Cefic-LRI ECO51 – SWiFT: Strengthening weight of evidence for FET data to replace acute fish toxicity

WP3: Development and evaluation of the BN model for replacing AFT

Presentation for ECHA 01.03.2022 Jannicke Moe (NIVA)

Outline

- 1. Weight of Evidence our approach
- 2. Data sources
- 3. BN model
- 4. Model calibration and evaluation
- 5. Demonstration of web user interface

Weight of Evidence - our approach

Weight of evidence frameworks

Integrated Environmental Assessment and Management — Volume 9999, Namber 9999—pp. 1–7

Health & Ecological Risk Assessment

A Weight of Evidence Framework for Enviror

Assessments: Inferring Qualities

Glenn Suter, *† Susan Cormier, † and Mace Barron;

Strett Further, -1 SUBATI Surmann, -1 and Innex Exercion) ISE Eminemental Protection Agency, Office of Research and Development, National Foruminant, Phate 10 Environmental Hotection Agency, unite or Antana and Development, National Environmental Postection Agency, Office of Research and Development, National Research Laboratory, Guil Ecology Bination, Guil Berear, Florida

Weightd

The weighing of heterogeneous evidence such as convention The weighing of haterogeneous evidence such as conventional absersary community surveys is cleance by the environmental assessments. Evidence synthe community surveys is elaborate to environmental assessments, evidence symmi-causes of observed effects, harards posed by chemicals or other agents, the chiefs or undervice ements, narance power of summarian or under agems, we environmental qualities. As part of its guidelines for weight of evidence envisivenenai quarristi vis part et its guddienes ter weight en evennen Environmental Protection Agency has developed a generally applicable trani evidence, weight the evidence, and weigh the body of evidence. Use of the fr invalues, indiget the invalues, and enough the barry or invalues, where there is high or Wolf practices and provide greater transparancy than a diverse and e Assess Manag 2017;0.000-000. Published 2017. This article is a US Govern

Keywords: Weight of evidence Rick assessment. Systematic review Ecole

INTRODUCITION Inferences in environmental assessment often involve and quantit interences in environmental assessment often involve multiple and hieterogeneous pieces of evidence. For examjudgment. ple, inferring the cause of an observed biological impairment evidence, ple, intering the cickle of an observed oxyongrice imperiment entering, could involve evidence derived from conventional laboratory heterogen could involve evidence derived iron conventional aboratory memorgen tests, ambient media tests, biomarkers, biological surveys, ments an tests, amount mena tests, biomanens, biological surveys, ments an chemical analyses, and models. Such inferences require evidence chemical anagoes, and modes. Such inversions require examination weighing the evidence. Although the weighing is often done whereas weighing the evidence. Attriough the weighing is often done weighes by instructured narratives or narratives guided by a flat of are alter or unanucanes nanatures or nanatures gueses by a list of lieu ane and considerations, an explicit weight-of-evidence (WoE) process convert considerations, an explicit weight-on-evidence (would a can increase the defensibility of results (Weed 2005). an increase the detensionity of results (Weed 2005), addition The US Environmental Protection Agency (USEPA 2016) quantit The US Environmental Protection Agency (USETA 2016) quantum has developed and adopted WoE guidelines for ecological of evid has developed and adopted work guidemite for ecological or ever assessments. The approach is potentially applicable to human quarter estementes. Interapproach potentisary approaches tomanan quantit health and welfare assessments, but ecological assessors and p Next and wenter assessments, but ecosystal assessment and p have, ingeneral, been more accepting of the formal weighing quart news, in general, usern more accepting or the formal weighing quart of multiple types of evidence. The framework presented here qualt or mumper system existence, the earnitework presented time quality on be used to integrate multiple pieces of evidence and to limits une used to integrate meaning people or evenesion or in our infer qualities such as causation, teratogenicity, or impairment. inter qualities such as causation, teratogenicity, or impairment. However, WoE may also be employed to derive quantities nowever, wore may also be employed to derive quantities woot such as benchmark concentrations or half-lives of chemicals et a (Suter et al. this issue). Address correspondence to sater-glern@opa.gov Published 14 June 2017 on wileyorlinelibrary.com/jou DOI: 10.1002/

Integr Environ Assess Manag 2017:1-7

Integrated Environmental Assessment and Management --- Volume 9999, Number 9999---pp. 1-7 Received: 29 March 2017 Returned for Revision: 1 May 2017 Accepted: 23 May 2017

Health & Ecological Risk Assessment

A Weight of Evidence Framework for Environmental Assessments: Inferring Ouantities

Glenn Suter, *† Susan Cormier, † and Mace Barront

†US Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, Ohio

US Environmental Protection Agency, Office of Research and Development, National Health and Environmental, Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, Florida

ABSTRACT

The US Environmental Protection Agency (USEPA) has developed a generally applicable framework for a weight-of-evidence (WoE) process for deriving quantitative values from multiple estimates. These guidelines are intended for environmental assessments that require the generation of quantitative parameters such as degradation rates or that develop quantitative products such as criterion values or magnitudes of effects. The basic steps are to weigh evidence for the environmental quality to be quantified, generate the value by merging estimates or by identifying the best estimate, and weight the results to determine confidence in the numerical value. When multiple data sets or outputs of multiple models are available, it may be appropriate to weigh the evidence. Use of the framework to weigh multiple estimates may increase the accuracy of quantitative results compared to a single estimate from a default method. Its use can provide greater transparency compared to ad hor weighing of evidence. Integr Environ Assess Manag 2017;X:000-000. Published 2017. This article is a US Government work and is in the public domain in the USA.

Keywords: Weight of evidence Risk assessment Water quality criteria Parameter estimation Meta analysis

INTRODUCTION

Environmental assessments typically involve the derivation numerical values apply to some quality that is being of quantities such as levels of effects or numerical criteria, quantified, Before deriving quantities, we must determine Often those quantities could be estimated by multiple data whether the system actually has a quality to be quantified analysis methods applied to multiple data sets from multiple (e.g., impairment or causation) or alternative qualities that experimental or observational approaches. None of the could be quantified (e.g., assessment endpoints or hazards) multiple estimates is the unknown true value, but each using the method described in the companion paper (Suter provides evidence concerning the true value. Hence, it is et al. this issue). Note that even if a quality is assessed appropriate to weigh each piece of evidence (estimate) to probabilistically (e.g., the probability of impairment), it is still determine the best-supported value of the effect level, a quality. The quality may be determined a priori (e.g., a threshold concentration, or whatever value is being derived standard assessment endpoint). If it is not, it may be derived (LISEPA 2016)

body of information to derive quantities by weighing. Having defined the quality, the appropriate quantity is qualitative and quantitative evidence. Note this is not a derived.

health assessments.

Table 1 presents examples to illustrate how qualities and quantities are associated in environmental assessments. THE FRAMEWORK

* Address correspondence to suter.glenn@eps.gov Published 14 June 2017 on wileyonlinelibrary Integr Environ Assess Manag 2017:1-7

When we estimate quantities such as cleanup levels, those by WoE, because there are likely to be multiple pieces of We present a formal framework that uses the richness of the evidence concerning the quality that must be considered.

guantification of the weight-of-evidence (WoE) process to This framework for inferring guantities by WoE, like other derive qualities (e.g., Linkov et al. 2009); rather it is the US Environmental Protection Agency (USEPA) frameworks, is derivation by WoE of a numerical output that is an actual intended to be sufficiently general and flexible to apply to quantity. This approach is illustrated by examples from the range of assessment problems addressed by the various ecological assessments, but it is equally applicable to human offices of the USEPA and other organizations that may find it useful.

The framework for WoE to derive a quantity has 3 basic steps (Figure 1). First, the quality to be quantified is determined. This step uses the WoE framework for inferring a quality (Suter et al. this issue). The other steps are distinct

DOI: 10.1002/jeam.1953 Published 2017 SETAC



EPA/100/R-16/001 December 2016 www.epa.gov/osa

Weight of Evidence in Ecological Assessment



Bayesian inference



The posterior probability is proportional to the product of the likelihood function and the prior probability

Your prediction is a compromise between your new evidence and your assumptions

Bayesian networks basics



Conditional probability table (CPT)

Toxicity to	1														
Substance gr.	LHS	LHU	MLN	MLS	MLU	MMN	MMS	MMU	MHN	MHS	MHU	HLN	HLS	HLU	HM
-inf9	0.0114	0.009	0.0059	0.009	0.0082	0.0024	0.0034	0.0055	0.0048	0.0067	0.0096	0.0326	0.0392	0.0463	0.02
-96	0.2518	0.2628	0.2048	0.2356	0.2476	0.1547	0.1823	0.1839	0.1923	0.218	0.229	0.3792	0.4259	0.4268	0.32
-65	0.2109	0.2141	0.194	0.2124	0.2055	0.1771	0.1792	0.1909	0.1934	0.2043	0.2091	0.2119	0.2134	0.2135	0.20
-54	0.2206	0.219	0.2281	0.2242	0.223	0.2189	0.2274	0.2293	0.225	0.2224	0.2182	0.1852	0.1652	0.1669	0.19
< -43	0.1624	0.1569	0.1904	0.1699	0.1683	0.2079	0.2008	0.1931	0.1861	0.1835	0.176	0.1116	0.0945	0.0908	0.13
j -3 - 0	0.1405	0.1358	0.1725	0.1465	0.1443	0.2306	0.2008	0.1913	0.1931	0.1615	0.1548	0.0777	0.0609	0.0551	0.10
् 0 - inf	0.0024	0.0024	0.0043	0.0024	0.0031	0.0084	0.0061	0.006	0.0053	0.0036	0.0033	0.0018	0.0009	0.0006	0.00
í	۲												111		
Jannicke Moe				01.03.2	022									6	

The model uses Bayes rule to combine **new evidence** (e.g. EC50 values) with **prior probabilities** (from the conditional probability table) to predict **posterior probabilities**

How can our Bayesian Network be used in a WoE approach?

EFSA JOURNAL

Scientific Opinion 🖞 Open Access 💿 💽 🕥

Guidance on the use of the weight of evidence approach in scientific assessments

EFSA Scientific Committee, Anthony Hardy, Diane Benford, Thorhallur Halldorsson, Michael John Jeger, Helle Katrine Knutsen, Simon More, Hanspeter Naegeli, Hubert Noteborn ... See all authors 🗸

First published:03 August 2017 | https://doi.org/10.2903/j.efsa.2017.4971 | Citations: 47

Our BN-WoE should be

- consistent with WoE approaches recommended for regulatory frameworks (EFSA, ECHA, US EPA, OECD)
- quantitative
- intuitive
- flexible

Main steps for WoE assessment (EFSA):



Data sources for BN development and evaluation

Data sources

TRAINING



AFT

AFT

TESTING

BN model structure

Schematic model structure



- Prior probabilities of toxicity to fish organised by substance group
- User: Enter information on substance
- 3 lines of evidence
- User: Enter toxicity values

- Posterior probability
- User: check warning (applicability domain)
- User: get model prediction

The full BN model: structure



The full BN model: probability distributions



«Warning node»

- Purpose: identify substances for which the BN predictions may not be reliable
- Reasoning: response of fish embryo may not representative of later fish stages



Current criteria to trigger a warning:

- Fish embryo toxicity = Low
- Daphnia/algae toxicity ratio = High
- Touch-evoke response = Yes / Unmeasured
- Biotransformation rate = Fast

Ongoing work:

- Refine Embryo toxicity criterion (< Daphnia & Algae toxicity)
- Include Metabolic activity (embryo vs. juvenile fish) as alternative variable?

Define the applicability domain of the BN

Critieria for exclusion from the applicability domain:

- 1) "Warning node": substances with certain combination of toxicity values and other properties
- 2) Substances without QSAR values: metals
- 3) ... [to be continued]
- Applicability domain is a scientific decision, not a technical issue
- Criteria must be defined in the Technical user manual (task 4.2) and the Guidance document for stakeholders (task 5.3)

How can our Bayesian Network be used in a WoE approach?





Model evaluation

Model evaluation: 6x4x4 criteria



Data selection criterion no. 2 (155 substances)

Model evaluation: example

Data selection criterion no. 2 (155 substances)

Evaluation criterion no. 2: LC50 < 10 mg/L



Predicted vs. observed toxicity level (difference)

Predicted toxicity = observed toxicity

➔ accurate prediction

Predicted toxicity < observed toxicity less protective

Predicted toxicity > observed toxicity → more protective

Web-based user interface (WP4, Anders Madsen - HUGIN)

Web-based user interface architecture



https://swift.hugin.com/models/extended3/