Changes to the Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants

Michael St. J. Warne [1], Graeme E. Batley [2] and Rick van Dam[3]

1. Department of Science, Information Technology and Innovation (DSITI)
2. Commonwealth Scientific and Industrial Research Organisation (CSIRO)
3. Environmental Research Institute of the Supervising Scientist (ERISS)

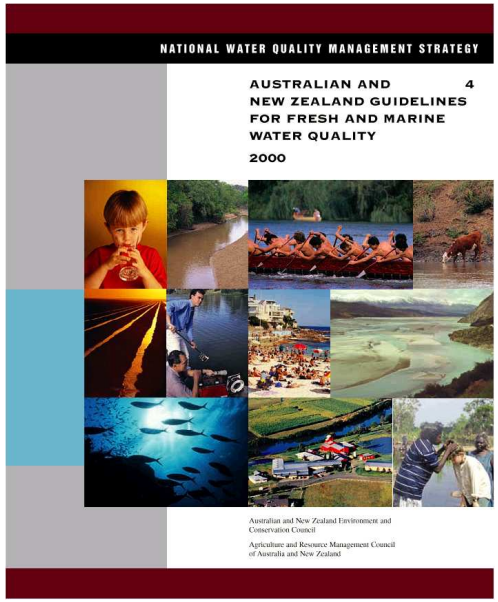
Managing Water Quality in Australia

National Water Quality Management Strategy (NWQMS)

Aim to achieve the sustainable use of the nation's water resources by protecting and enhancing their quality while maintaining economic and social development

Australian and New Zealand Guidelines for Fresh and Marine Water Quality (WQGs)

(1 of 25 documents)



WQGs aim to protect water resources for various uses (values), e.g., Human Health, Ecosystem Protection, Irrigation, Cultural

Risk-based protection (% of species) is provided by meeting Guideline Values (GVs) previously called Trigger Values (TVs)

Water Quality Guideline Value (GV)

- A numerical concentration limit or descriptive statement recommended to support a designated environmental value (use)

Science Based

- The best scientific estimate of concentrations below which there is a low probability of adverse environmental effects

Water Quality Objective

- Specific targets that become indicators of management performance

Stakeholder Based

- Derived through co-operative management involving all stakeholders
- Can consider non-scientific factors

Revision of the Water Quality Guidelines (2009 – 2016)

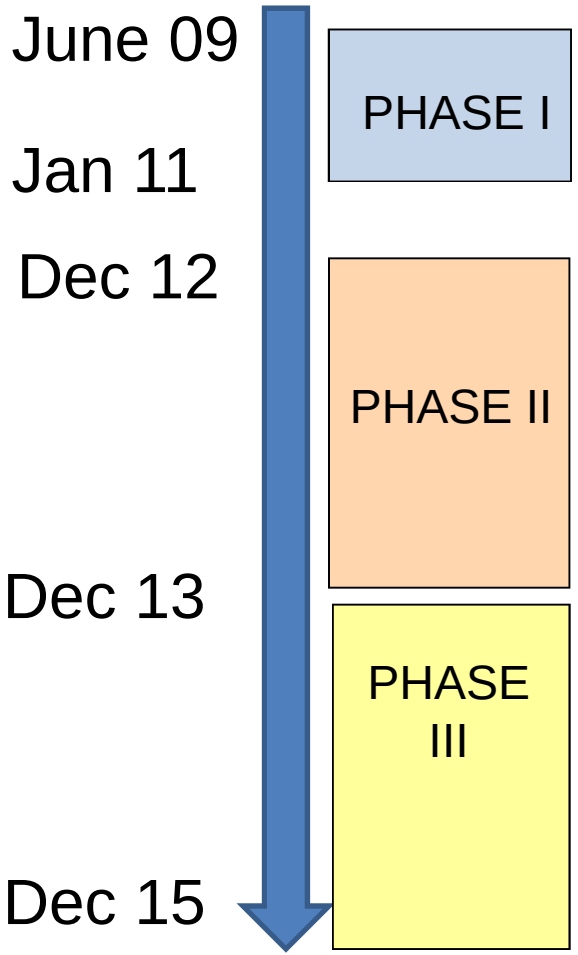
Why?

- It is a regulatory requirement to review and update the water quality guidelines
- There is the need to maintain scientific rigour and relevance to the user
- To correct mistakes and develop guidelines for new chemicals
- To improve the methodology for deriving TVs
- To increase the use of SSD approach and decrease use of the Assessment Factor method
- To use new data and advances in the science

What are being revised?

- Document 4 – Australian and New Zealand Guidelines for Fresh and Marine Water Quality
- Document 7 – Monitoring and Reporting

Revision of the WQGs



Develop scope of work
Work out timelines

Develop and test new version of BurrliOZ
Derive GVs for four priority chemicals
Finalise Sediment Quality Guidelines
Finalise list of chemical to derive GVs

New method for toxicant WQGs
Derive GVs for ≈ 35 selected chemicals

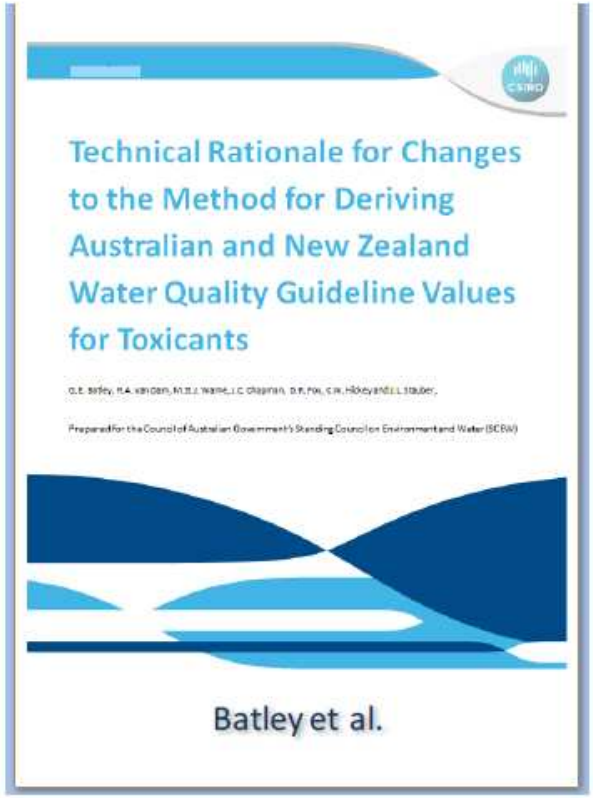
Key outputs for toxicants

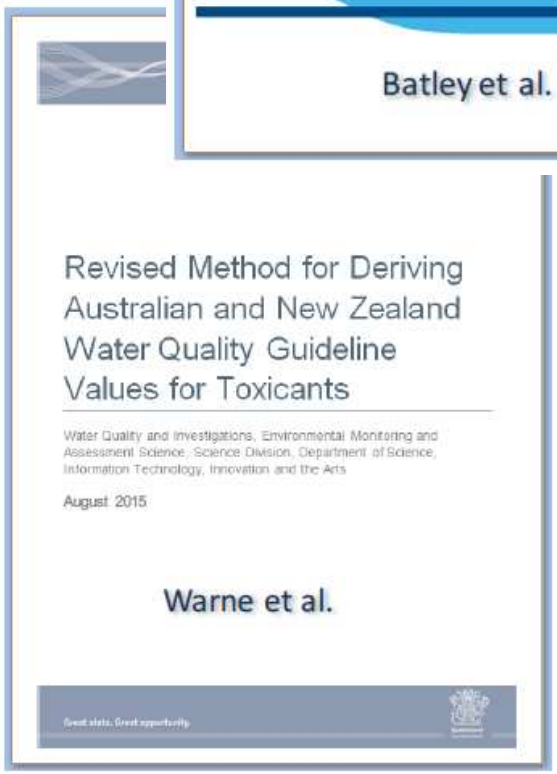
Sediment Quality Guidelines

- Simpson et al. 2013
  - Incorporation of weight of evidence

Water Quality Guidelines

- Batley et al. 2014
  - Technical rationale for changes to the method for GV derivation
- Warne et al. 2015
  - Actual method for calculating GVs – detailed step by step instructions
- New version of BurrliOZ





Guidelines

- The Australian and New Zealand Water Quality Guidelines are **Guide**lines. They provide **guidance**.
- Default is to use the GVs and methods provided
- Exceptions are permitted provided they are:
  - scientifically defensible
  - transparent and
  - pass critical review


CAUTION
THIS MACHINE
HAS NO BRAIN
USE YOUR OWN

Methods used to derive GVs

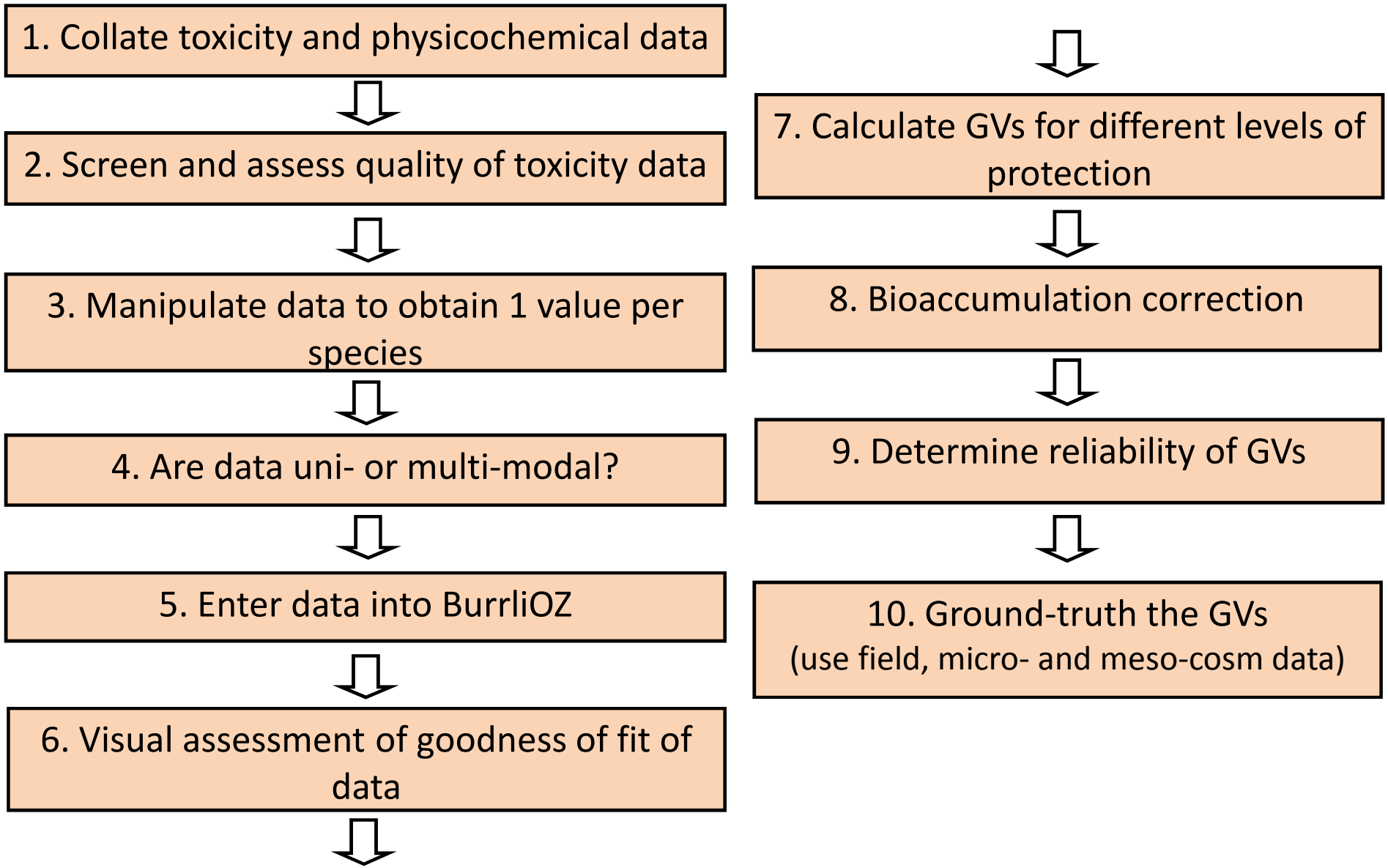
1. Species Sensitivity Distribution (SSD) method (preferred)
   - BurrliOZ V2.0 software*
   - Uses Burr Type III distributions and log-logistic

Types of data
1. Field, micro- and meso-cosm data
2. Laboratory-based data

2. Assessment Factor (AF) method
(only used when there are insufficient data for SSD method)

* Available from: https://research.csiro.au/software/burrlioz/

Overview of the method for deriving toxicant GVs using a SSD method



Acceptable sources of data

- Previously only published articles were used
- Therefore much data was excluded, e.g. consulting reports and commercial-in-confidence data
- Aim is to use all appropriate quality ecotoxicology data as SSDs are a statistical method
- We do not differentiate between standard tests (e.g. OECD methods), GLP, research or commercial ecotoxicology results
- Two steps
  - Develop method to include commercial-in-confidence data (e.g. NICNAS, APVMA data)
  - All grey literature data used provided copy publically available.

Acceptable endpoints

- Ecologically relevant endpoints (e.g., lethality, immobilisation, growth, population growth rate)
- Non-traditional endpoints (e.g., photosynthesis inhibition, in-vivo biochemical and physiological endpoints, behavioural endpoints, and genotoxicity and mutagenicity) **can** be used:
  - to derive GVs **provided their ecological relevance can be proved**
  - in a weight of evidence manner if ecological relevance not proved

Assessing data quality - I

- Based on Hobbs et al. (2005) and Heemsbergen et al. (2009)
- Covers fresh and marine water, sediment and soil
- Asks a series of questions about the experimental design, organism, statistics and chemistry
- We developed an Excel spreadsheet to
  - increase transparency
  - preserve the assessment

Hobbs et al. 2005. *Integr Environ Assess Manag*, 1, 174 -18
Heemsbergen et al. 2009. *Sci Tot Environ*, 407, 2546-2556

Table 1
Questionnaire for assessing the quality of terrestrial toxicity data adapted from Hobbs et al. (2005)

	Question	Credits
1	Was the duration of the exposure stated (e.g., 48 or 96 h)?	10 or 0
2	Was the biological endpoint (e.g., immobilization or population growth) stated and defined (10 marks)? Award 5 marks if the biological endpoint is only stated	10, 5 or 0
3	Was the biological effect stated (e.g., LC or NOEC)?	5 or 0
4	Was the biological effect quantified (e.g., 50% effect, 25% effect)? The effect for NOEC and LOEC data must be quantified	5 or 0
5	Were appropriate controls (e.g., a no-toxicant control and/or solvent control) used?	5 or 0
6	Was each control and contaminant concentration at least duplicated?	5 or 0
7	Were test acceptability criteria stated (e.g., mortality in controls must not exceed a certain percentage)? OR Were test acceptability criteria inferred (e.g., test method used [USEPA, OECD, ASTM etc]) (award 2 marks). Note: Invalid data must not be included	5, 2 or 0
8	Were the characteristics of the test organism (e.g., length, mass, age) stated?	5 or 0
9	Was the type of test media used stated?	5 or 0
10	Were the contaminant concentrations measured?	4 or 0
11	Were parallel reference contaminant toxicity tests conducted?	4 or 0
12	Was there a concentration-response relationship either observable or stated?	4 or 0
13	Was an appropriate statistical method or model used to determine the toxicity?	4 or 0
14	For NOEC/LOEC data was the significance level 0.05 or less? OR For LC/EC/BEC data was an estimate of variability provided?	4 or 0
15	Were the following parameters measured?	
	pH,	3, 1 or 0
	OM or OC content	3, 1 or 0
	Clay content	3, 1 or 0
	CEC	3, 1 or 0
16a	Was the temperature measured and stated?	3 or 0
17	Was the grade or purity of the test chemical stated? For metal salts 3 points are automatically scored.	3 or 0
18	Were other cations and or major soil elements measured? Or Were elements known to affect bioavailability measured? (e.g. Mo for Cu and Cl for Cd)?	3 or 0
19	For soils spiked with metal salts: Were the soils leached after spiking?	3 or 0
20	Was the incubation conditions and duration stated	3, 1 or 0
	Total score	
	Total possible score for the various types of data and contaminants: 102	
	Quality score ([Total score / 102] * 100)	
	Quality class (H _80%, 51–79% A, U _ 50%)	

The quality of a study will be considered unacceptable to derive soil and amended soil trigger values if the quality score is <50%.

Assessing data quality - II

- Quality Score = $\frac{\text{total score}}{\text{total possible score}} \times 100$
- Three quality classes

Quality Score	Quality Classification	Acceptable to drive GVs
>80	High	Yes
>50 to ≤80	Moderate	
<50%	Unacceptable	No

- Tests with nominal concentrations cannot be high quality and usually end up being unacceptable

Assessing data quality - III

- This is quite different to the Klimisch et al. (1997) data quality assessment method
- The numerical quality score helps with subsequent weight of evidence type approaches

New definitions of acute and chronic

- **Acute toxicity**: A lethal or adverse sub-lethal effect over a short exposure period relative to the organism's life span.

- **Chronic toxicity**: An adverse lethal or adverse sub-lethal effect over a substantial portion of the organism's life span (usually a long-term exposure) or a sub-lethal adverse effect on an early life stage.

Examples of acute and chronic toxicity tests

	Toxicity Test	Life Stage	Duration
Acute	Fish & Amphibs	Adults/Juv'les	<21 days
		Larvae/eggs	<7 days
	Invertebrates		<7 days
	Plants		<7 days
	Micro-organisms		≤24 h
Chronic	Fish & Amphibs	Adults/Juv'les	≥21 days
		Larvae/eggs	≥7 days
	Invertebrates	Adults/Juv'les	≥14 or 7 days or 3 broods
		Larvae (dev)	≥48 h
		Embryos (fert)	≥1h
	Plants		≥7d
	Macroalgae	Early life stages	≥1h
	Micro-organisms		>24 h

Data requirements and adequacy

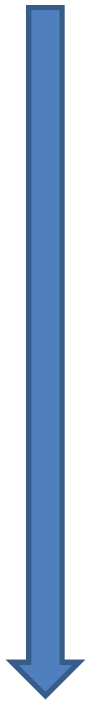
Sample size	Adequacy of sample size
≥15 species for ≥4 taxa	Preferred
8 – 14 species for ≥4 taxa	Good
5 – 7 species for ≥4 taxa	Adequate
5 species for ≥4 taxa	Minimum*
<5 species	Low to very low

* aim is to increase the adequacy of sample size over time

No requirement for data for any specific type of organism – as it violates a key assumption of SSD methods

Order of Preference for Data Use

Highest — Chronic



Lowest — Acute

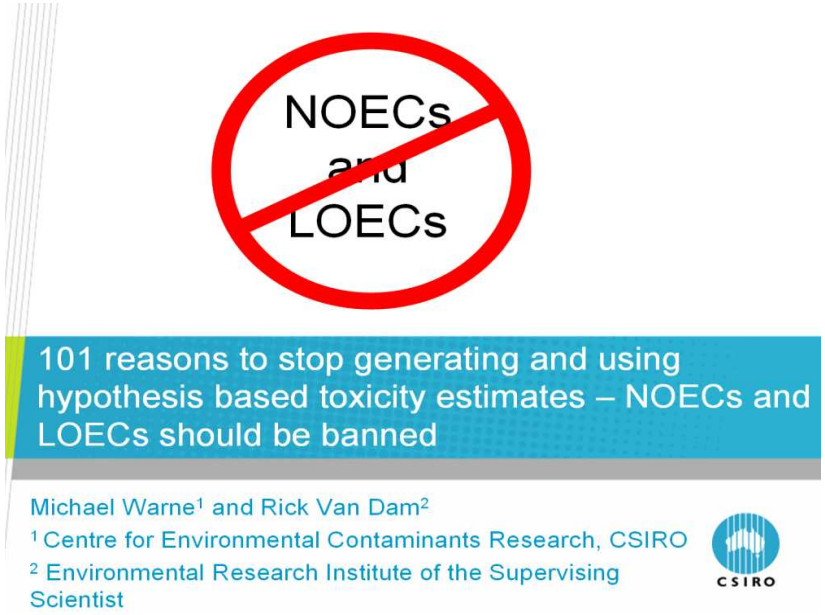
1. NEC
2. EC10
3. BEC10
4. EC15 – 20
5. NOEC
6. est. values from EC\LC\IC50, LOEC or MATC

NOEC data

NOEC data have been extensively criticised

- have a misleading name
- are inappropriate for the intended use
- are an invalid use of statistical methods

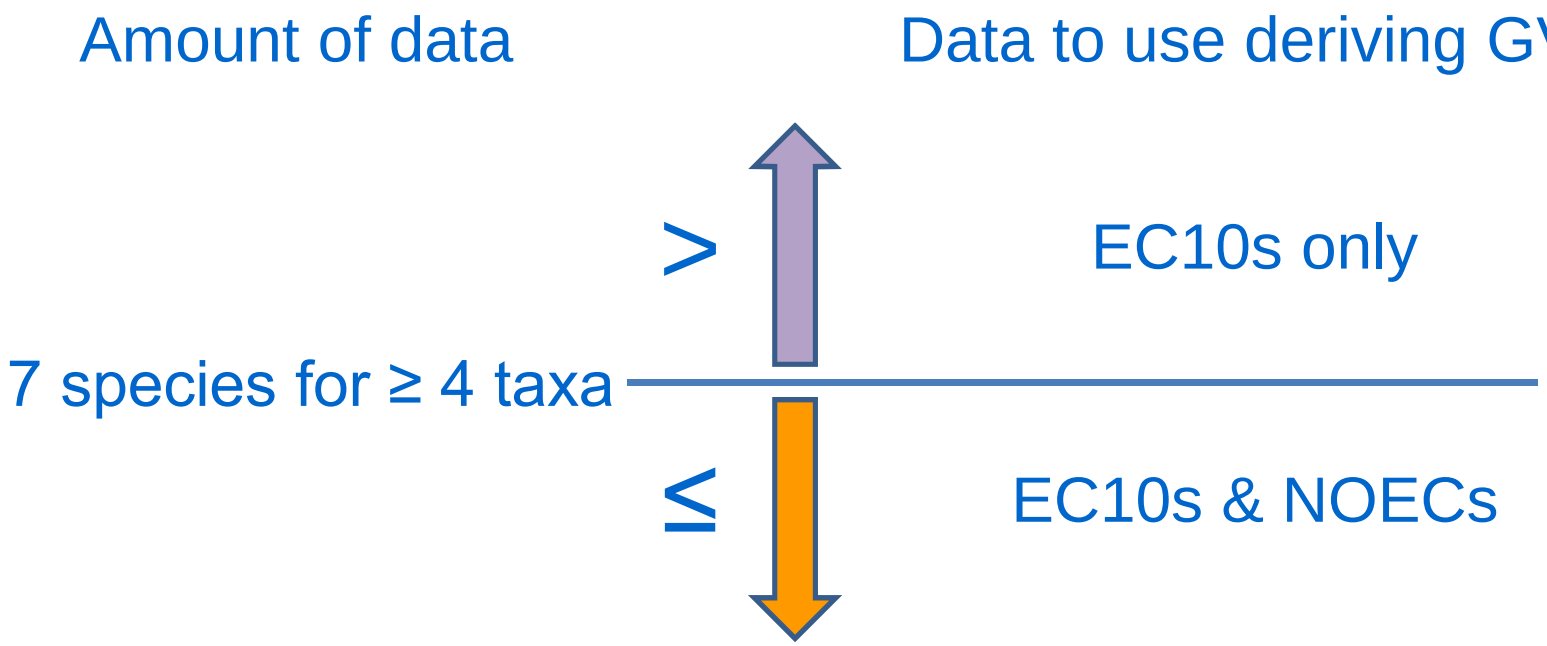
But they form the majority of existing chronic toxicity data



e.g. Hoekstra & Van Ewijk. 1993. *ET&C* 12, 187-194; Chapman et al. 1996. *ET&C 15, 77-79;* Fox 2008. *Aus J Ecotox 14(1), 6-9;* Newman. 2008. *ET& Chemistry 27,* 1013-1019; Warne & Van Dam. 2008. *Aus J Ecotox 14(1), 1-5.*

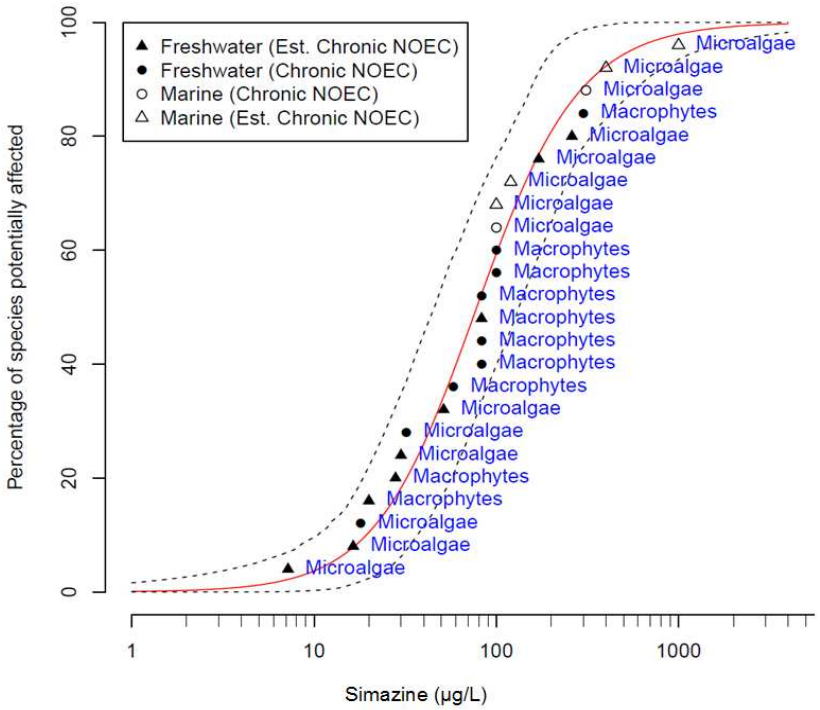
Dealing with NOEC data

- We are phasing out the use of NOECs



Improvements to BurrliOZ

Aim: to increase the calculation transparency and understanding of precision



Three main changes

1. Two distributions fitted to data
   - log-logistic when n = 5 - 7 and
   - Burr type III when n ≥ 8
2. Symbols for different types of organisms and data (e.g., chronic, est. chronic, acute)
3. Calculate the 95% CLs associated with any PC value

Guideline values

Type of ecosystem	Guideline Value (degree of protection)
High conservation value – minimal modification	PC99 (HC1)
Slightly to moderately modified	PC95 (HC5)
Highly modified	PC90 or PC80 (HC10 or HC20)

GVs are the outputs of the SSD. They are not then divided by an AF.

Reliability of GVs

Previously

- High reliability
- Moderate reliability
- Low reliability
  - Interim
  - Environmental concern level (ECL)

Based on

(1) the no. of species & taxa

(2) the type of data

Currently

- Very high reliability
- High reliability
- Moderate reliability
- Low reliability
- Very low reliability

Based on

(1) the no. of species & taxa

(2) the type of data

(3) fit of model to data

SAMPLE SIZE[a]	DATA TYPE	ADEQUACY OF		RELIABILITY
		SAMPLE SIZE	SSD FIT	
≥15	Chronic	Preferred	Good	Very high
			Poor	Moderate
8–14	Chronic	Good	Good	High
			Poor	Moderate
5–7	Chronic	Adequate	Good	Moderate
			Poor	Low
≥15	Chronic & c. acute	Preferred	Good	Moderate
			Poor	Low
8–14	Chronic & c. acute	Good	Good	Moderate
			Poor	Low
5–7	Chronic & c. acute	Adequate	Good	Moderate
			Poor	Low
≥15	Converted acute	Preferred	Good	Moderate
			Poor	Low
8–14	Converted acute	Good	Good	Moderate
			Poor	Low
5–7	Converted acute	Adequate	Good	Low
			Poor	Very low

Options for GV derivation

- Derivation of site-specific GVs is encouraged
- GVs (including default) can and are being derived by non-government scientists – could include industry
- But all proposed default GVs must be rigorously assessed before being endorsed

Acknowledgements

- USEPA for the invitation to attend
- Members of Technical Working Group 4 for WQG Revision
- Partners and families of scientists involved

Thank you

Any questions?