SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 32, No. 8, pp. 1918–1927, 2013

Printed in the USA

DEVELOPMENT OF SHORT, ACUTE EXPOSURE HAZARD ESTIMATES: A TOOL FOR ASSESSING THE EFFECTS OF CHEMICAL SPILLS IN AQUATIC ENVIRONMENTS

ADRIANA C. BEJARANO*† and JAMES K. FARR‡
†Research Planning, Columbia, South Carolina, USA
‡National Oceanic and Atmospheric Administration, Seattle, Washington, USA

(Submitted 11 January 2013; Returned for Revision 15 March 2013; Accepted 16 April 2013)

Abstract: Management decisions aimed at protecting aquatic resources following accidental chemical spills into rivers and coastal estuaries require estimates of toxic thresholds derived from realistic spill conditions: acute pulse exposures of short duration (h), information which often is unavailable. Most existing toxicity data (median lethal concentration or median effective concentration) come from tests performed under constant exposure concentrations and exposure durations in the 24-h to 96-h range, conditions not typical of most chemical spills. Short-exposure hazard concentration estimates were derived for selected chemicals using empirical toxicity data. Chemical-specific 5th percentile hazard concentrations (HC5) of species sensitivity distributions (SSD) from individual exposure durations (6–96 h) were derived via bootstrap resampling and were plotted against their original exposure durations to estimate HC5s and 95% confidence intervals (CIs) at shorter exposures (1, 2, and 4 h). This approach allowed the development of short-exposure HC5s for 12 chemicals. Model verification showed agreement between observed and estimated short-exposure HC5s (r^2 adjusted = 0.95, $p < 0.0001$), and comparison of estimated short-exposure HC5s with empirical toxicity data indicated generally conservative hazard estimates. This approach, applied to 2 real spill incidents, indicated hazard estimates above expected environmental concentrations (acrylonitrile), and suggested that environmental concentrations likely exceeded short-exposure hazard estimates (furfural). Although estimates generated through this approach were likely overprotective, these were derived from environmentally realistic exposure durations, providing risk-assessors with a tool to manage field decisions. *Environ Toxicol Chem* 2013;32:1918–1927. © 2013 SETAC

Keywords: Species sensitivity distributions Bootstrap Short-exposure duration Chemical spill

INTRODUCTION

Several hundred potentially toxic chemicals are transported in navigable waters of the United States at volumes high enough that they may pose significant risks to aquatic environments in the event of accidental spills. In fact, the US Coast Guard's National Response Center reports approximately 7000 accidental releases of chemicals into US waters every year, most of which involve oil and oil-related chemicals. Although most spills do not exceed 1000 L, a few incidents have involved much larger releases, typically related to accidents with ships, railcars, tankers, and barges carrying large amounts of hazardous materials. While most spills of chemicals go unnoticed in peer-reviewed literature, accidental spills have caused negative consequences to aquatic resources. Examples include the 1991 spill of metam sodium in the Sacramento River, California, USA [1], and the 1993 spill of a complex mixture of aromatic concentrates in the Nemadji River, Wisconsin, USA [2]. Management decisions regarding recovery options following a chemical spill rely on scientific support and multiagency coordination [3], as well as on emergency planning and response tools [4]. Forecast and trajectory models [5,6] are often used to characterize aqueous concentrations at the spill site and downstream by incorporating specific information about the receiving water body (depth, velocity, and volume), combined with physicochemical parameters, transport characteristics of chemicals (partitioning, evaporation, dissolution), and chemical fate information (hydrolysis, biodegradation). Although these models provide initial estimates of the potential concentrations of concern, risk-assessors, scientists, and environmental managers also need quantitative measures of concentrations associated with toxicological effects. Acute toxicity data (median lethal concentrations [LC50s] and median effect concentrations [EC50s]) derived from standard laboratory exposure conditions (static tests, or flow-through with limited dilution) and durations (e.g., 24–96 h) provide valuable information that can be used to characterize potential effects on aquatic receptors; however, these conditions are not representative of some spill conditions. Typical exposures under spill conditions are of short duration (a few hours) with rapid dilution into the water column, particularly when chemicals are spilled in moving waters. This issue is neither new [7] nor unique to chemical spills, and it extends to the use of dispersants during oil spills [8] as well as to episodic spills of pesticides [9–11] and episodic release of common pollutants [12]. Furthermore, applying safety factors to acute toxicity data from longer exposure durations falls short of providing toxicity values for shorter exposures and can lead to overestimation of risks [13]. Therefore, immediate assessments of spills based on short exposures (a few hours) are critical, but acute toxicity data derived from toxicity testing at short-exposure durations (e.g., <8 h) are scarce for most chemicals.

The main objective of the present study is to propose the use of a novel methodology for estimating hazard concentrations for aquatic species under environmentally realistic spill-exposure durations (a few hours). The outcomes of this approach would aid risk-assessors and environmental managers in their immediate characterization of the potential risk to aquatic resources from spills, while providing information for management decisions. It is important to note that this approach is not intended to address long-term environmental impacts caused by

All Supplemental Data may be found in the online version of this article.
* Address correspondence to abejarano@researchplanning.com.
Published online 27 April 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.2255

spills but to provide first-tier acute toxicity information to responders.

METHODS

Chemicals of concern

Hazard concentration estimates for short-exposure durations (a few hours) were derived for selected chemicals with varying amounts of acute toxicity data. Selection of chemicals was based on the following criteria: hazardous materials commonly transported in the United States, high potential risk for spills in large quantities, potentially toxic to aquatic organisms, and involvement in spills.

Short-exposure hazard estimates

Queries for toxicity data for selected chemicals were performed in databases compiled by US governmental agencies: the Chemical Aquatic Fate and Effects (CAFÉ) database (National Oceanic and Atmospheric Administration Emergency Response Division [NOAA ERD], Seattle, WA, USA, unpublished data; Supplemental Data, Figure S1) and the US Environmental Protection Agency ECOTOX database [14]. Data used in these analyses included toxicity data (LC50) for aquatic taxa (fish, crustaceans, mollusks, amphibians, insects, invertebrates, and worms) from tests ranging from 1 h to 96 h in duration. Given the limited data for most of the chemicals selected, analyses also included data on effects concentrations reported as EC50s, results from various laboratory exposure regimens (e.g., static, flow-through), toxicity values reported as measured or nominal, and tests performed under fresh or seawater conditions. Because data were queried from databases that compiled information from numerous sources, all data used here were assumed to be independently collected; even within the same original data source, chemical-specific toxicity data are, by definition, independently collected from each exposure duration test.

The 5th percentile hazard concentration (HC5; based on LC50 and EC50 data) of species sensitivity distributions (SSDs) and its associated confidence interval were used as a measure of chemical hazard [15]. The HC5, derived from a variety of species with different relative chemical sensitivities, corresponds to the chemical concentration that is assumed to be protective of 95% of the species tested, or 5% of the cumulative SSD curve. In the present study the HC5 was used to demonstrate the applicability of this approach; however, a more protective value (e.g., 1st percentile hazard concentration [HC1]) is recommended when the protectiveness of the mean HC5 is in question. Although there are several shortcomings in the use and interpretation of SSDs [16], these curves are valuable in representing the variability in responses across several species.

For each chemical, exposure duration–specific SSDs were generated using the geometric mean of LC50 and EC50 toxicity values by species, requiring a minimum of 5 species per curve. When no systematic differences in sensitivity between taxa were found, SSDs were generated for the combined data set, avoiding potential taxa bias by excluding taxa not equally represented across the different exposure durations. In cases where there were statistically significant differences in sensitivity to a chemical between taxa (via generalized linear interactive modeling [17]), curves were generated using only data for the most sensitive taxa. For chemicals with limited measured toxicity values, estimates of acute toxicity were generated using interspecies correlation estimation (ICE) models [18–20], which use robust regression analyses to generate effect concentrations for 1 or several species based on the known acute toxicity for a surrogate species. The ICE models were used only for chemicals that had data available for the same species at different exposure durations, selecting the most sensitive species when such information was available for more than 1 species. To limit propagation of uncertainty, only ICE models meeting general rules of thumb for best model selection [20] were included in these analyses. These rules included ICE models with relatively low mean square errors (<0.22) and high cross-validation success rate (>90%), degrees of freedom ($df > 8$), and coefficient of determination ($r^2 > 0.6$) [20]. Furthermore, only the closest taxonomic distance (within the same genus and family) [20] were included, except when information for the most closely related species was absent. In those cases, greater taxonomic distances were considered (within the same order, class, and phylum).

Several criteria were used to minimize propagation of uncertainty through the analyses. Although 3 empirical family distributions (log-logistic, log-normal, and Weibull) were used to fit SSDs [21], the log-normal model generally produced smaller (more conservative) estimates at the lower tail of the curve and therefore was selected for these analyses. The probability density function of the log-normal distribution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma \mathrm{x}}\ \mathrm{e}\ \left(-\frac{[\ln(\mathrm{x}) - \mu]^2}{2\sigma^2}\right), \mathrm{X}(0, \infty)$$

with parameters μ (mean) and σ (standard deviation). All SSDs were tested for goodness of fit ($\alpha = 0.01$) using the Anderson–Darling and the Kolmogorov–Smirnov test statistics. The former is more sensitive at the tail ends [22–24], while the latter is robust at the middle of the distribution [23]. Curves failing either goodness-of-fit tests were evaluated for gross outliers via distribution-based outlier detection methods [25] under the assumption of a log-normal family distribution. Based on this outlier detection method, values identified as gross outliers are unlikely to be drawn from the same log-normal distribution as the remainder data points [25]. Gross outliers were considered for removal (see Supplemental Data, Figure S2) only if their removal improved the fit of the SSD. Curves used to estimate HC5s included those that passed both goodness-of-fit tests and curves passing the Kolmogorov–Smirnov test, but with insufficient data points ($n < 7$) to compute the Anderson–Darling test.

Resampling theory was used to derive the HC5s and associated 95% confidence intervals (95% CI) by bootstrapping individual SSDs (within the same exposure duration) 2000 times under the assumption of a log-normal family distribution (Figure 1A). For each chemical, independent HC5 values for short-exposure durations (e.g., 1, 2, and 4 h) were calculated by plotting the bootstrapped mean response HC5 and its 95% CI (derived from the individual SSDs) on a log-log scale versus the original exposure duration, followed by bootstrap [26] of all the possible regressions occurring between the given HC5 and 95% CI values across exposure durations (Figure 1B). The form of this regression is described by a simple power law function

$$\mathrm{HC5} = \beta_0 \times \text{exposure duration(h)}^{\beta 1}$$

where β_0 and β_1 are the intercept (at exposure duration = 0) and slope, respectively. In all cases, this model was constrained

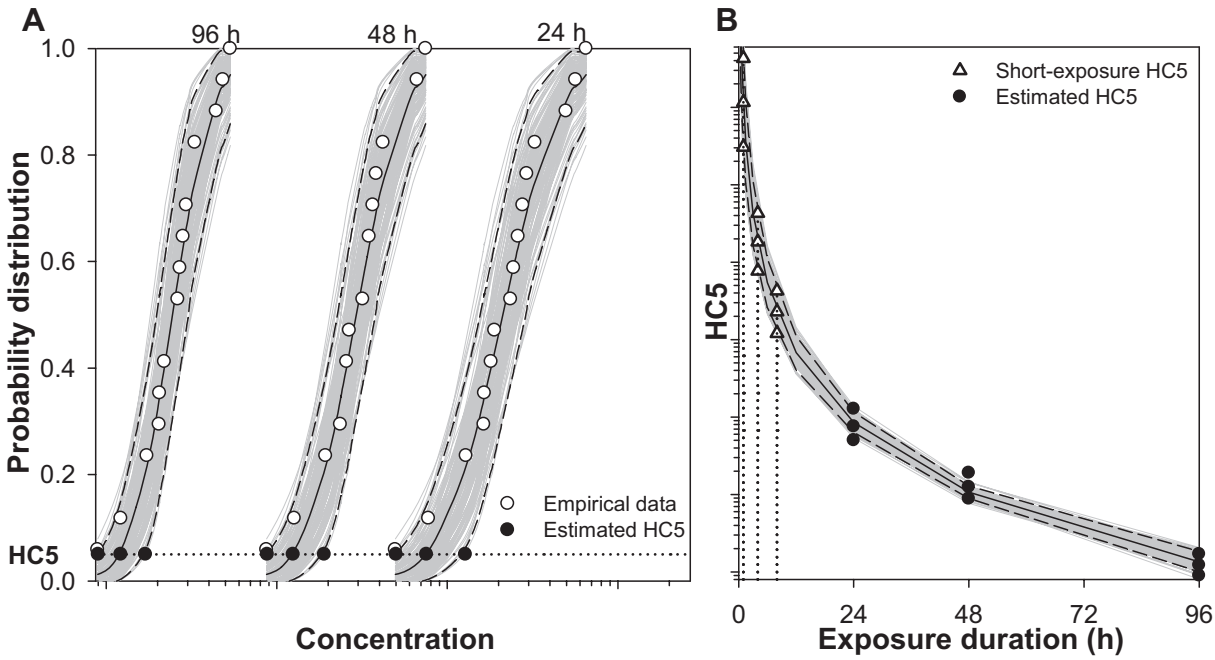

Figure 1. Schematic representation of the method used to derive 5th percentile hazard concentrations (HC5s) for short-exposure durations. (**A**) Empirical data (open circles) for several exposure durations (24, 48, 96 h) with their bootstrapped species sensitivity distributions (SSDs; gray lines), mean and 95% confidence interval (95% CI) concentrations (solid and dashed lines, respectively), and estimated HC5 (closed circles). (**B**) Estimated HC5 from each exposure duration (24, 48, 96 h; closed circles), the bootstrapped HC5s including mean and 95% CI concentrations (solid and dashed lines, respectively), and the estimated HC5s for short-exposure durations (1, 2, 4 h; triangles and dotted lines). Concentrations and HC5 scales are unitless for demonstration purposes.

arbitrarily to a 0.5-h minimum exposure duration. The selection of this model was not ad hoc, but it was based on a preliminary analysis that included the evaluation of other exposure duration–HC5 relationship models (3-parameter exponential decay or Haber's law, baseline toxicity, and 2 hyperbolic functions [27–29]; see Supplemental Data, Table S1). This analysis showed that the simple power law function generally had the smallest Akaike's Information Criterion value across several individual tests, indicating that this model provides the best fit. Short-exposure duration (1, 2, and 4 h) mean HC5s and 95% CIs were calculated using the simple power law function via bias-corrected accelerated bootstrap [26,30–32]. The use of bias-corrected accelerated bootstrap is recommended when estimating bootstrapped confidence intervals, as it is a robust method that adjusts for sample bias, and it does not require distribution assumptions [26]. For chemicals with data for at least 4 exposure durations that failed linear bootstrap (slope not statistically significant from 0 at $\alpha = 0.05$), the mean HC5 (but not its 95% CI) was used to estimate short-exposure duration HC5s. Only regressions that successfully passed linear bootstrap (slope statistically significant from 0 at $\alpha = 0.05$) were used in these analyses.

All data processing was recorded and considered when the reliability of hazard estimates were assessed (high, moderate, or low; see Supplemental Data, Table S2). Chemicals with highly reliable estimates included those with at least 3 different exposure duration SSDs each with at least 7 species, and curves that passed both goodness-of-fit tests ($\alpha = 0.01$). Moderately reliable estimates included those with at least 3 different exposure duration SSDs each with at least 5 species, curves with insufficient data points to compute the Anderson–Darling test, curves that violated the Anderson–Darling test ($\alpha = 0.01$), and curves with gross outliers that required outlier removal. Chemicals with low estimate reliability included those that had the same criteria as the moderately reliable estimates, and also included curves with data from ICE models, and curves where only the mean HC5, but not its 95% CI were used to estimate short-exposure duration HC5s.

Model verification

To quantitatively assess the accuracy of hazard estimates, available empirical data from short-exposure durations, not used in hazard estimate calculations, were compared with the estimated hazard values. Two approaches were used to verify hazard estimates. When empirical acute toxicity data for short-exposures (1–6 h) were available, HC5s and 95% CIs were estimated for these short-exposure SSDs, and values were compared with the HC5s estimated from longer exposure toxicity data using the methodology described above. However, this verification approach was limited to a few chemicals, and therefore an alternate approach was developed. Short-exposure (1, 2, and 4 h) HC5s estimated from longer-exposure toxicity data (6–96 h) were compared with empirical short-exposure toxicity data for individual species. This approach does not necessarily address the accuracy of hazard estimates, but rather it serves as a measure of their level of protectiveness.

RESULTS

Short-exposure hazard estimates

Short-exposure duration hazard concentration estimates (HC5s based on LC50 and EC50 data) for various chemicals were classified as having high, moderate, or low reliability (see Supplemental Data, Table S2). Highly reliable short-exposure HC5s derived from longer exposure SSDs were obtained for 5 chemicals, while estimates with moderate and low reliability were obtained for 4 and 5 chemicals, respectively (Table 1; see Supplemental Data for details). For example, at least 4 exposure durations SSDs were available for potassium permanganate

Table 1. Short-exposure duration (1, 2, and 4 h) bootstrapped 5th percentile hazard concentrations (HC5s) and 95% confidence intervals (95% CIs) for selected chemicals using empirical species sensitivity distributions (SSDs) available for longer exposure durations (6–96 h), and combining information from one or more taxa[a]

Chemical (CAS no.)	Empirical exposure duration SSDs (h)	Species taxa	Exposure duration (h)	Bootstrapped short-exposure estimates (μg/L)		Reliability
				Mean HC5	95% CI HC5	
Formalin (50000)	6, 24, 48, 96	C, F, M	1	684[b]	260–1652[b]	High
			2	392[b]	178–805[b]	
			4	225[b]	122–389[b]	
Parathion (56382)	24, 48, 72, 96	C, I	1	9	3–25	High
			2	7	3–16	
			4	5	2–10	
Malathion (121755)	24, 48, 72, 96	C, I, S	1	54	14–200	High
			2	33	11–99	
			4	20	8–48	
Zinc (7440666)	24, 48, 96	C	1	40[b]	1.5–1183[b]	High
			2	16[b]	1–262[b]	
			4	6[b]	0.7–58[b]	
Potassium permanganate (7722647)	6, 24, 48, 96	C, F	1	14[b]	7–29[b]	High
			2	9[b]	5–16[b]	
			4	6[b]	3.5–9[b]	
Furfural (98011)	24, 48, 96	C, F	1	319[b]	111–906[b]	Moderate
			2	189[b]	79–447[b]	
			4	112[b]	56–221[b]	
Acrylonitrile (107131)	24, 48, 96	A, C, F, S, M, W	1	138[b]	36–533[b]	Moderate
			2	87[b]	28–266[b]	
			4	55[b]	23–133[b]	
Zinc (7440666)	24, 48, 96	F	1	132[b]	5–3243[b]	Moderate
			2	54[b]	3–754[b]	
			4	22[b]	2–177[b]	
Chlorine (7782505)	12, 24, 48, 96	C, F, S	1	366	100–1300	Moderate
			2	227	80–665	
			4	141	62–323	
2-propanone (67641)	24, 48, 96	C, F	1	7374[b]	7100–7700[b]	Low
			2	7000[b]	6795–7261[b]	
			4	6660[b]	6501–6854[b]	
Trichloromethane (67663)	12, 24, 48, 96	F	1	33[b]	29–38[b]	Low
			2	30[b]	27–33[b]	
			4	28[b]	25–30[b]	
Styrene (100425)	24, 48, 72, 96	C, F	1	128[b]	30–555[b]	Low
			2	80[b]	24–271[b]	
			4	50[b]	19–132[b]	
2-propenal (107028)	24, 48, 72, 96	C, F, S	1	132	75–260	Low
			2	101	63–176	
			4	78	53–120	
Malathion (121755)	24, 48, 72, 96	F	1	136	96–181	Low
			2	118	89–149	
			4	102	82–122	

[a] All the toxicity data used to derive these estimates were obtained through data queries of CAFÉ and ECOTOX databases [14]. See Supplemental Data for details.
[b] Units in $\times 10^3$.
A = amphibians; C = crustaceans; F = fish; I = invertebrates; S = insects; M = mollusks; W = worms.

(high reliability) and chlorine (moderate reliability) (Figure 2), although the latter required more data processing (outlier identification via distribution-based outlier detection methods and gross outlier removal) for derivation of the HC5. By contrast, SSDs for trichloromethane (low reliability) used empirical data plus data generated through ICE models with *Oncorhynchus mykiss* as the surrogate species (taxonomic distance = 1 [20]), and statistically significant linear bootstrap was achieved by using only mean HC5 values, but not their bootstrapped 95% CIs.

Differences in the sensitivity across taxa were not found for 2-propenal and chlorine ($p > 0.05$) and were undetermined for a handful of chemicals with limited toxicity data or with little data for 1 or more taxa (e.g., potassium permanganate, furfural, trichloromethane, styrene). Consequently, and to err on the side of caution, most of the acute toxicity data used to estimate HC5s included data for the most sensitive taxa, which were generally fish and crustaceans, although other taxa were included for selected chemicals with limited empirical data. Chemicals with differences in sensitivity across taxa included: 1) formalin: crustaceans slightly more sensitive than fish ($p = 0.04$); 2) parathion: crustaceans and insects more sensitive than fish ($p = 0.005$); 3) malathion: crustaceans, insects, and invertebrates more sensitive than other taxa ($p = 0.01$); 4) acrylonitrile: crustaceans appear more sensitive than other taxa, but there were not sufficient data to assess this group independently; 5) zinc: crustaceans more sensitive than fish ($p = 0.02$); and 6) 2-propanone: fish and crustaceans slightly more sensitive than other taxa ($p = 0.04$). For zinc, short-exposure HC5 estimates for crustacean and fish were within 1 order of magnitude of each other, reflecting sensitivity differences between taxa; estimates for crustaceans were more reliable than those for fish, as the latter required outlier removal (Figure 3). Similarly, for malathion, estimates for the most sensitive taxa (crustacean, invertebrate,

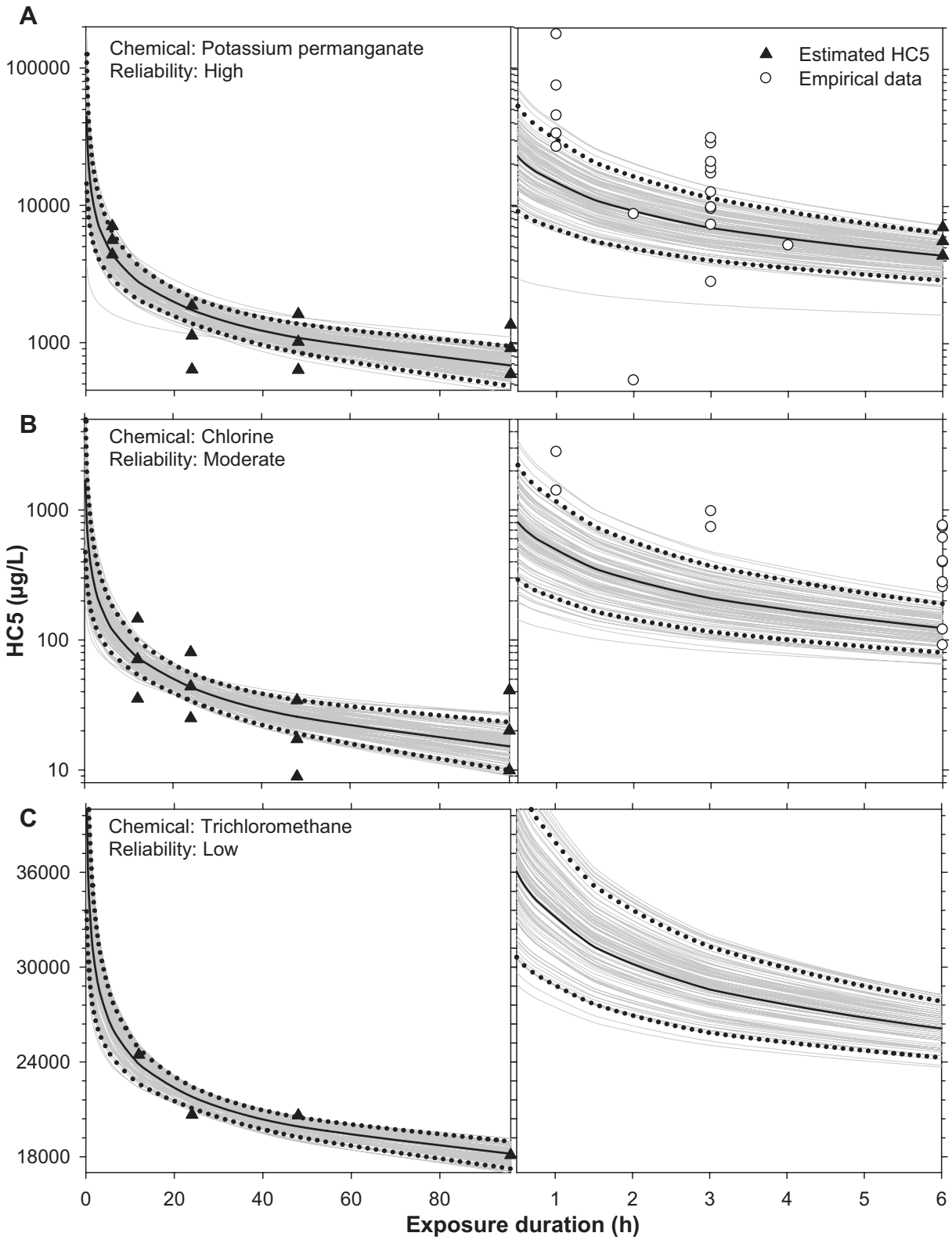

Figure 2. Bootstrapped 5th percentile hazard concentrations (HC5s) and 95% confidence intervals (95% CIs; triangles) used to estimate shorter exposure mean HC5s (solid line) and 95% CIs (dotted lines) via bootstrapping (gray lines) of a simple power law function. Estimates are shown for 3 chemicals with different degrees of reliability: (**A**) potassium permanganate, (**B**) chlorine, and (**C**) trichloromethane. Right-hand panels show a close-up of the shorter exposure estimates with all available chemical-specific empirical data (open circles).

and insect) were 1 order of magnitude less than those derived from fish data; estimates for the most sensitive taxa were more reliable than those of fish, as the latter required outlier removal and only the mean HC5 was used to estimate short-exposure duration HC5s. These observations indicate that potential differences in species sensitivities across taxa need to be considered when one is assessing hazards under spill conditions. That is particularly the case for pesticides, which, given their

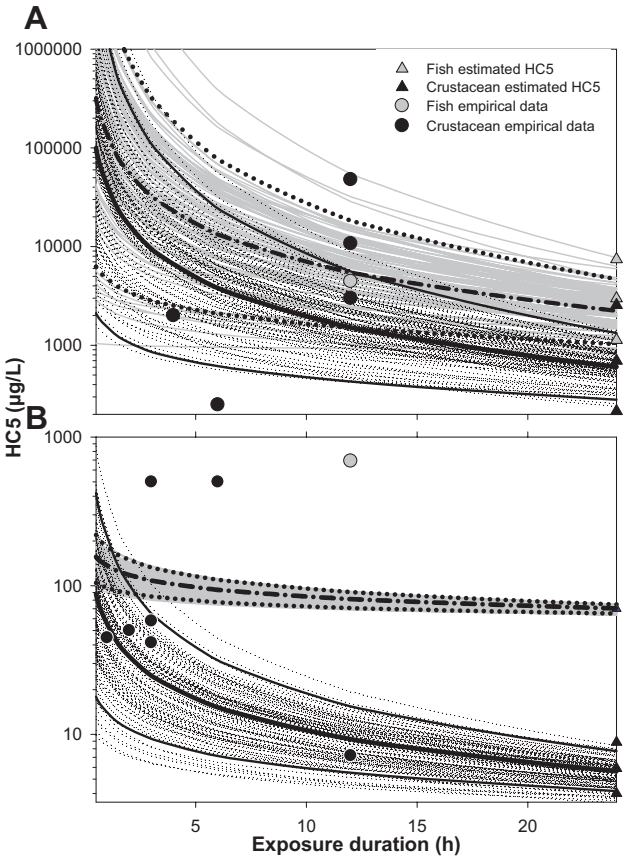

Figure 3. Mean short-exposure 5th percentile hazard concentrations (HC5s; fish = solid line; crustacean = dash-dot line) and 95% confidence intervals (95% CIs; fish = solid line; crustacean = dotted lines) estimated via bootstrapping (fish = thin dotted lines; crustacean = gray lines) for (**A**) zinc and (**B**) malathion, with differences in sensitivities between fish and crustaceans. Data for fish are shown in black-gray colors. All available chemical-specific empirical data (open circles) are shown for comparison.

mode of action, are likely to exhibit higher toxicity to invertebrates compared with fish.

Model verification

A total of 6 SSDs and their associated mean HC5s and 95% CIs were estimated via bootstrap for 4 chemicals that had sufficient acute toxicity data for short-exposure durations (1–6 h). These estimates were compared with estimates derived using longer exposures (24, 48, and 96 h) using the approach described earlier. The HC5s from these individual SSDs, not used in hazard estimate calculations, were in agreement (adjusted coefficient of determination, r^2 adjusted = 0.95, $p < 0.0001$) with those predicted from longer exposure durations (Figure 4). Of the 6 empirical HC5s, 3 fell within the estimated HC5 confidence intervals derived from longer exposures, while 2 fell above and 1 below the predicted confidence intervals. In all of these cases, estimates were within 2-fold of the empirical data values. Similarly, comparison of observed empirical acute toxicity data from short-exposure durations were in agreement with short-exposure HC5s derived from longer exposures (r^2 adjusted = 0.63, $p < 0.0001$; Figure 4). Of the 80 empirical data points available, 23 fell within the HC5 confidence intervals, while 46 were above (overprotective), and 11 below (underprotective) the HC5 confidence intervals derived from longer exposures. Comparison of estimates with observed empirical data is not a true validation of model fitness, but it provides an indication of how

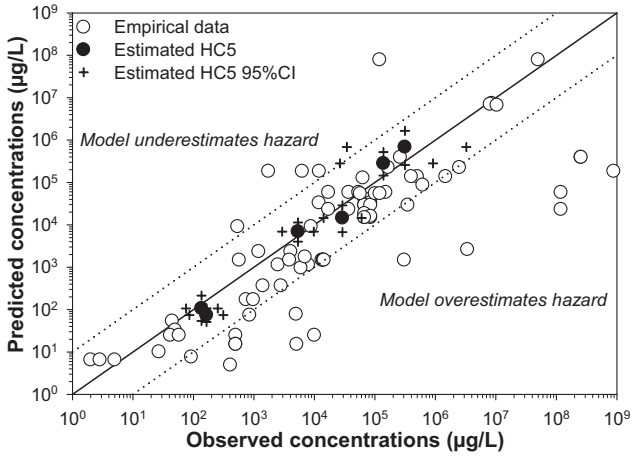

Figure 4. Verification of hazard estimates by comparison of predicted values (bootstrapped 5th percentile hazard concentrations [HC5s] and 95% confidence intervals [95% CIs]) with observed empirical data not used in derivation of hazard estimates. Comparisons were made relative to HC5s and their 95% CIs from species sensitivity distributions of chemicals with sufficient acute toxicity data for short-exposure durations (1–6 h; closed circles), and relative to all chemical-specific empirical data available for short-exposure durations (1–6 h; open circles). Points below the solid line indicate that model overestimates hazards (more protective than empirical data). The solid line represents the 1:1 ratio between observed and predicted concentrations, and the dotted lines represent a 10-fold difference relative to the solid line.

close hazard estimates are relative to the empirical toxicity data available for a small number of chemicals. The analyses above indicate that HC5s derived from longer exposure durations are more likely to overestimate than to underestimate hazards, and that the model does not grossly underestimate hazards. However, when concerns regarding the protectiveness of the HC5 exist, a more protective HCx (e.g., the lower 95% CI of the HC5, or the HC1) is highly recommended to ensure adequate protection of aquatic species.

Model application

Two real chemical spills were used to assess the applicability of hazard estimates. On 16 July 2010, a barge collision in the lower Mississippi River (USA) resulted in the release of 600 barrels of acrylonitrile (CAS number 107-13-1) [33], with the potential release of an additional 50 000 barrels. Based on conditions of river flow, NOAA's environmental forecast models estimated a potential concentration of ~1000 µg/L within 30 min of a complete release up to 9 km downstream from the collision site. Model estimates for toxicity of acrylonitrile derived from longer exposure durations with data from several taxa (mostly fish) yielded an HC5 of 138 000 µg/L for an exposure lasting 1 h. A few empirical data points were available for short exposures (≤12 h) to acrylonitrile (Figure 5) of which all but 1 were within the estimated 95% CI HC5, confirming that this approach can provide a numerically protective value for short exposures when empirical data are limiting. Although crustaceans appear to be more sensitive to acrylonitrile than fish, short-exposure hazard estimates were conservative enough and well above expected environmental concentrations to be protective of the most sensitive group.

On 25 January 2000, an overturned tank truck released 26 500 L of furfural (CAS number 98-01-1) into a drainage canal (6–12 m wide) that led to San Martin Lake, Texas, USA [34]. At the time, water depth in the canal was 1 m, and water flow was between 0.3 and 0.6 m^3/min. The concentration of furfural was 10 000 000 µg/L in water samples collected

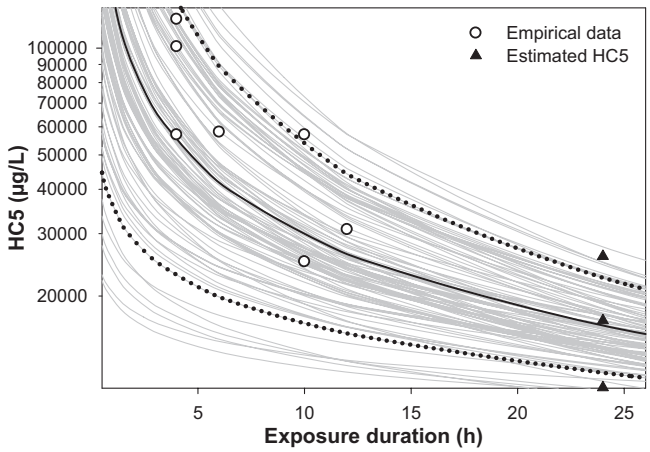

Figure 5. Bootstrapped (gray lines) 5th percentile hazard concentration (HC5) and 95% confidence interval (95% CI) estimates (solid and dotted lines, respectively) for short exposures to acrylonitrile. These estimates were derived from HC5s of species sensitivity distributions (SSDs) from longer exposures (24, 48, and 96 h) using acute toxicity data from all taxa. The triangles represent the 24-h SSD HC5 and 95% CI, and the open circles display empirical data for short exposures (1–12 h).

within a day near the spill site. A few days later, concentrations of furfural in the same area were as high as 10 000 µg/L, with concentrations in water samples from San Martin Lake 10 d after the spill between 7000 and 9000 µg/L. Subsequent water samples showed concentrations $\leq$3000 µg/L. The HC5s derived from 24-, 48-, and 96-h exposure SSDs (Figure 6) were 32 000 µg/L, 15 000 µg/L, and 11 000 µg/L, respectively, which are below the ranges observed during the spill. Although water samples were not collected for chemical analyses within the first hours after the spill, high furfural concentrations above the estimated SSD HC5s in the days after the spill suggested that environmental concentrations likely exceeded short-exposure hazard estimates derived using this approach. Data limitation prohibited the assessment of the sensitivity of fish versus crustaceans, but field observations during the spill suggested a greater sensitivity of fish.

DISCUSSION

The present study presents an approach to help responders and managers make immediate, informed decisions regarding the potential short-term hazards to aquatic resources from chemical spills. Because of the specific goal of the present study, this approach may not be suited to address long-term and chronic effects and environmental damage, which require more detailed toxicity data and information on environmental persistence and partitioning. These more thorough investigations are typically performed as part of subsequent natural resource environmental damage assessments. Consequently, risk hazards derived from this approach are intended to provide a first-tier assessment of acute effects following a spill.

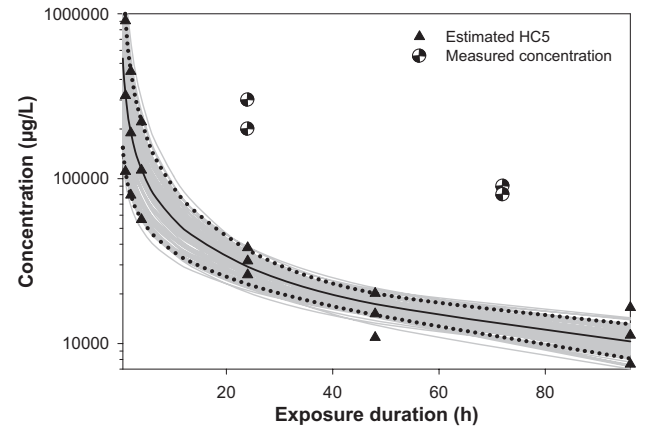

Figure 6. Comparison of bootstrapped 5th percentile hazard concentration (HC5) estimates (triangles) for different exposure durations (1, 2, 4, 24, 48, 96 h) relative to water concentrations (partially filled circles) of furfural measured following the spill near Brownsville Port, TX, USA.

The development of this approach was largely motivated by the lack of data for short exposures (a few hours), which are the data needed to assess the potential acute toxicity resulting from exposures to spilled chemicals. The greatest challenge in developing this approach was lack of data for most chemicals of interest. Of the 120 chemicals initially considered, only a few had sufficient acute toxicity data for short-exposure hazard estimates; and for some, data limitations resulted in estimates with relatively low reliability. In some cases, data restrictions were overcome by the use of ICE models [18–20], which generated acute toxicity data for several predicted species based on the known toxicity to a surrogate species (e.g., *Claassenia sabulosa*, *Hyalella azteca*, *Micropterus salmoides*, *Oncorhynchus mykiss*, *Oncorhynchus clarkii*, and *Pimephales promelas*), thus allowing the generation of SSDs. Acute toxicity data for several other chemicals (not shown here) were also evaluated. For most of these chemicals, short-exposure hazard estimates were not derived because of limited acute toxicity data (LC50 and EC50). For others, a close proximity among SSD HC5s and 95% CIs (based on LC50 and EC50 data) across SSD-specific exposure duration, or a nonlinear relationship between HC5s and exposure duration, led to nonstatistically significant ($p > 0.05$) linear bootstrap regressions. For some chemicals, the lack of a negative correlation between SSD HC5s and SSD-specific exposure durations, the main assumption needed to generate short-exposure hazard estimates, was explained by the type and number of species in each SSD. In some cases, chemicals with large data sets for 96-h exposures had greater HC5s than HC5s from less commonly tested exposure durations (e.g., 24 and 72 h), possibly due to a larger number of species and to the presence of more tolerant species in 96-h SSDs. It is apparent from this research that acute toxicity data derived from environmentally realistic exposure conditions, different from standard toxicity testing procedures, are needed to assess potential effects on aquatic receptors from chemical spills. The need for this information also extends to oil spills [8], although studies are starting to incorporate short-exposure durations (a few hours) in their toxicity testing [35,36]. One shortcoming of the hazard estimates derived with this methodology is that these were derived with information from the parent compound, and thus chemicals that degrade or hydrolyze rapidly into more or less potent metabolites can lead to over- or underestimation of hazard estimates; these are not accounted for in these calculations. Consequently, the use of hazard estimates derived using this methodology needs to consider chemical-specific properties to adequately assess potential risks to aquatic receptors.

Uncertainties in short-exposure hazard estimates were likely influenced by the data sources and the model used in their calculation. Chemical-specific HC5s for short-exposure durations were based on a variety of exposure conditions and tests, including measured versus nominal concentrations, static versus flow-through tests, freshwater versus seawater exposure conditions, and differences in life stages within a single species [14],

all of which add uncertainty to hazard estimates. Although uncertainties could be reduced by performing analyses based on water type (freshwater vs seawater), life stages (adults vs early life stages) [37,38], and taxa (fish vs crustaceans), data were limited and such detailed analyses were not possible for most chemicals of interest. A close examination of the acute toxicity data used in the development of this approach indicated that most data were for freshwater test species, and when both freshwater and seawater data were available, no apparent differences in sensitivities were observed, except for malathion. This chemical appears to be less toxic to seawater crustaceans, and inclusion of these more tolerant species could have resulted in underestimation of hazard estimates. However, because freshwater data comprised nearly 5 times the data for seawater species, inclusion of the more tolerant species did not have any effects on the HC5. For example, the 1-h malathion HC5 for the most sensitive species (crustaceans, invertebrates, and insects) derived using all freshwater and seawater data (mean HC5 54 μg/L, 95% CI 14–200 μg/L) were not different from estimates based on the most sensitive freshwater species (mean HC5 53 μg/L, 95% CI 13–230 μg/L). Although previous studies showed greater sensitivity of seawater species than freshwater species to malathion [37,39], comparisons between studies were not possible because information on data selection (taxa and exposure durations) were not included. Nevertheless, underestimation or overestimation of hazards can result from ignoring the influence of water type on acute toxicity, and therefore analyses should be performed taking into consideration water type, as well as any other identified sources of uncertainty. Furthermore, and as previously acknowledged, the chronic toxicity and cumulative environmental consequences of highly toxic chemical spills, such as one involving malathion, should be carefully addressed with alternate approaches, as the one presented here is not intended to substitute for systematic environmental assessments.

Another source of uncertainty in short-exposure hazard estimates may have been introduced by the use of a simple power law model. Although the selection of this model from a number of plausible candidate models was based on the lowest Akaike's Information Criterion value (an approach based on minimum discrepancy estimation), in some cases other models (baseline toxicity model and 3-parameter exponential decay model) may have provided a better model estimation for short-exposure HC5s. However, because of limited data specific for the purpose of the approach presented here, it was not possible to test under which circumstances one model may be preferred over another. A more systematic investigation of relationships between duration of exposures and HC5s may be re-evaluated as more data become available.

The applicability of this approach was also evaluated relative to chemical-specific modes of action. While the mode of action of most chemicals used here was baseline narcosis [40], this approach also worked well with reactive electrophiles/proelectrophiles and acetylcholinesterase inhibitors, suggesting that this approach could be used for a number of chemicals with various modes of toxicity. However, this approach likely works best for chemicals that have a mode of action that is rapidly reversible and has no latency (e.g., industrial chemicals that act by baseline narcosis), and not as well for chemicals that have a receptor-mediated mode of action. Consequently, data for chemicals with different modes of toxicity are needed to legitimize the suitability of this approach.

Despite limited data for a more complete model validation, the data available showed that this approach can predict HC5s for short-exposure durations within 1- to 2-fold of empirical value derived from SSDs and that hazards are not grossly underestimated by the model. However, if any hazard underestimation is deemed unacceptable, the same approach can be implemented to derive hazard values that protect a larger fraction of the species (e.g., HC1). In the present study, the 3-h HC5 for formaldehyde derived from longer exposure durations (6, 24, 48, and 96 h) was 281 000 μg/L (95% CI 145 000–522 000 μg/L), which was not different from the HC5 derived from the 3-h SSD (mean 139 000 μg/L, 95% CI 26 000–912 000 μg/L). Given uncertainties in the relative sensitivity of field species to a particular chemical, it is often preferable to err on the side of caution and favor hazard overestimation, such as in the cases of potassium permanganate and chlorine, when assessing potential effects under spill conditions. Furthermore, comparison of hazard estimates derived using this approach with either estimated environmental concentrations (acrylonitrile example) or measured concentrations in samples from impacted water bodies (furfural example) provides an alternative scientific approach for assessing potential concentrations of concern under typical spill conditions. However, the estimation of HC5s is dependent on available data, and further validation of this approach requires collection of water samples for chemical analysis within the first hours of a spill.

The practical application of this approach to chemical spills can be further expanded by using the 2 examples presented in the present study. In the case of the acrylonitrile spill, and under the assumption that the HC5 is a conservative estimate of hazard, responders and managers may have elected the slow release of this chemical into the river at a rate not to exceed the estimated HC5 value. Alternatively, if sensitive crustacean species or life stages were suspected to be in the area, one could consider the use of the same approach but with a more conservative hazard estimate (e.g., HC1), followed by the assessment of salvage operations that would minimize the release of this chemical into the river. By contrast, the furfural spill suggested that environmental concentrations likely exceeded short-exposure hazard estimates and consequently, responders and managers may have chosen to implement all viable options to limit the migration of this chemical downstream. In this particular spill, the need for additional analyses is evident to better characterize the immediate potential hazards to aquatic resources, by, for example, estimating the fraction of potentially affected species [22,41,42] as a function of exposure duration and expected environmental concentrations. Hypothetically, an environmental concentration of 35 000 μg furfural/L 24 h post spill (above the estimated 24-h HC5 of 32 000 μg/L) would have affected 33% of the species on the SSD. Clearly, an expansion of the approach presented in this paper with the integration of expected environmental concentrations would allow estimates of exposure–effects joint probabilities. This expanded approach would add scientific value to the information used by risk-assessors, allowing more scientifically driven management of decisions regarding the immediate risks of spills.

Although there are inherent limitations with the use of SSDs [15,16,43,44], significant efforts have been made to refine their use in environmental assessments [45–47]. The unknown sensitivity of untested species, the ecological relevance of aquatic species within the impacted area, and the impact of their unknown sensitivity on the uncertainty associated with the lower end of SSDs also should be considered when using SSDs to assess hazards from spills. Because species included in SSDs are surrogates for other species, risk-assessors and environmental managers should consider the relative sensitivity and ecological roles of resources at risk within the area impacted by the spill, as

well as their taxonomic distance from species more commonly used in toxicity testing. The HC5s provided here (mean and 95% CIs) may provide a range of protectiveness that can be used by risk-assessors and managers in their assessment of risks from a chemical spill. While estimates derived using this approach should be considered tentative, these provide risk-assessors and environmental managers with a scientifically based tool for assessing immediate potential acute toxicity effects and managing field decisions. Future efforts will focus on adding short-exposure hazard estimates to the CAFÉ database (NOAA ERD, Seattle, WA, USA, unpublished data) and developing models that take into account the mode of action and other suspected sources of uncertainty. Finally, this approach is an attempt to tackle a question that, until now, has not been answered in a rigorous fashion.

SUPPLEMENTAL DATA

Tables S1 and S2.
Figures S1–S5. (1.9 MB PDF).

Acknowledgment—The authors would like to thank K. Doe (Environment Canada) for providing acute toxicity data for acrylonitrile and C. Russom (US Environmental Protection Agency) for her contributions to CAFÉ. We would also like to thank the CAFÉ team members for their help with the development of this database: A. Mearns, G. Shigenaka, and P. Jenne (NOAA ERD), and A. Hielscher (GenWest).

REFERENCES

1. Brett MT, Goldman CR, Lubnow FS, Müller-Solger A, Bracher A, Brandt D, Brandt O. 1995. Impact of a major soil fumigant spill on the planktonic ecosystem of Shasta Lake, California. *Can J Fish Aquat Sci* 52:1247–1256.
2. Caldwell CA. 1997. Aromatic hydrocarbon pathology in fish following a large spill into the Nemadji River, Wisconsin, USA. *Bull Environ Contam Toxicol* 58:574–581.
3. US Environmental Protection Agency. 1994. National oil and hazardous substances pollution contingency plan (NCP). 40 CFR 300.145. Washington, DC.
4. Daniels WJ, Miller A. 2001. Computer resources for planning and responding to chemical emergencies. *Appl Occup Environ Hyg* 16:645–648.
5. Barker CH, Galt J. 2000. Analysis of methods used in spill response planning: Trajectory analysis planner TAP II. *Spill Sci Technol Bull* 6:145–152.
6. French McCay DP, Whittier N, Ward M, Santos C. 2006. Spill hazard evaluation for chemicals shipped in bulk using modeling. *Environ Modell Softw* 21:156–169.
7. Brungs WA, Mount DA. 1978. Introduction to a discussion of the use of aquatic toxicity tests for evaluation of the effects of toxic substances. In Cairns J, Dickson K, Maki A, eds, *Estimating the Hazard of Chemical Substances to Aquatic Life. STP 6578*. American Society for Testing and Materials, Philadelphia, PA, pp 15–26.
8. National Research Council. 2005. *Oil Spill Dispersants: Efficacy and Effects*. National Academies Press, Washington, DC.
9. Heming TA, McGuinness EJ, George LM, Blumhagen KA. 1988. Effects of pulsed- and spiked-exposure to methoxychlor on early life stages of rainbow trout. *Bull Environ Contam Toxicol* 40:764–770.
10. Heming TA, Sharma A, Kumar Y. 1989. Time-toxicity relationships in fish exposed to the organochlorine pesticide methoxychlor. *Environ Toxicol Chem* 8:923–932.
11. Hosmer AJ, Warren LW, Ward TJ. 1998. Chronic toxicity of pulse-dosed fenoxycarb to *Daphnia magna* exposed to environmentally realistic concentrations. *Environ Toxicol Chem* 17:1860–1866.
12. Seager J, Maltby L. 1989. Assessing the impact of episodic pollution. *Hydrobiologia* 188:633–640.
13. Chapman PM, Fairbrother A, Brown D. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. *Environ Toxicol Chem* 17:99–108.
14. US Environmental Protection Agency. 2012. ECOTOX Release 4.0. Duluth (MN). [Cited 13 March 2013]. Available from: www.epa.gov/ecotox.
15. Posthuma L, Suter GW, Traas TP. 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, Boca Raton, FL, USA.
16. Vaal MA, Van Leeuwen CJ, Hoekstra JA, Hermens JLM. 2000. Variation in sensitivity of aquatic species to toxicants: Practical consequences for effect assessment of chemical substances. *Environ Manage* 25:415–423.
17. Piegorsch WW, Bailer AJ. 1997. *Statistics for Environmental Biology and Toxicology*. Chapman & Hall, London, UK.
18. Dyer SD, Versteeg DJ, Belanger SE, Chaney JG, Mayer FL. 2006. Interspecies correlation estimates predict protective environmental concentrations. *Environ Sci Technol* 40:3102–3111.
19. Dyer SD, Versteeg DJ, Belanger SE, Chaney JG, Raimondo S, Barron MG. 2008. Comparison of species sensitivity distributions derived from interspecies correlation models to distributions used to derive water quality criteria. *Environ Sci Technol* 42:3076–3083.
20. Raimondo S, Vivian DN, Barron MG. 2007. Web-based Interspecies Correlation Estimation (Web-ICE) for Acute Toxicity: User Manual, Ver 2.0. EPA 600/R-07/071. US Environmental Protection Agency, Gulf Breeze, FL.
21. R Development Core Team. 2011. R: A Language and Environment for Statistical Computing, Ver 2.13, Vienna, Austria.
22. Aldenberg T, Jaworska JS, Traas TP, Posthuma L. 2002. Normal species sensitivity distributions and probabilistic ecological risk assessment. In Posthuma L, Suter GW, Traas TP, eds, *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, Boca Raton, FL, USA, pp 49–102.
23. D'Agostino RB, Stephens MA. 1986. *Goodness-of-Fit Techniques*. CRC Press, Boca Raton, FL, USA.
24. Stephens MA. 1974. EDF statistics for goodness of fit and some comparisons. *J Am Statist Assoc* 69:730–737.
25. van der Loo MPJ. 2010. *extremevalues: A Package for Outlier Detection in Unvariate Data*. R Package Version 2.0. Statistics Netherlands, The Hague, The Netherlands.
26. Efron B, Tibshirani RJ. 1994. *An Introduction to the Bootstrap*. Chapman & Hall/CRC Press, Boca Raton, FL, USA.
27. Ashauer R, Boxall A, Brown C. 2006. Predicting effects on aquatic organisms from fluctuating or pulsed exposure to pesticides. *Environ Toxicol Chem* 25:1899–1912.
28. Rozman KK, Doull J. 2000. Dose and time as variables of toxicity. *Toxicology* 144:169–178.
29. Miller FJ, Schlosser PM, Janszen DB. 2000. Haber's rule: A special case in a family of curves relating concentration and duration of exposure to a fixed level of response for a given endpoint. *Toxicology* 149:21–34.
30. Canty A, Ripley B. 2011. boot: Bootstrap R (S-Plus) Functions. R Package Version 1.3-2. R project [cited 13 march 2013]. Available from: http://cran.r-project.org/web/packages/boot/boot.pdf.
31. Davison AC, Hinkley DV. 1997. *Bootstrap Methods and Their Applications*. Cambridge University Press, Cambridge, UK.
32. Grist EPM, Leung KMY, Wheeler JR, Crane M. 2002. Better bootstrap estimation of hazardous concentration thresholds for aquatic assemblages. *Environ Toxicol Chem* 21:1515–1524.
33. US Coast Guard. 2010. Update: Coast guard responds to poution incident on Mississippi River. News Release. July 27, 2010. [Cited 13 March 2013]. Available from: http://www.uscgnews.com/go/doc/4007/1404455/Update-Coast-Guard-responds-to-pollution-incident-on-Mississippi-River.
34. National Oceanic and Atmospheric Administration. 2000. Incident news: Furfural spill, Port Brownsville. January 25, 2000. [Cited 13 March 2013]. Available from: http://incidentnews.noaa.gov/entry/507739.
35. McIntosh S, King T, Wu D, Hodson PV. 2010. Toxicity of dispersed weathered crude oil to early life stages of Atlantic herring (*Clupea harengus*). *Environ Toxicol Chem* 29:1160–1167.
36. Greer CD, Hodson PV, Li Z, King T, Lee K. 2012. Toxicity of crude oil chemically dispersed in a wave tank to embryos of Atlantic herring (*Clupea harengus*). *Environ Toxicol Chem* 31:1324–1333.
37. Hutchinson TH, Scholz N, Guhl W. 1998. Analysis of the ecetoc aquatic toxicity (EAT) database IV—Comparative toxicity of chemical substances to freshwater versus saltwater organisms. *Chemosphere* 36:143–153.
38. Hutchinson TH, Solbe J, Kloepper-Sams PJ. 1998. Analysis of the ecetoc aquatic toxicity (EAT) database III—Comparative toxicity of chemical substances to different life stages of aquatic organisms. *Chemosphere* 36:129–142.
39. Wheeler JR, Leung KMY, Morritt D, Sorokin N, Rogers H, Toy R, Holt M, Whitehouse P, Crane M. 2002. Freshwater to saltwater toxicity extrapolation using species sensitivity distributions. *Environ Toxicol Chem* 21:2459–2467.

40. Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 16:948–967.
41. Aldenberg T, Jaworska JS. 2000. Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. *Ecotoxicol Environ Safe* 46:1–18.
42. Verdonck FAM, Aldenberg T, Jaworska J, Vanrolleghem PA. 2003. Limitations of current risk characterization methods in probabilistic environmental risk assessment. *Environ Toxicol Chem* 22:2209–2213.
43. Forbes VE, Calow P. 2002. Species sensitivity distributions revisited: A critical appraisal. *Hum Ecol Risk Assess* 8:473–492.
44. van Straalen NM. 2002. Threshold models for species sensitivity distributions applied to aquatic risk assessment for zinc. *Environ Toxicol Pharm* 11:167–172.
45. Hayashi TI, Kashiwagi N. 2010. A Bayesian method for deriving species-sensitivity distributions: Selecting the best-fit tolerance distributions of taxonomic taxa. *Hum Ecol Risk Assess* 16:251–263.
46. Wang B, Yu G, Huang J, Hu H. 2008. Development of species sensitivity distributions and estimation of HC(5) of organochlorine pesticides with 5 statistical approaches. *Ecotoxicology* 17:716–724.
47. Grist EPM, O'Hagan A, Crane M, Sorokin N, Sims I, Whitehouse P. 2006. Bayesian and time-independent species sensitivity distributions for risk assessment of chemicals. *Environ Sci Technol* 40:395–401.