lead acetate
	15347-57-6	Lead acetate (generic)
	12612-47-4	Lead chloride (unspecified)
Manganese	7773-01-5	Manganese chloride
	7785-87-7	Manganese(II) sulfate
	7439-96-5	Manganese (elemental)
	11132-78-8	Manganese chloride (generic)
	10124-55-7	Manganese sulfate (generic)
	10377-66-9	Manganese nitrate
Nickel	7718-54-9	Nickel chloride
	7786-81-4	Nickel sulfate
	7440-02-0	Nickel (elemental)
	37211-05-5	Nickel chloride (unspecified)
	13138-45-9	Nickel nitrate
	1250264-70-0	Nickel(I) nitrate
	373-02-4	Nickel acetate
	13931-83-4	Nickel chloride (“NiCl”)
Vanadium	13718-26-8	Sodium metavanadate
	1314-62-1	Vanadium pentoxide
	7440-62-2	Vanadium (elemental)
	13721-39-6	Sodium orthovanadate
	14333-18-7	Orthovanadate ion
Zinc	7733-02-0	Zinc sulfate
	7646-85-7	Zinc chloride
	7440-66-6	Zinc (elemental)
	7779-88-6	Zinc nitrate
	557-34-6	Zinc acetate
	7699-45-8	Zinc bromide

DRAFT

3.2.2 Species identification

An SSD requires one value per species, so species need to be identified consistently across records. ECOTOX records the scientific name as the original authors reported it, which means the same species can appear under a current name, an older synonym, or a subspecies or variety name (Olker et al. 2022). FEMMtool resolves each ECOTOX name to an accepted scientific name using an internal taxonomy reference table that covers roughly 2,300 ECOTOX names, resolving to about 2,200 accepted taxa, occurring in ECOTOX records for FEMMtool stressors. Synonyms are pooled under the accepted name, and subspecies and varieties are treated as their parent species, so that each species contributes one value to the SSD (§5). Taxonomic ranks (phylum to family) follow the NCBI Taxonomy database (Schoch et al. 2020), which is used as the single reference so that family and class comparisons are internally consistent; accepted names and synonyms were curated against it by hand, and each entry records its source.

The reference table also supplies information that ECOTOX does not, and that is needed to apply the taxonomic completeness criteria (§5): climate zone (temperate, tropical or both) and the taxonomic groups used by the US EPA and **CCME** frameworks (chordate, invertebrate or plant/algae; planktonic or benthic). It carries phylum, subphylum, class and family as well, and where ECOTOX's own taxonomy for a record is incomplete these are filled from the reference table so that a species is not counted as a taxonomic gap because a field was blank in the source. Taxa identified only above family level cannot satisfy criteria that depend on family (§5.2 to §5.5). Each row records how the accepted name was assigned (an exact match in the NCBI Taxonomy database, a genus-level match, or a manual assignment).

3.2.3 Inclusion criteria

Each record is evaluated against FEMMtool's inclusion criteria, which follow FEMM's recommendation that SSDs meet established data-quality criteria as closely as possible (Weighman et al. 2026, Section 6.2) and adapt the data-quality requirements of US EPA (Stephan et al. 1985) and the CCME (2007) criteria for primary data to the information available in ECOTOX. A record must satisfy all active criteria to be eligible. The outcome for every criterion, and the reason for any exclusion, is recorded and appears in the report's data appendix.

For some criteria a blank counts as a failure, because the attribute is essential to interpreting the result; for others a blank passes and only a reported value outside the range fails (Table 3.2b).

Table 3.2b. Inclusion criteria. A record must meet every criterion.

Requirement	Accepted	If not reported
Accepted endpoint	NOEC, NOEL, LOEC, MATC, or ECx/LCx with x $\leq$ 50 (§4.2)	Fails

DRAFT

Requirement	Accepted	If not reported
Aquatic test	ECOTOX records the test organism's habitat as water	Fails
Accepted effect	Mortality, growth, reproduction or development; for plants and algae also population-level effects (e.g. cell density, biomass yield)	Fails
Laboratory test	Conducted in the laboratory	Fails
Control reported	A control treatment is reported	Fails
Number of test concentrations, including the control	At least 4	Passes
Test pH (mean or range)	6.0 to 9.0	Passes

Whether exposure concentrations were analytically measured or nominal is recorded for each record and shown in the report's data appendix, but is not currently used as an inclusion criterion.Records whose result unit cannot be converted to an aqueous concentration (for example, body-weight or dietary units) are excluded; all others are converted to mg/L (§3.2.4). Criteria and thresholds are fixed in the software and cannot be changed by users.

3.2.4 Units

All concentrations are converted to mg/L (from µg/L, ng/L, ppm or ppb) before extrapolation and fitting, and records in any other unit are excluded (§3.2.3); and exposure durations are converted to hours before acute or chronic classification (§4.3).

DRAFT

4. Data extrapolation

FEMMtool SSDs are built from chronic, negligible-effect, same-medium data. For many metals such data are scarce, and FEMM v1.0 identifies data limitation as the main near-term obstacle to marine SSDs (Weighman et al. 2026, Sections 6.2 and 6.6). Many otherwise acceptable ECOTOX records fall short of the standard in one respect: a 96-hour LC50, an EC50 rather than an EC10, a freshwater test for a marine project. Rather than discard such records, FEMMtool standardises them by dividing by an **EF** and tracks the uncertainty this introduces (Weighman et al. 2026, Section 6.4). Conclusions drawn from extrapolated data are less certain than those drawn from data that meet the standard directly, so extrapolated data enter the SSD only where they are needed (§5) and the **total EF** is reported for every value. Extrapolation is automatic and cannot be switched off or altered by users.

4.1 Extrapolation factors

FEMMtool calls the factors applied to individual effect concentrations extrapolation factors (EFs). Other frameworks use terms such as "uncertainty factor", "safety factor" or "conversion factor" for the same purpose. FEMMtool uses "EF" so that these record-level factors are not confused with the AF, which is applied to the SSD-derived **HC5** (§6.1).

Each retained record's chronic-equivalent concentration is

```
Eq. (1)    C_adj = C_raw / (EF_percent-effect × ACR × EF_cross-medium)
```

where each factor is given in Table 4.1 and equals 1 when it does not apply. The denominator is the record's total EF, so unextrapolated records have a total EF of 1. Records with a total EF above 100 are excluded from the SSD; a product of exactly 100 (e.g. 10 × 10) is retained.

Table 4.1. Extrapolation factors. The three factors are multiplied to give a record's total EF (Eq. 1); records with a total EF above 100 are excluded.

Factor	Applies to	EF	Details
Percent effect	EC0–EC20, LC0–LC20, NOEC, NOEL	1.0	§4.2
	EC21–EC25, LC21–LC25	2.5	
	EC26–EC30, LC26–LC30	3.0	
	EC31–EC35, LC31–LC35	3.5	
	EC36–EC40, LC36–LC40	4.0	
	EC41–EC45, LC41–LC45	4.5	
	EC46–EC50, LC46–LC50	5.0	
	LOEC	2.5	
	MATC	2.0	
	ECx/LCx with x > 50; log-transformed endpoints	Excluded	

DRAFT

Factor	Applies to	EF	Details
Acute to chronic (ACR)	Records classified acute	10	§4.3
	Records classified chronic	1.0	
Cross-medium	Records from the other medium	10	§4.4
	Same-medium records	1.0	

The total EF has three roles downstream. It defines the preference order for data selection, with unextrapolated data first and then increasing total EF (§5). It determines whether a species counts as "extrapolated" in the completeness reporting (§6.1). And it is reported for every selected value, so that reviewers can see how much of the SSD rests on extrapolation.

4.2 Non-negligible percent effects

The SSD is intended to represent negligible-effect concentrations. Endpoints reporting a larger percent effect are divided by an EF that scales them toward that basis (Table 4.1).

Endpoint labels are normalised before lookup, so that variants such as "EC50*" or "EC50 (growth)" are read as EC50 and fractional values such as EC12.5 are read as the integer below. The scheme is adapted from the Australian and New Zealand guideline-derivation method (Warne et al. 2018, 2025, Sections 3.2.4 and 3.4.2.1), which treats **ECx** values with x $\leq$ 20 and NOECs as negligible-effect estimates needing no conversion and, when such data are insufficient, divides chronic EC50/LC50, LOEC and **MATC** values by conversion factors of 5, 2.5 and 2 respectively. FEMMtool uses the same three anchor factors. Rather than treating every endpoint between EC21 and EC50 as an EC50, it interpolates the factor in steps from 2.5 at EC21 to 25 up to 5 at EC46 to 50.

4.3 Acute-to-chronic extrapolation

Each record is classified acute or chronic from its exposure duration. A test is chronic when its duration is at or above the criterion for the organism's group, its life stage at the start of the test and, where the criterion distinguishes them, the type of effect measured. ECOTOX does not classify organisms into these groups, so FEMMtool's taxonomy reference table (§3.2.2) assigns every taxon to a group under each set of criteria, and the criteria are applied per record using the assigned group, the reported life stage and the endpoint's effect family.

4.3.1 Location-specific and default criteria

Where available, records are classified using the acute/chronic exposure durations prescribed by the regulatory framework associated with the project location; these are currently implemented for Canada (CCME 2007; Table 4.3b). Where location-specific criteria have not yet been implemented, FEMMtool uses a set of default criteria derived from the Australian and New Zealand guidelines (Warne et al. 2018, 2025; Table 4.3a). Further frameworks will be added as their criteria are implemented.

DRAFT

Where location-specific criteria exist, the tool uses them in preference to the default. The default criteria are used wherever the location-specific set is silent: for an organism group it does not address (CCME does not specify a chronic duration for vascular plants, reptiles, protozoa or other microorganisms), or for a taxon that cannot be matched to any of its groups. The lookup order for a Canadian project is therefore:

1. CCME criterion for the species, by life stage and effect type;
2. CCME criterion for the species' CCME organism group;
3. Default criterion for the species;
4. Default criterion for the species' default organism group.

Where the project location has no location-specific criteria, only the default criteria are used (steps 3 and 4). If no criterion is found at any step, the record's duration classification is recorded as unknown and the source database's own acute/chronic designation is used to decide whether the ACR applies. Each record carries the criterion applied (ccme, ccme_group, default or default_group) and a flag showing whether a Canadian record fell back to the default, so reviewers can see how much of a CCME dataset was classified on CCME's own terms.

Location-specific criteria do not vary by medium: CCME prescribes the same durations for freshwater and marine tests.

4.3.2 Default criteria

The default criteria are the chronic exposure durations of Table 1 of the Australian and New Zealand guideline method (Warne et al. 2025), applied by organism type and life stage at the start of the test; for early-life macroalgae the criterion also depends on the endpoint. "Juveniles" means post-larval juveniles and is grouped with adults. Two simplifications are made: all macroinvertebrate embryo tests use the 48-hour criterion that Warne et al. specify for development and metamorphosis endpoints, and a 24-hour microalgal or microbial test is classed as chronic where Warne et al. require more than 24 hours.

Table 4.3a. FEMMtool default chronic duration criteria

Organism type	**Life stage at start**	**Endpoint**	**Chronic if duration ≥**
Fish and amphibians	Adult, juvenile	All	21 days
Fish and amphibians	Embryo, larva	All	7 days
Macroinvertebrates	Adult, juvenile	All	14 days
Macroinvertebrates	Larva	All	7 days
Macroinvertebrates	Embryo	All	48 hours

DRAFT

Organism type	Life stage at start	Endpoint	Chronic if duration ≥
Macroinvertebrates	Gamete	All	1 hour
Microinvertebrates	Adult, juvenile, larva	All	7 days
Microinvertebrates	Embryo	All	48 hours
Microinvertebrates	Gamete	All	1 hour
Macrophytes	All	All	7 days
Macroalgae	Mature	All	7 days
Macroalgae	Early life stage	Lethality	7 days
Macroalgae	Early life stage	Development	48 hours
Macroalgae	Early life stage	Fertilization	1 hour
Microalgae	All	All	24 hours
Microorganisms	All	All	24 hours

4.3.3 CCME criteria

CCME (2007) sets chronic durations by organism group and, for invertebrates, by whether the measured effect is lethal or non-lethal. Invertebrates are divided into short-lived taxa (cladocerans, copepods, rotifers, insects, amphipods, isopods, mysids, polychaetes and oligochaetes, among others) and long-lived taxa (molluscs, decapod crustaceans, echinoderms, cnidarians, barnacles and other sessile forms). FEMMtool's taxonomy reference table assigns every taxon to one of these groups.

The durations are those given under "Exposure Period Definition" in CCME (2007, Part II, Section 1, pp. 7–8). CCME gives them in prose rather than as a table; two points required interpretation. CCME states that lethal endpoints from short-lived invertebrates in tests shorter than 21 days "will be considered on a case-by-case basis"; FEMMtool applies the 21-day criterion to all invertebrate lethal endpoints, which is the conservative reading. CCME gives no general duration for vascular plants, stating only that standard *Lemna* tests are long-term and other species are case-by-case; FEMMtool applies the default macrophyte criterion (7 days), which the standard 7-day *Lemna* test meets.

Table 4.3b. CCME chronic duration criteria (Canadian projects)

CCME organism group	Life stage at start	Effect	Chronic if duration ≥
Fish	Adult, juvenile	All	21 days
Fish	Embryo, larva	All	7 days

DRAFT

CCME organism group	Life stage at start	Effect	Chronic if duration ≥
Amphibians	Adult, juvenile	All	21 days
Amphibians	Embryo, larva	All	7 days
Short-lived invertebrates	All	Lethal	21 days
Short-lived invertebrates	All	Non-lethal	96 hours
Long-lived invertebrates	All	Lethal	21 days
Long-lived invertebrates	All	Non-lethal	7 days
Algae (micro- and macroalgae)	All	All	24 hours
Vascular plants	All	All	Not specified by CCME (case-by-case); default criterion applies (7 days)
Protozoa and other microorganisms	All	All	Not specified by CCME; default criterion applies (24 hours)
Reptiles	Adult, juvenile	All	Not specified by CCME; default criterion applies (21 days)

Mortality endpoints are read as lethal effects; growth, reproduction, development and fertilization endpoints as non-lethal effects. Where an endpoint's effect family is unknown, only criteria that do not distinguish effect type are used.

4.3.4 Unreported life stage

Where the life stage is not reported, criteria defined for "all" life stages are used if the group has one (invertebrates, algae, plants and microorganisms under CCME; macrophytes, microalgae and microorganisms under the default). For fish and amphibians, whose criteria differ by life stage, an unreported life stage is treated as adult or juvenile, so the 21-day criterion applies; this is the more protective reading. Macroalgae with no reported life stage are treated as mature.

4.3.5 Acute-to-chronic ratio

Records classified acute (§4.3.1 to §4.3.4) are divided by the acute-to-chronic ratio (ACR) of 10 (Table 4.1). Empirical ACRs, derived from paired acute and chronic tests on the same species and chemical, vary widely across taxa and chemicals because toxicokinetics and toxicodynamics differ between short and long exposures (Weighman et al. 2026, Section 6.4.1). Where no empirical ACR is available, a generic factor must be used. A default ACR of

DRAFT

10 is the value the Australian and New Zealand method prescribes in the absence of a chemical-specific ratio (Warne et al. 2018, 2025, Section 3.4.2.2, which traces the recommendation to US EPA and OECD guidance). The classification and the factor applied are recorded for each record.

4.4 Cross-medium extrapolation

Ideally a saltwater assessment would rely only on marine species. FEMM v1.0 (Weighman et al. 2026, Section 6.4.2) recognises that marine data are scarce and that saltwater assessments are routinely built from combined freshwater and marine data, citing ECHA (2008a) guidance that a factor of at least 10 applies when a PNEC is derived from a combined SSD, and a substantially higher factor where taxonomic diversity is limited. ECHA's reasoning (2008a, Section R.10.3.2) is that freshwater and marine fish, crustaceans and algae show no systematic difference in sensitivity and may be used interchangeably, but that the greater taxonomic diversity of marine communities implies a broader distribution of sensitivities and therefore additional uncertainty, which the factor covers. Warne et al. (2018, 2025) likewise permit combining fresh and marine data in one dataset under specific circumstances.

FEMMtool applies this at the record level. A project has an exposure medium (marine or freshwater), and records from the other medium are divided by a cross-medium EF of 10 (Table 4.1); same-medium records are unaffected. The taxonomic AF (§6.1) separately penalises limited diversity. Both media are searched by default (§8.3) so that marine SSDs, for which ECOTOX data are often sparse, can draw on freshwater data at the cost of a tenfold EF penalty. Users may select to restrict data selection to exclusively same-medium data as a supplementary assessment option, but these assessments can only be presented in a report in addition to a standard assessment, drawing from all available stressor-relevant data.

DRAFT

5. Data selection

An SSD requires exactly one value per species (accepted scientific name). FEMMtool selects that value, and decides which species to include, in two stages. In doing so it works toward a set of eight taxonomic requirements (§5.2 to §5.5). Where available, projects use taxonomic requirements derived from the regulatory framework associated with the project location (currently implemented only for Canada; CCME 2007). Where location-specific criteria have not yet been implemented, FEMMtool uses a set of default taxonomic criteria derived from the US EPA guidelines (Stephan et al. 1985). The CCME and default requirements specify different taxa, but FEMMtool applies the same standard of completeness to both: eight species are required for a complete dataset, and where a framework specifies fewer than eight taxa, the remainder are required to come from families not already represented (§5.4, §5.5). This keeps the AF consistent across locations while the taxa required remain those specified for the project's location. The procedure is fully automatic.

Stage 1: unextrapolated data (total EF = 1). Every species with at least one record at total EF = 1 is included, and within a species the value is chosen by endpoint preference (§5.1). Where a climate zone applies (temperate under the Canadian requirements; under the default requirements, the zone selected by the user, if any, §2.5), species from that zone are preferred, and unextrapolated data from the other zone are used only to fill criteria the preferred data leave unmet.

Stage 2: extrapolated data (total EF > 1), used only to fill taxonomic gaps. Two considerations pull in opposite directions here. Taxonomic breadth is essential to an SSD's representativeness, and extrapolated records are often the only way to reach it for a data-sparse metal. At the same time, every extrapolated value has been divided by an EF that is protective by design rather than an empirical estimate of the species' sensitivity, and each such value shifts the shape of the fitted distribution. FEMMtool therefore admits extrapolated data only where they add taxonomic coverage the requirements call for, so that the SSD gains breadth while the influence of these adjusted values on the distribution as a whole is kept as small as possible.

The tool checks whether the eight species requirements for AF = 1 that apply to the project's location and medium are met by the Stage 1 species. If not, it searches the extrapolated pool one requirement at a time, in order, because later requirements are defined relative to earlier ones, and takes the single best candidate for each unmet requirement. "Best" is ordered by total EF (lowest first), then climate zone (preferred zone first, where one applies), then endpoint preference. Each requirement has an ideal definition and, in most cases, a loosened definition used only when the ideal cannot be met. A species is added only if it fills an unmet requirement, selection stops when all requirements are met or the pool is exhausted, and a species is never replaced once selected. Records with a total EF above 100 are never eligible.

The eight species each set of requirements looks for are described below (§5.2 to §5.5). The same eight requirements drive both data selection here and the AF in §6.1, which gives the rules for turning fully and partially met requirements into an AF.

DRAFT

5.1 Endpoint preference

Within a species, the most protective allowed endpoint is preferred, in this order:

1. Percent-effect endpoints (ECx, LCx), lowest x first. At the same x, a non-lethal ECx is preferred over a lethal LCx, but a lower-x lethal value is preferred over a higher-x non-lethal value.
2. LOEC, then MATC.
3. NOEC, then NOEL, used only when no percent-effect concentration exists for the species.

This hierarchy follows the Australian and New Zealand method (Warne et al. 2018, 2025, Section 3.2.4). That method prefers directly estimated negligible-effect concentrations, accepts ECx ($x \leq 20$) and NOEC without conversion, uses LOEC, MATC and EC50 only after conversion, and prefers a sublethal ECx to an LCx unless a reliable LCx is lower. Preference applies within a total-EF band, so an unextrapolated NOEC is preferred to an extrapolated EC10. When a species has several equally preferred values (same total-EF band, same most-preferred endpoint) from different sources, the lowest value is used, which is the most protective choice. Endpoint preference is not user-adjustable.

5.2 Default (US EPA): Saltwater

US EPA guidance (Stephan et al. 1985) considers a saltwater dataset complete when it contains: two chordates from different families; at least one non-chordate that is not an arthropod; a mysid or penaeid shrimp; three further non-chordates, each from a family not already represented; and one additional species from any family not yet used. FEMMtool looks for these eight species in the order below. Each entry gives what the tool looks for first and, where the ideal cannot be found, what it will accept instead.

Table 5.2. Species sought during selection: Default, saltwater. Scoring: Table 6.1a.

#	Species sought	Ideal	Accepted if the ideal is unavailable
1	First chordate	Any chordate (fish, amphibian, or other vertebrate/chordate)	None
2	Second chordate	A chordate from a different family than the first	Any other chordate, even from the same family
3	Plant or alga	Any aquatic plant or alga. This fills the “non-chordate, non-arthropod” requirement	None

DRAFT

#	Species sought	Ideal	Accepted if the ideal is unavailable
4	Mysid or penaeid shrimp	A crustacean in the family Mysidae or Penaeidae	Any other crustacean
5	First additional invertebrate	An invertebrate from a family not used by the shrimp	Any other invertebrate
6	Second additional invertebrate	An invertebrate from a family not used by species 4 or 5	Any other invertebrate
7	Third additional invertebrate	An invertebrate from a family not used by species 4 to 6	Any other invertebrate
8	Unrepresented family	Any species (chordate, invertebrate or plant) from a family not used by species 1 to 7	None

Species identified only to a level above family cannot fill a requirement that depends on family.

5.3 Default (US EPA): Freshwater

For freshwater, US EPA calls for: a salmonid; a second bony fish from a different family (ideally a commercially or recreationally important warm-water species); a third chordate from another family; a planktonic crustacean such as a cladoceran or copepod; a benthic crustacean such as an amphipod, isopod or crayfish; an insect such as a mayfly, stonefly or midge; a non-chordate that is not an arthropod (typically a plant, algae or mollusc); and one further species from an unused family.

Table 5.3. Species sought during selection: Default, freshwater. Scoring: Table 6.1b.

#	Species sought	Ideal	Accepted if the ideal is unavailable
1	Salmonid	A bony fish in the family Salmonidae (trout, salmon, char)	Any other bony fish

DRAFT

#	Species sought	Ideal	Accepted if the ideal is unavailable
2	Second bony fish	A bony fish from a different family than species 1	Any other bony fish, even from the same family
3	Third chordate	Any chordate (a fish, amphibian or other) from a family not used by species 1 or 2	Any other chordate
4	Planktonic crustacean	A crustacean that lives in the water column (e.g. *Daphnia*, *Ceriodaphnia*, copepods)	Any other crustacean
5	Benthic crustacean	A crustacean that lives on or in the bottom (e.g. amphipods, isopods, crayfish)	Any other crustacean
6	Insect	Any aquatic insect	None
7	Plant or algae	Any aquatic plant or algae. This fills the "non-chordate, non-arthropod" requirement	None
8	Unrepresented family	Any species from a family not used by species 1 to 7	None

"Bony fish" is implemented as the class Actinopteri (ray-finned fishes).

5.4 Canada: Saltwater

The CCME protocol (2007) considers a marine dataset complete with: three fish species, at least one of which is temperate; two aquatic invertebrates from different classes, at least one of which is temperate; and one temperate marine vascular plant or alga. This specifies six taxa. FEMMtool therefore requires two further species from families not already represented, so that eight species are needed for a complete dataset, as under the default requirements.

Table 5.4. Species sought during selection: Canada, saltwater. Scoring: Table 6.1c.

DRAFT

#	Species sought	Ideal	Accepted if the ideal is unavailable
1	Temperate fish	A fish species whose native range is temperate	Any fish
2	Second fish	Any fish	None
3	Third fish	Any fish	None
4	Temperate invertebrate	An aquatic invertebrate whose native range is temperate	Any aquatic invertebrate
5	Second invertebrate	An aquatic invertebrate from a different class than species 4 (e.g. a mollusc if species 4 is a crustacean)	Any other aquatic invertebrate
6	Temperate plant or alga	A marine vascular plant or alga whose native range is temperate	Any marine plant or alga
7	Unrepresented family	Any species from a family not used by species 1 to 6	None
8	Second unrepresented family	Any species from a family not used by species 1 to 7	None

DRAFT

Climate zone comes from the taxonomy table (§3.2.2). Under the Canadian requirements, temperate species are preferred at every stage of selection. Tropical data are used only where temperate data are missing.

5.5 Canada: Freshwater

For freshwater, CCME calls for: three fish including at least one salmonid and one non-salmonid; three invertebrates including at least one planktonic crustacean (aquatic life stages of semi-aquatic insects count); and one freshwater vascular plant or algae. FEMMtool adds one further species from an unrepresented family to make eight.

Table 5.5. Species sought during selection: Canada, freshwater. Scoring: Table 6.1d.

#	Species sought	Ideal	Accepted if the ideal is unavailable
1	Salmonid	A fish in the family Salmonidae	Any fish
2	Non-salmonid fish	A fish from a different family than species 1	Any other fish
3	Third fish	Any fish	None
4	Planktonic crustacean	A crustacean that lives in the water column	Any other crustacean
5	Second invertebrate	Any aquatic invertebrate	None
6	Third invertebrate	Any aquatic invertebrate	None
7	Plant or algae	Any freshwater vascular plant or algae	None
8	Unrepresented family	Any species from a family not used by species 1 to 7	None

DRAFT

6. Risk assessment outputs

This section describes the quantities FEMMtool reports for each stressor and compartment, in the order they are derived. The statistical fitting that produces the HC5 and **HC10** is described in §7.

6.1 Assessment factor

The assessment-factor approach follows ECHA (2008a, Equation R.10-1), as in FEMM v1.0 (Weighman et al. 2026, Section 6.1): the HC5 is divided by an AF chosen to reflect the remaining uncertainty in the SSD, having regard to the quality of the database and the endpoints covered, the diversity and representativity of the taxonomic groups, and the statistical uncertainty around the 5th percentile. ECHA sets the range at 5 to 1 for an SSD that meets its minimum-data expectation and leaves the choice within that range to expert judgement.

FEMMtool adopts this logic but makes the AF a deterministic function of the dataset. Extrapolation uncertainty is handled at the record level by the EFs (§4) and statistical uncertainty by the confidence level (§6.4), so the AF itself accounts specifically for gaps in taxonomic coverage, scored against the eight species requirements that apply to the project's location and medium (§5.2 to §5.5). Where regulatory guidance would decline to derive a criterion from an incomplete dataset, FEMMtool applies a graduated AF instead, consistent with the principle (§1.2) that the tool should return a conservative answer rather than hand a judgement back to the user. The one exception is a dataset that lacks an entire trophic level, which the tool does not assess (§6.1.2).

A single scoring scheme is used for every location and medium. The taxa it looks for differ between the default and Canadian requirements, and between freshwater and saltwater, but the way those requirements are scored and converted to an AF is identical. The AF is computed on the final species set, including any extrapolated species added in Stage 2 of §5, and the report shows completeness both with and without the extrapolated species (§8.4).

6.1.1 Taxonomic completeness scoring

The eight species requirements are grouped into criteria, one per broad taxonomic group, because leaving an entire group unrepresented is a larger gap than missing one of several required species within a group. Each criterion receives a score between 0 (nothing found) and 1 (fully met). A species found in its ideal form earns full credit for its share of the criterion, a species accepted under the loosened alternative earns half of that share, and a species not found at all earns nothing. Each species is counted toward one criterion only; the criteria are evaluated in the order listed, and a species used by an earlier criterion is not available to a later one. The scores are averaged across the criteria and

Eq. (2) $AF = (1 / \text{average score})^5$

The result is rounded to two decimals and is never less than 1. A dataset that meets every requirement in full has an average score of 1 and receives AF = 1. The exponent of 5 applies to all locations and both media; it was chosen so that AFs for incomplete datasets remain

DRAFT

comparable with the deterministic factors of other frameworks (an average score of 0.63 gives AF ≈ 10; of 0.46, AF ≈ 50).

A species earns its full weight if found in the ideal form, half if only the alternative is found (Tables 6.1a–d). “Family”, “class” and the group labels refer to the columns of the taxonomy reference table (§3.2.2).

Table 6.1a. Completeness scoring: Default, saltwater. Selection: Table 5.2.

Criterion	Weight	Ideal (full weight)	Alternative (half weight)
Two chordates	½	Any chordate	—
	½	A chordate from a different family	A chordate from the same family
Non-chordate, non-arthropod*	1	Plant, alga or non-chordate, non-arthropod invertebrate	—
Mysid or penaeid	1	Mysidae or Penaeidae	Any other crustacean
Three more non-chordates	⅓ each	A non-chordate from a family not yet used	A non-chordate from a family already used
Unrepresented family	1	Any species from an unused family	—

* Usually filled by a plant or alga, since every SSD must include one (§6.1.2).

Table 6.1b. Completeness scoring: Default, freshwater. Selection: Table 5.3.

Criterion	Weight	Ideal (full weight)	Alternative (half weight)
Three chordates	⅓	Salmonid	Bony fish† of another family
	⅓	Bony fish from a different family	Bony fish from the same family
	⅓	Chordate from a family not yet used	Any other chordate
Three invertebrates	⅓	Planktonic crustacean	Any other crustacean
	⅓	Benthic crustacean	Any other crustacean
	⅓	Insect	—
Non-chordate, non-arthropod*	1	Plant, alga or non-chordate, non-arthropod invertebrate	—
Unrepresented family	1	Any species from an unused family	—

* Usually filled by a plant or alga, since every SSD must include one (§6.1.2).

DRAFT

† "Bony fish" is implemented as the class Actinopteri (ray-finned fishes).

Table 6.1c. Completeness scoring: Canada, saltwater. Selection: Table 5.4.

Criterion	Weigh t	Ideal (full weight)	Alternative (half weight)
Three fish	⅓	Any fish	—
	⅓	Any fish	—
	⅓	Temperate fish	Non-temperate fish
Two invertebrates	½	Any invertebrate	—
	½	Invertebrate from a different class (temperate preferred)	Invertebrate from the same class
Marine plant or alga	1	Temperate	Non-temperate
Two unrepresented families	1 each	Any species from an unused family	—

Table 6.1d. Completeness scoring: Canada, freshwater. Selection: Table 5.5.

Criterion	Weigh t	Ideal (full weight)	Alternative (half weight)
Three fish	⅓	Salmonid	Bony fish† of another family
	⅓	Fish from a different family	Fish from the same family
	⅓	Any fish	—
Three invertebrates	⅓	Planktonic crustacean	Any other crustacean
	⅓ each	Any invertebrate (×2)	—
Vascular plant or alga	1	Any	—
Unrepresented family	1	Any species from an unused family	—

† "Bony fish" is implemented as the class Actinopteri (ray-finned fishes).

Temperate preference under CCME is applied in data selection (§5.4) for all requirements, but enters the score only where stated above: the third fish and the plant or algae in saltwater. Invertebrate climate zone does not affect the score.

6.1.2 Datasets that cannot be assessed

An SSD is fitted only if the selected species include at least one primary producer (plant or algae), one invertebrate and one chordate. A dataset lacking any of these three trophic levels is not assessed and the tool reports why. This is the point at which FEMMtool

DRAFT

declines to answer rather than apply a large AF: an SSD with no producers or no vertebrates does not describe a community, and no factor on its 5th percentile makes it one.

Within that gate the AF is always calculable from Eq. (2), and there is no separate default or fallback factor. A species set with no usable taxonomy at all would receive AF = 50; in practice the trophic gate prevents this case from reaching the AF step. Every result records the completeness path used (US or CCME, saltwater or freshwater) and the score and met/unmet status of each criterion.

6.2 PNEC

The PNEC is

Eq. (3) PNEC = HC5 / AF

where HC5 is the 5th percentile of the model-averaged SSD (§7) and AF is the taxonomic completeness factor of §6.1. One PNEC is derived per stressor and shared by all compartments of the project. The bootstrap confidence limits on HC5 (§7.5) propagate directly, at both the 95% and 90% levels: PNEC_L95 = HC5_L95 / AF, PNEC_U95 = HC5_U95 / AF, and likewise for the 90% bounds.

6.3 Risk quotient and risk level

To grade the severity of a PNEC exceedance, FEMMtool also evaluates the 10th percentile of the same fitted SSD (§7) and divides it by the same AF:

Eq. (4) HC10 / AF

This is a severity threshold, not a no-effect concentration. HC10 is evaluated from the fitted SSD parameters as a point value and is not bootstrapped, so HC10 / AF carries no confidence bounds. Both values are stored with each result.

For each compartment, the **risk quotient (RQ)** is

```
Eq. (5)    RQ = PEC / PNEC
```

with bounds RQ_upper = PEC / PNEC_lower and RQ_lower = PEC / PNEC_upper. This is the PEC/PNEC ratio used for risk characterisation under REACH (ECHA 2008a) and in FEMM v1.0 (Weighman et al. 2026, Section 6.1). The risk level is then as given in Table 6.3, which also gives the confidence level for Low Risk results (§6.4).

Table 6.3. Risk level and confidence level.

Risk Level	Confidence Level	Where the PEC sits	What it means
Low	High	Below the PNEC's lower 95% bound (PEC < PNEC_L95)	Protection goal met with margin; holds even for a less favourable SSD. **Recommended target.**
Low	Moderate	Below the lower 90% bound, but above the lower	Met, but not robustly. A small reduction in PEC, or more

DRAFT

Risk Level	Confidence Level	Where the PEC sits	What it means
		95% bound (PNEC_L95 ≤ PEC < PNEC_L90)	species in the SSD, would likely clarify the risk level.
Low	Low	Below the PNEC, but above its lower 90% bound (PNEC_L90 ≤ PEC ≤ PNEC)	Met on the point estimate only; a different sample of species could put the PEC above the PNEC. Should not be relied on without more data or a lower PEC.
Low	Unknown	Below the PNEC; the SSD has no 90% bootstrap bounds	Low Risk stands, but its robustness cannot be assessed.
Moderate	—	Above the PNEC, at or below HC10 / AF	Not acceptable. Relatively minor redesign may reach Low Risk (e.g. lower dose, slower release, relocation of the point of input).
High	—	Above HC10 / AF	Not acceptable. Substantial redesign or advanced assessment is needed.

6.4 Confidence

6.4.1 Definition

Each Low Risk result carries a confidence level of High, Moderate, Low or Unknown. It answers one question: how robust is the Low Risk finding to sampling uncertainty in the SSD? The rating is derived from where the PEC sits relative to the lower bootstrap confidence bounds of the PNEC (§6.2, §7.5), and it is framed protectively: High Confidence means the PEC would still be below the PNEC if the SSD had been fitted to a less favourable sample of species. Moderate and High Risk results carry no confidence level; the exceedance stands on the point estimate, and the two levels are distinguished by whether the PEC also exceeds HC10 / AF (§6.3).

Notation: PNEC_L95 and PNEC_L90 are the lower 95% and 90% confidence bounds of the PNEC. The conditions for each confidence level are given in Table 6.3 (§6.3).

If the required bounds are missing or invalid, the rating is Unknown; this occurs when the 90% bootstrap bounds (§7.5) were not computed for the SSD, and the reason is recorded with the result. When ratings are combined across compartments or stressors, the weakest is reported, and any Unknown forces Unknown.

The rating reflects only sampling uncertainty in the SSD. It does not capture uncertainty in the PEC, in the EFs or in the AF, which the report addresses qualitatively.

DRAFT

6.4.2 Interpreting risk level and confidence together

FEMMtool recommends that project planners aim for a Low Risk, High Confidence result in every compartment. That combination means the PEC is below the PNEC by a margin that sampling uncertainty in the SSD does not erode. Any other combination indicates either that the protection goal is not met or that the data cannot show robustly that it is. A Low Risk level is necessary for an acceptable outcome but is not by itself sufficient: a Low Risk result with Low Confidence rests on the point estimate alone. Which combinations a regulator will accept, and what confidence they require, is at the regulator's discretion; Table 6.3 (§6.3) describes what each outcome says about the assessment.

6.4.3 Acceptable outcomes

Only a Low Risk level can support an acceptable outcome; Moderate and High are not acceptable at any confidence. In the FEMM pathway a Low Risk result at the highest-risk compartment is what supports proceeding at the screening tier (Weighman et al. 2026, Section 5 and Figure 7), and FEMMtool's recommendation is that it be a Low result with High Confidence. Moderate Risk is an indicator that acceptable risk is close: the HC10 is a percentile of the same SSD, so a PEC between the PNEC and HC10 / AF exceeds the protection goal by an amount the distribution still resolves well, and relatively minor redesign (lower dose, slower release, relocation of the point of input) may bring it to Low. Some jurisdictions permit a concentration protective of 90% of species in already highly disturbed ecosystems under defined conditions; the Australian and New Zealand guidelines, for example, allow the 90% or 80% species-protection value for highly disturbed systems depending on the extent of disturbance, prospects for improvement and stakeholder agreement (Warne et al. 2018, 2025; ANZG 2018). Whether any such provision applies is a matter for the regulator with jurisdiction over the project site. FEMMtool does not apply it; it reports Moderate Risk as unacceptable everywhere and leaves the regulatory judgement to the reader. Both Moderate and High are FEMM's trigger for project redesign or for advanced risk assessment (§1.1; Weighman et al. 2026, Figure 7), in which mechanistic models refine the PEC, the PNEC or both. FEMMtool reports the trigger; the advanced tier is outside its scope.

DRAFT

7. SSD fitting procedure

The choice of distribution is one of the most consequential decisions in SSD practice, and no single distribution is preferred across datasets (Newman et al. 2000). FEMMtool therefore fits a fixed set of candidate distributions and averages across them, following the `ssdtools` approach (Thorley and Schwarz 2018) that the Australian and New Zealand guidelines adopted in 2018 in place of single-distribution fitting (Warne et al. 2018; Fox et al. 2021). The candidate set includes a mixture distribution to accommodate bimodal data, which ECHA (2008a, Section R.10.3.1.3) had earlier noted as a common cause of poor fit when a mode of action is specific to some taxa. The fitting algorithm, candidate set and model-averaging threshold are administrative settings (§8.3) and are the same for all users and projects.

7.1 Minimum dataset

FEMM v1.0 (Weighman et al. 2026, Section 6.1) follows the ECHA (2008a, Section R.10.3.1.3) convention that confidence can be placed in an SSD-derived PNEC when the dataset contains at least ten species spanning eight taxonomic groups, and preferably more than fifteen species. In FEMMtool the eight groups are the taxonomic requirements that apply to the project's location (§5), and any shortfall against them is carried into the AF (§6.1). The ten-species minimum is applied separately: SSDs with fewer than ten species are flagged in the results and the report (§8.4) as not adequately supported for screening, though the tool will fit an SSD from as few as three species, and only when the species include at least one plant or algae, one invertebrate and one chordate (§6.1.2).

7.2 Candidate distributions and maximum likelihood

All concentrations are on the mg/L scale after extrapolation (§4). Six candidate distributions are fit by maximum likelihood; a candidate is dropped if the optimizer fails or returns non-finite parameters or likelihood.

Distribution	Parameters	CDF F(x)	Estimation
Log-normal	μ, σ	$\Phi((\ln x - \mu)/\sigma)$	Closed form; $\hat{\mu}$ and $\hat{\sigma}$ are the mean and standard deviation (n divisor) of ln x
Log-logistic (Fisk)	c, scale	1 / (1 + (x/s)^−c)	Numerical MLE, location fixed at 0
Gamma	shape, scale	Regularised lower incomplete gamma	Numerical MLE, location fixed at 0

DRAFT

Distribution	Parameters	CDF F(x)	Estimation
Log-Gumbel (inverse Weibull)	location, scale (log scale)	exp(−exp(−(ln x − a)/b))	Gumbel MLE on ln x, likelihood corrected for the change of variable
Weibull	shape, scale	$1 - \exp(-(x/\lambda)^k)$	Numerical MLE, location fixed at 0
Two-component log-normal mixture	meanlog_1, sdlog_1, meanlog_2, sdlog_2, pmix	pmix Φ_1 + (1 − pmix) Φ_2	Numerical MLE on an unconstrained reparameterisation; requires n ≥ 6; $\text{meanlog}_1 \leq \text{meanlog}_2$; pmix bounded away from 0 and 1 as in `ssdtools`

7.3 Model averaging

For each candidate with k parameters and maximised log-likelihood ℓ on n species,

$$\text{Eq. (6)} \quad AICc = 2k - 2\ell + 2k(k+1)/(n-k-1)$$

(Burnham and Anderson 2002). **AICc** is treated as infinite when $n - k - 1 \leq 0$, which excludes the mixture for $n \leq 6$. Candidates whose AICc exceeds the best by more than 7 are dropped, and the rest receive Akaike weights $w_i = \exp(-\Delta_i/2) / \Sigma \exp(-\Delta_i/2)$, where $\Delta_i = AICc_i - AICc_min$. The model-averaged SSD is the weighted mixture of the candidate CDFs,

$$\text{Eq. (7)} \quad G(x) = \Sigma_i \, w_i \, F_i(x)$$

and the HC5 for proportion p is the root of G(x) = p. Averaging on the CDF scale (the `ssdtools` "multi" method) keeps G a proper distribution and avoids the bias that arises from averaging quantiles across distributions of different shape **(Thorley and Schwarz 2018; Fox et al. 2021).** All components, weights and the Δ threshold are stored with the SSD so that any HC5 can be recomputed from stored parameters.

7.4 Goodness of fit and diagnostics

For each candidate, **Anderson–Darling**, **Cramér–von Mises** and **Kolmogorov–Smirnov** statistics are computed on the fitted-CDF-transformed data, with p-values obtained by parametric bootstrap (simulate n values from the fit, re-estimate, recompute; 200 replicates). ECHA (2008a) recommends the Anderson–Darling test alongside Kolmogorov–Smirnov because it weights the tails, where the HC5 is read. These tests have low power at typical SSD sample sizes, so the report also provides a P–P plot for each

DRAFT

candidate, with the lower quarter of the distribution shaded, and a leave-one-out sensitivity table giving the model-averaged HC5 with each species removed in turn. A large shift from removing a single species indicates an HC5 that rests heavily on one test.

7.5 Uncertainty

Sampling uncertainty in the HC5 is estimated by nonparametric bootstrap: species values are resampled with replacement, all candidates are refit and re-weighted, and the HC5 is recomputed (HC10 is evaluated from the point fit, §6.3). The production interval uses 1,000 replicates and a fixed seed for reproducibility. Equal-tailed percentile intervals are reported at 95% (2.5th to 97.5th percentiles) and 90% (5th to 95th), and both feed the confidence level in §6.4. The bootstrap captures uncertainty about which species happened to be tested; it does not capture within-test measurement error or the structural uncertainty of the EFs in §4.

8. Software architecture

8.1 Components

femm-core is the scientific package. It decides which studies qualify, fits the SSD, and calculates the assessment factor, PNEC, risk quotient, and confidence. It is tested on its own, separately from the website.

- **Backend** stores the ECOTOX data, projects, assessments, and reports, and runs those calculations. An SSD is saved and reused when a later assessment uses the same stressor and the same data and fitting choices. The project report is a PDF.
- **Frontend** (femmtool.org) guides the user through a project, an exposure, the result, and the report. Sign-in is by email or Google.
- **Deployment**: the website and the database are hosted on Render.

8.2 Provenance and reproducibility

Each result stores what is needed to reproduce it: which toxicity records were used and why, the fitting choices, the assessment factor and PNEC, and the exposure compared with the PNEC. The same inputs produce the same result.

8.3 Report structure

The project report is a PDF. The cover names the project, location, regulatory salinity, and date, and states the screening disclaimer. An executive summary lists each selected run (compartment, stressor, PEC, PNEC, risk level, and confidence). Each run then gives the exposure, the risk quotient, the SSD and HC5, and the assessment factor. A mixture summary is included when one was run. The glossary follows. The appendices are:

1. Relation to regulatory guidance. Where location-specific taxonomic requirements are not used, the completeness path is the default path (§5).

DRAFT

2. Data available for the SSD: the candidate studies, whether each was included, and why.
3. Data selection methodology, including the chronic-duration criteria for the framework in use.
4. SSD statistical fitting: model weights, goodness of fit, parameter estimates, HC5, and the effect of removing one species.
5. Alternative assessments, only when a supplemental data pool (for example temperate-only or tropical-only) was included.

9. Future directions

FEMMtool v1.0 is a first release, and this manual describes only what the tool does today. The developments below are planned or under consideration. None affects the results of a v1.0 assessment, and this section is kept separate so that it can move to release notes or the user guide as items are delivered.

Additional stressors. The stressor list (§2.3) will grow as toxicity data for further stressors are compiled and validated. Further trace metals are the near-term priority; mercury and strontium are already staged in the catalogue. Carbonate-chemistry parameters relevant to ocean alkalinity enhancement, such as pH, total alkalinity and carbonate saturation state, are planned for a later version, since the toxicity data and test criteria for these differ from those for dissolved metals (Weighman et al. 2026, Section 3).

Location-specific criteria for more frameworks. Acute/chronic exposure durations drawn from US EPA guidance are in preparation (§4.3). Location-specific taxonomic requirements and durations for the European Union (ECHA 2008a) and Australia/New Zealand (Warne et al. 2018, 2025) will follow, so that projects there no longer rely on the default criteria (§1.2, §5). Support for locations outside the four jurisdictions currently accepted (§2.1) depends on identifying an applicable framework for each.

Mixtures. FEMM describes the assessment of multiple stressors through a concentration-addition Risk Index (Weighman et al. 2026, Section 6.5). FEMMtool contains a preliminary implementation of this approach that is not yet user-ready; it will be documented here once its outputs and their presentation have been settled.

Post-deployment monitoring. The four-day averaging window applied to MEC time series (§2.4) follows US EPA practice. Making the window a user-selectable setting would allow MECs to be checked against guidelines that use other averaging periods, such as those of CCME or of the Australian and New Zealand guidelines.

Literature cited

Aldenberg, T., and J. S. Jaworska. 2000. Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. *Ecotoxicology and Environmental Safety* 46:1–18.

DRAFT

ANZECC and ARMCANZ. 2000. *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*. Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra. †

ANZG (Australian and New Zealand Governments). 2018. *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*. Canberra. https://www.waterquality.gov.au/anz-guidelines †

Burnham, K. P., and D. R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. Springer, New York.

CCME (Canadian Council of Ministers of the Environment). 2007. *A Protocol for the Derivation of Water Quality Guidelines for the Protection of Aquatic Life 2007*. In: *Canadian Environmental Quality Guidelines, 1999*. Canadian Council of Ministers of the Environment, Winnipeg. †

ECHA (European Chemicals Agency). 2008a. *Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.10: Characterisation of Dose [Concentration]-Response for Environment*. Helsinki. †

ECHA (European Chemicals Agency). 2008b. *Guidance on Information Requirements and Chemical Safety Assessment. Appendix R.7.13-2: Environmental Risk Assessment for Metals and Metal Compounds*. Helsinki. †

ECHA (European Chemicals Agency). 2016. *Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.16: Environmental Exposure Assessment*, version 3.0. Helsinki. doi:10.2823/515110.

Etterson, M. 2020. *Technical Manual: SSD Toolbox Version 1.0*. EPA/600/R-19/104. US Environmental Protection Agency, Duluth, MN. †

European Commission. 2018. *Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 27: Technical Guidance for Deriving Environmental Quality Standards*. Updated version 2018. European Commission, Brussels.

Fox, D. R., R. A. van Dam, R. Fisher, G. E. Batley, A. R. Tillmanns, J. Thorley, C. J. Schwarz, D. J. Spry, and K. McTavish. 2021. Recent developments in species sensitivity distribution modeling. *Environmental Toxicology and Chemistry* 40:293–308.

MEE (Ministry of Ecology and Environment of the People's Republic of China). 2022. *Technical Guideline for Deriving Water Quality Criteria for Freshwater Organisms* [淡水生物水质基准推导技术指南] (HJ 831-2022). Ministry of Ecology and Environment, Beijing.

Newman, M. C., et al. 2000. Applying species-sensitivity distributions in ecological risk assessment: assumptions of distribution type and sufficient numbers of species. *Environmental Toxicology and Chemistry* 19:508–515.

Olker, J. H., C. M. Elonen, A. Pilli, A. Anderson, B. Kinziger, S. Erickson, M. Skopinski, A. Pomplun, C. A. LaLone, C. L. Russom, and D. Hoff. 2022. The ECOTOXicology Knowledgebase: a curated database of ecologically relevant toxicity tests to support

DRAFT

environmental research and risk assessment. *Environmental Toxicology and Chemistry* 41:1520–1539. doi:10.1002/etc.5324.

Posthuma, L., G. W. Suter II, and T. P. Traas (eds). 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers (CRC Press), Boca Raton, FL.

Schoch, C. L., S. Ciufo, M. Domrachev, C. L. Hotton, S. Kannan, R. Khovanskaya, D. Leipe, R. McVeigh, K. O'Neill, B. Robbertse, S. Sharma, V. Soussov, J. P. Sullivan, L. Sun, S. Turner, and I. Karsch-Mizrachi. 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* 2020:baaa062.

Stephan, C. E., D. I. Mount, D. J. Hansen, J. H. Gentile, G. A. Chapman, and W. A. Brungs. 1985. *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*. PB85-227049. US Environmental Protection Agency, Office of Research and Development, Duluth, MN; Narragansett, RI; Corvallis, OR. †

Thorley, J., and C. Schwarz. 2018. ssdtools: An R package to fit species sensitivity distributions. *Journal of Open Source Software* 3(31):1082.

US EPA. 1992. *Guidelines for Exposure Assessment*. EPA/600/Z-92/001. Risk Assessment Forum, Washington, DC.

US EPA. 1998. *Guidelines for Ecological Risk Assessment*. EPA/630/R-95/002F. Risk Assessment Forum, Washington, DC.

US EPA. 2008. *White Paper: Aquatic Life Criteria for Contaminants of Emerging Concern. Part I: General Challenges and Recommendations*. OW/ORD Emerging Contaminants Workgroup, Washington, DC. †

US EPA. 2026. *ECOTOX Knowledgebase*. Office of Research and Development, Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division, Duluth, MN. https://cfpub.epa.gov/ecotox/**.** The release loaded into the FEMMtool database is identified in each report (§8.2).

Warne, M. St J., G. E. Batley, R. A. van Dam, J. C. Chapman, D. R. Fox, C. W. Hickey, and J. L. Stauber. 2018. *Revised Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants*. Australian and New Zealand Governments and Australian state and territory governments, Canberra. †

Warne, M. St J., G. E. Batley, R. A. van Dam, J. C. Chapman, D. R. Fox, C. W. Hickey, J. L. Stauber, and R. Fisher. 2025. *Method for Deriving Australian and New Zealand Water Quality Guideline Values for Protecting Aquatic Ecosystems from Toxicants – Update of 2018 Version*. Prepared for the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. Australian and New Zealand Governments and Australian state and territory governments, Canberra, ACT. †

Weighman, K., E. Jankowska, and M. G. Andrews. 2026. *Framework for Ecotoxicological Modeling of mCDR (FEMM) v1.0*. Hourglass Climate. doi:10.5281/zenodo.18487758.

DRAFT