SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 35, No. 6, pp. 1576–1586, 2016

Hazard/Risk Assessment

THE CHEMICAL AQUATIC FATE AND EFFECTS DATABASE (CAFE), A TOOL THAT SUPPORTS ASSESSMENTS OF CHEMICAL SPILLS IN AQUATIC ENVIRONMENTS

ADRIANA C. BEJARANO,*† JAMES K. FARR,‡ POLLY JENNE,‡ VALERIE CHU,§ and AL HIELSCHER§
†Research Planning, Columbia, South Carolina, USA
‡National Oceanic and Atmospheric Administration, Seattle, Washington, USA
§Genwest Systems, Edmonds, Washington, USA

(Submitted 2 September 2015; Returned for Revision 5 October 2015; Accepted 23 October 2015)

Abstract: The Chemical Aquatic Fate and Effects (CAFE) database is a centralized repository that allows for rapid and unrestricted access to data. Information in CAFE is integrated into a user-friendly tool with modules containing fate and effects data for 32 377 and 4498 chemicals, respectively. Toxicity data are summarized in the form of species sensitivity distributions (SSDs) with associated 1st and 5th percentile hazard concentrations (HCs). An assessment of data availability relative to reported chemical incidents showed that CAFE had fate and toxicity data for 32 and 20 chemicals, respectively, of 55 chemicals reported in the US National Response Center database (2000–2014), and fate and toxicity data for 86 and 103, respectively, of 205 chemicals reported by the National Oceanic and Atmospheric Administration (2003–2014). Modeled environmental concentrations of 2 hypothetical spills (acrylonitrile, 625 barrels; and denatured ethanol, 857 barrels) were used to demonstrate CAFE's practical application. Most species in the 24-h SSD could be potentially impacted by acrylonitrile and denatured ethanol during the first 35 min and 15 h post spill, respectively, with concentrations falling below their HC5s (17 mg/L and 2676 mg/L) at 45 min and 60 h post spill, respectively. Comparisons of CAFE-based versus published HC5 values for 100 chemicals showed that nearly half of values were within a 2-fold difference, with a relatively small number of comparisons exceeding a 10-fold difference. The development of CAFE facilitates access to relevant environmental information, with potential uses likely expanding beyond those related to assessment of spills in aquatic environments. *Environ Toxicol Chem* 2016;35:1576–1586. © 2015 SETAC

Keywords: Interactive tool Environmental fate Aquatic toxicity Species sensitivity distributions Hazard concentrations

INTRODUCTION

Databases and web-based applications, containing aquatic and terrestrial toxicity data and related fate information, play an important role in supporting environmental decision processes. These tools have proved to be critical in facilitating access to data valuable in characterizing adverse ecological effects, assessing hazards, and quantifying species-specific or trophic-guild–specific risks [1–3]. Environmental databases vary in complexity and scale (e.g., species to national levels), as well as in their objectives, which are likely driven by specific scientific and management needs. For example, the Web-based Interspecies Correlation Estimation (Web-ICE) application is a database that contains over a thousand ICE models describing intrinsic relationships among species. This particular database was developed to augment toxicity estimates in the absence of empirical data, facilitating estimates of acute toxicity to aquatic (fish and invertebrates) and terrestrial (birds and mammals) species in risk assessment [3–5]. Databases can also be site-specific, like ECORISK [6]. This database contains ecological screening levels for a variety of chemicals, with ecological screening levels being available for terrestrial and aquatic organisms (e.g., avian, mammalian, invertebrate, and plant species) and several environmental compartments (soil, water, sediment, and air). This database has been used to assess potential ecological risks from chemical exposures [1,7]. Larger and more complex databases include, for example, the Toxics Release Inventory (TRI) database. The TRI was created in the late 1980s to annually disclose the toxic chemical releases and waste management activities for nearly 650 chemicals and 20 chemical categories from industrial and federal facilities. Data from the TRI have been used for several purposes including assessments of the human toxicity potential of several hundred compounds [8] and assessments of the role of TRI in driving a reduction in chemical releases [2,9].

Although several databases contain important environmental information, many of these have not been specifically designed to address concerns related to accidental releases of chemicals into aquatic environments. Several hundred potentially toxic chemicals are transported in navigable waters of the United States at volumes high enough to pose significant risks to aquatic environments in the event of accidental spills. Recent statistics from the US National Response Center [10] show that, on average, over 30 000 chemical incidents were reported each year between 2000 and 2014, with nearly half of these reaching water bodies. Most cases, however, involved relatively small spill volumes (<1000 L). Although only a small number of incidents (1% annually) reaching water bodies involved non–oil chemicals, spills of potentially acutely toxic chemicals can have devastating consequences on aquatic ecosystems [11–13]. As a result, the primary objective of developing the Chemical Aquatic Fate and Effects (CAFE; See *Data availability*) database was to synthesize fate and toxicity data in a meaningful way, to improve the decision-making process during chemical spills, as well as to provide rapid access to data. The goals of the present study are to provide a description of the steps undertaken during the development of CAFE, give examples of key elements of the database, and illustrate by means of exercises the potential use of this tool in environmental assessments of chemical spills. Data in CAFE may facilitate a better

* Address correspondence to abejarano@researchplanning.com
Published online 24 October 2015 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.3289

understanding of the effects of chemical spills, and may help assist with the development of risk estimates in aquatic environments by allowing data selection based on the specific needs of the end user. Given the broad distribution of CAFE within the larger scientific community, it is expected that this database will be used in the field of environmental sciences for purposes other than its originally intended use.

MATERIALS AND METHODS

Data sources

Fate and effects data were obtained from multiple databases and data sources. Fate data were obtained primarily from PHYSPROP [14], the Hazardous Substances Data Bank [15], the US National Institute of Science and Technology Standard Reference Database [16], and the Estimation Programs Interface Suite from the US Environmental Protection Agency (USEPA) [17]. Effects data were obtained primarily from ECOTOX [18], and secondarily from the Aquatic Toxicity (EAT3) database of the European Centre for Ecotoxicology and Toxicology of Chemicals [19], as well as from peer-review publications, and other open literature. Toxicity data for oil, dispersants, and physically and chemically dispersed oil are also included (DTox) [20]. The recurrent version of CAFE (1.1) uses all data sources updated to January 2015. However, it is expected that annual updates of CAFE will be routinely undertaken, particularly following large updates of primary data sources (e.g., ECOTOX).

CAFE's quality assurance/quality control

Several steps were undertaken to assess the quality and accuracy of data compiled in CAFE, as well as to document all data management procedures and manipulations implemented during the development of the tool (detailed in US National Oceanic and Atmospheric Administration [21]). Several data requirements were implemented during the data inclusion phase, which focused primarily on the effects module of CAFE. Each toxicity record is required to have the chemical name of the tested compound or its associated Chemical Abstracts Service (CAS) number, as well as the reported common or scientific name. Nearly the entire toxicity dataset was derived from controlled laboratory exposures using aquatic species and toxicological endpoints based on live, whole organisms, reporting environmental chemical concentrations associated with an explicitly reported exposure duration. Although not a requirement, data associated with each record also include other information, for example the life stage and water type (freshwater or saltwater) used in toxicity testing. With a few exceptions, CAFE primarily includes toxicity tests performed with only 1 chemical.

Several standardization procedures were undertaken to facilitate further data integration. For example, all scientific names were verified and checked for changes in official names, as well as for spelling errors. This verification was made through online sources including the Integrated Taxonomic Information System on-line database [22] and the World Register of Marine Species [23]. Following scientific name verification, common names of all records sharing a scientific name were standardized to a single common name, and species were assigned to 1 of the following taxonomic groups: coral, crustacean, fish, mollusk, other. Other standardization procedures included conversion to a single toxicity unit (μg/L). Missing information was also evaluated and filled in to enhance database completeness. For example, when water type was not reported, missing information was filled in using entries from other records matching both the scientific name and the life stage information. A similar approach was used to fill in missing life stage data. Efforts were also made to identify and remove duplicate data (~3% of all records) introduced through the use and combination of multiple data sources. Data were stored in a database format that was migrated into a platform (FileMaker® Pro 12) with expanded visualization capabilities. The CAFE database was tested prior to its public release, allowing the identification of major issues. To the extent possible, the tool was created to address the needs of end users with disabilities (e.g., color blind; Section 508 29 U.S.C. '749d).

Assessment of data availability

A critical step in the present study was to demonstrate the practical application of CAFE. To this end, 2 separate chemical incident databases were queried to identify the type of information associated with each chemical currently available in CAFE. The first source was the US National Response Center database [10], which was queried to identify chemical incidents between 2000 and 2014 known to have reached water bodies. Only incidents of chemicals with reported CAS numbers were included in these analyses. The second source was the incidents database requiring scientific support from the Emergency Response Division of the National Oceanic and Atmospheric Administration (NOAA) [24]. This exercise used both individual chemicals and chemicals that commonly occur as mixtures (e.g., polychlorinated biphenyls and oils), with searches in CAFE based on matching names.

Practical application

Two incidents in the Mississippi River were used to demonstrate the practical application of toxicity data in CAFE in helping inform environmental assessments following a chemical spill. The first incident occurred in 2010 and involved 2 barges each carrying approximately 25 000 barrels of acrylonitrile (CAS number 107131) [24]. The second incident occurred in 2015 following derailment of a train transporting between 1600 and 3100 barrels of denatured ethanol (CAS number 64175) [24]. During both incidents, the total volume released into the river was not determined. For the purpose of this exercise, modeled environmental concentrations at the point of release and as a function of time post spill were estimated for a hypothetical release of acrylonitrile and denatured ethanol involving 625 and 857 barrels, respectively. Environmental concentrations were calculated using NOAA's River Dilution Calculator [25], which requires river-specific parameters. For the acrylonitrile spill, river parameters were 600 m river width, 18 m water column depth and 20 000 m^3/s volume flow rate. For the denatured ethanol spill, river parameters were 570 m river width, 2 m water column depth and 1.07 m^3/s volume flow rate. Data from CAFE were queried and exported to derive species sensitivity distributions (SSDs), which are cumulative distributions of acute toxicity data that allow for comparisons of the relative sensitivities of aquatic species to the same chemical [26]. The SSDs and their associated hazard concentration (HC) values, concentrations assumed to be protective of a predetermined percentile of species, were derived by fitting the data to a log-normal distribution function, and resampling this function 2000 times to derive central tendencies and 95% confidence intervals (95% CI) [27]. Comparisons were made between modeled expected environmental concentrations of these 2 chemicals and equivalent concentrations in the SSD. These comparisons allow for estimates of the proportion of aquatic species potentially

affected, as a function of time, which could be used to characterize potential impacts to aquatic receptors.

Verification of hazard concentration values

In addition to chemical spill demonstrations, verification of HC values from SSDs directly generated in CAFE were also made to confirm the potential use of this database in aquatic assessments. The estimated 5th percentile HC (HC5) values from SSDs generated in CAFE for a variety of chemicals (e.g., metals, pesticides, aromatics, etc.) were compared with published HC5 values from 2 sources: 1) HC5 values derived based on ambient water quality criteria documents using primarily invertebrate (24 h and 48 h) and fish (96 h) acute toxicity data [3], and 2) HC5 values derived using toxicity data for standard test species without an specified exposure duration [5]. All queries in CAFE were based on searches matching the reported chemical name.

RESULTS

Database structure and layouts

Key features of CAFE (Figure 1) include data sorting into 4 scenarios (chemical, oil only, dispersant only, and dispersant and oil), which are summarized as 2 modules: the Aquatic Fate Module (chemical scenario only), and the Aquatic Effects Module (all scenarios). Because most of the data in CAFE are under the chemical scenario, further discussions focus solely on this scenario. Other features of CAFE are presented, but not discussed in detail in the present study.

Aquatic fate module

Fate information is presented and summarized in 3 submodules: structure and physical properties, environmental fate, and analytical methods and uses. The structure and physical properties submodule contains a graphic representation of the molecular structure, a physical description (e.g., odor, color and form, and human exposure), and physical properties important in modeling environmental fate (e.g., water solubility, octanol–water partitioning coefficient; biodegradation rates; chemical partitioning). The environmental fate submodule contains fate information (e.g., soil adsorption, biodegradation half-life, estimates of volatilization from water, biodegradation, photo-oxidation, removal from sewage treatment, media partitioning) and detailed environmental explanations on fate and behavior. The analytical methods and submodule used to quantify the target chemical across different media contains analytical chemistry methods (e.g., method numbers) and a description of the current and intended use of individual chemicals, including citations and the original source of this information (Figure 2).

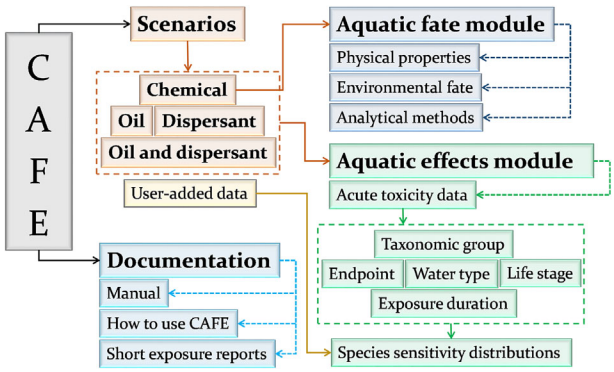

Figure 1. Diagram summarizing key features of the Chemical Aquatic Fate and Effects (CAFE) database.

The information contained in the fate module of CAFE can be used by environmental modelers, which, combined with site-specific information (e.g., river flow, water volume, salinity), could be useful in generating plume behavior models, 3-D and forecast trajectories, and predicted environmental concentrations. This information is also useful in informing the possible fate of the spilled chemical across different environmental media. To date, 32 377 chemicals in CAFE have fate information, although not all chemicals have data on all parameters. For example, most chemicals (99%) have information on the most commonly studied or estimated physical properties (e.g., water solubility, vapor pressure, octanol–water partitioning coefficient, Henry's law constant), whereas only a small number of chemicals ($\leq$4%) have information on less common parameters (e.g., ionization potential, flash point, explosive/flammable limits, heat of vaporization).

Aquatic effects module

Acute toxicity data in CAFE are sorted and displayed based on specific attributes, including: taxonomic groups (corals, crustaceans, fish, mollusks, and others), life stages (adults, embryos, juveniles, larvae, unknown), acute toxicity data (median lethal concentration [LC50], median effects concentration [EC50], lowest-observed-effect concentration [LOEC], and no-observed-effect concentration [NOEC]), water type (fresh, saltwater), and exposure duration (24 h, 48 h, 72 h, and 96 h). These attributes are displayed in CAFE across several windows where data selection options are given to facilitate query customization (Figure 3). Because of the wide variability in toxicity testing practices, an applicability score was assigned to each record in the database. This applicability score, developed with the sole purpose of guiding data selection to inform chemical spill response, was as follows: high applicability—toxicity data with reported concentrations on the basis of measured concentrations, performed under flow-through conditions, and using individual chemicals containing $\geq$90% active ingredient purity; moderate applicability—toxicity data with reported concentrations on the basis of measured concentrations, performed under static or static renewal conditions, and using individual chemicals containing 75% to less than 90% active ingredient purity; and low applicability—toxicity data with reported concentrations on the basis of nominal or unmeasured concentrations, toxicity data with unclear reported concentrations (nominal, unmeasured, or measured), and laboratory conditions, and using individual chemicals containing <75% active ingredient purity.

Once query selections are made, the resulting toxicity data are plotted, with standardized concentration units (μg/L) and by exposure duration (24 h, 48 h, 72 h, and 96 h), in the form of SSDs. In CAFE, SSDs are generated for datasets with a minimum of 5 species by fitting the empirical toxicity data to a logistic function defined as $F(x) = \frac{L}{(1+c*e^{(a*x)})}$ where L is the curve's maximum value, and a and c the regression coefficients. For each individual species and by scientific name, the geometric mean of all reported concentrations is calculated and used to derive SSDs. These curves are plotted over a colored background representing a common scale of relative toxicity (adopted from the USEPA Office of Pesticide Programs [28]) for aquatic organisms, which are as follows: very highly toxic (<100 μg/L), highly toxic (100–1000 μg/L), moderately toxic (1000–10 000 μg/L), slightly toxic (10 000–100 000 μg/L), and practically nontoxic (>100 000 μg/L). The SSDs are useful in that HCs can be derived as a measure of chemical risk. With

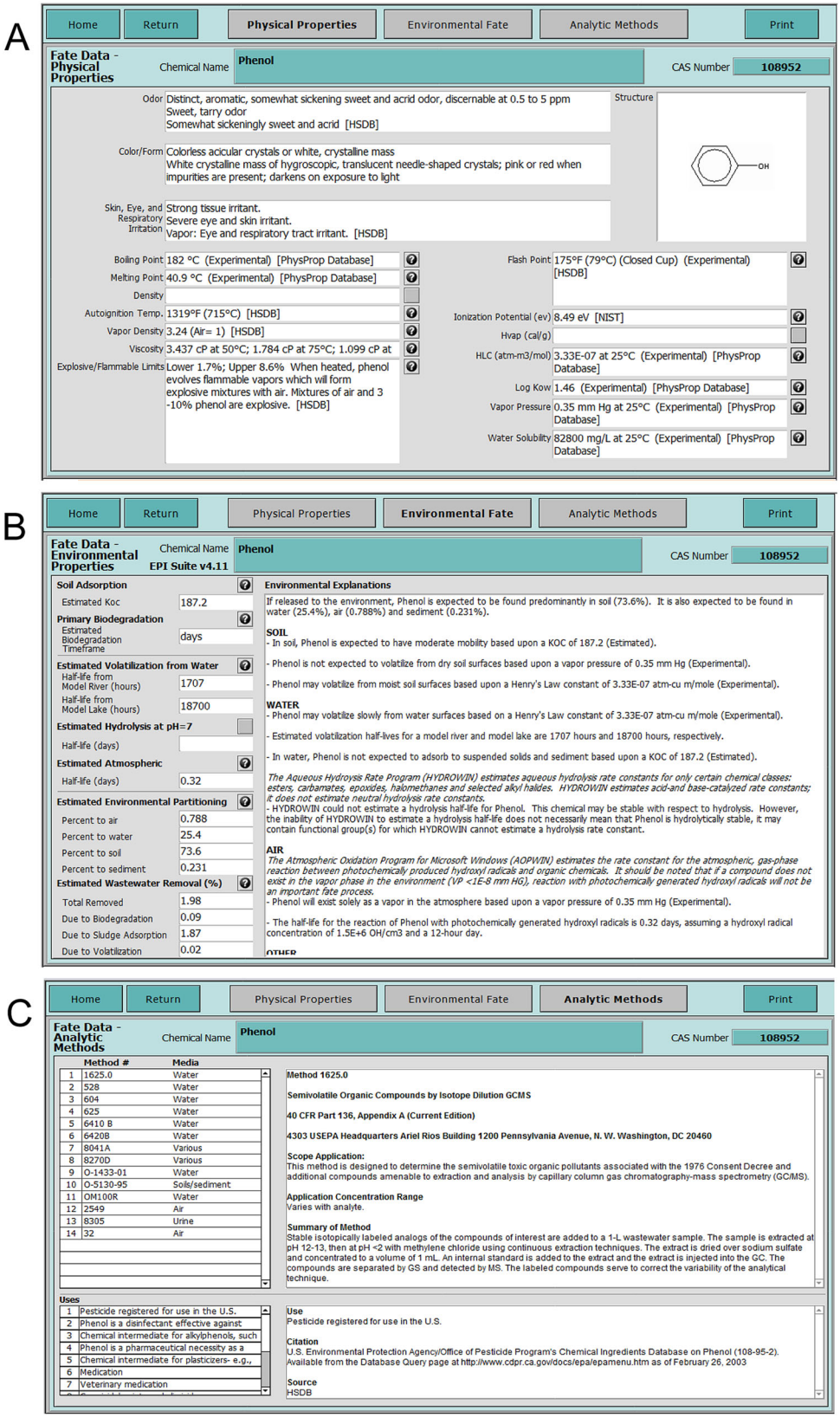

Figure 2. Components of the aquatic fate module of the Chemical Aquatic Fate and Effects database: Structure and physical properties (**A**), environmental fate (**B**), and analytic methods (**C**).

each SSD, CAFE displays HC1 and HC5 values, which are the concentrations assumed to protective of 99% and 95%, respectively, of the species in the SSD. Although there is debate in the scientific literature regarding the appropriateness of one versus another HC percentile, the HC5 was selected because this is the most commonly used percentile, whereas a lower percentile (HC1) would offer an additional safety factor that is preferable for very highly toxic chemicals. However, the

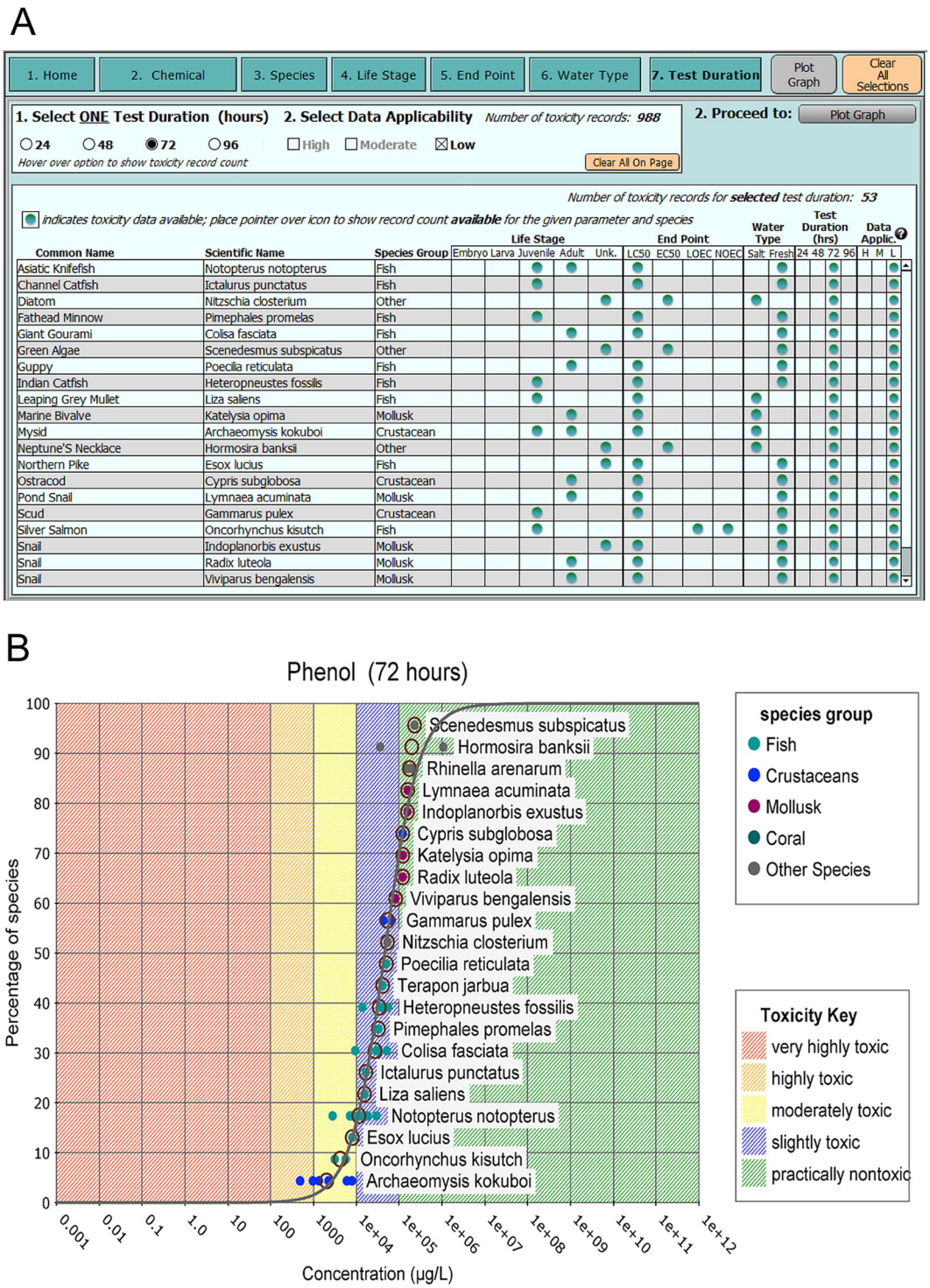

Figure 3. Examples of data visualization windows (**A**) and display of toxicity data in the form of species sensitivity distributions (**B**) within the effects module of the Chemical Aquatic Fate and Effects database.

use of HC1 values requires a greater degree of curve fitting certainty, because uncertainty tends to increase toward the tail ends of the SSDs.

To date, the effects module of CAFE contains toxicity data for 2533 unique aquatic species and 4498 chemicals, encompassing a total of 130 707 individual records, most of which (~95%; 123 548 records) came from ECOTOX [18] (Table 1). Most records in CAFE are for fish, toxicity values reported as LC50s, tests performed with freshwater, and records with low applicability for the purpose of chemical spills. The aquatic species with the most records are the water flea (*Daphnia magna*; 9475 records, 7%), rainbow trout (*Oncorhynchus mykiss*; 8840 records, 7%), fathead minnow (*Pimephales promelas*; 7546 records, 6%), bluegill (*Lepomis macrochirus*; 5707 records, 4%), a microalgae (*Selenastrum capricornutum*; 2495 records, 2%), and common carp (*Cyprinus carpio*; 2349 records, 2%). Each of the remaining aquatic species comprise ≤1% of the records, with the top 20 data-rich species comprising 43% of all records. Data for coral species are heavily under-represented in CAFE.

Table 1. Breakdown by key features of toxicity records in the effects module of the Chemical Aquatic Fate and Effects (CAFE) database

Feature	Metric	No. of records (% of total)
Taxonomic group	Coral	122 (0.1)
	Crustacean	30 346 (23)
	Fish	59 805 (46)
	Mollusk	8723 (7)
	Other	31 691 (24)
Life stage	Embryo	5016 (4)
	Larva	19 489 (15)
	Juvenile	46 508 (36)
	Adult	17 577 (13)
	Unknown	42 128 (32)
Endpoint	LC50	87 121 (67)
	EC50	20 125 (15)
	LOEC	11 444 (9)
	NOEC	12 017 (9)
Water type	Saltwater	26 553 (20)
	Freshwater	104 154 (80)
Exposure duration	24 h	32 893 (25)
	48 h	34 693 (27)
	72 h	10 719 (8)
	96 h	52 402 (40)
Applicability	Moderate	7287 (6)
	High	7366 (6)
	Low	116 054 (89)

LC50 = median lethal concentration; EC50 = median effective concentration; LOEC = lowest-observed effect concentration; NOEC = no-observed effect concentration.

The number of toxicity records per unique chemical varies widely, from 22% of chemicals having only 1 record to 4% of chemicals having >1000 records (Figure 4). The number of unique species also varies, resulting in only 22% of all chemicals having enough data to generate at least 1 SSD. Despite these limitations, the data currently in CAFE are sufficient to generate a total of 2442 SSDs, with most SSDs available for 96 h exposures (797 total SSDs), followed by 48-h, 24-h, and 72-h exposure durations (605, 529, and 211 total SSDs, respectively).

Other features of CAFE

In addition to the fate and effects modules, CAFE also contains a module that allows users to enter their own toxicity data, which can be displayed in conjunction with data already in CAFE. An additional module of CAFE includes reports for selected chemicals (developed following Bejarano and Farr [27]) describing estimated short acute exposures and HC values for durations not typically collected in standard toxicity testing (e.g., 1 h, 2 h, 4 h, or 8 h LC50 or EC50 data), but more closely associated with the type of exposures generally seen during chemical spills. These chemicals were selected because they followed 1 or more of the following criteria: chemicals involved in accidents, potentially toxic to humans and biological resources, reasonably water soluble, and shipped in bulk generally in large quantities.

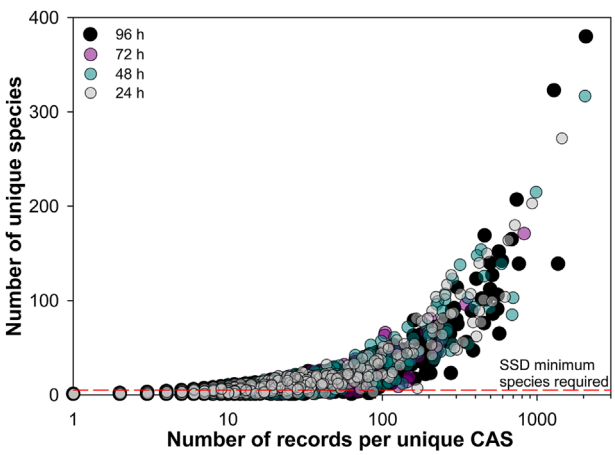

Figure 4. Number of toxicity records by exposure duration for each chemical in the effects module of the Chemical Aquatic Fate and Effects database. The dashed line represents the minimum number of species (5) required to generate species sensitivity distributions (SSDs).

Assessment of data availability

A total of 55 chemicals in the US National Response Center database (2000–2014) were involved in at least 5 incidents reaching water bodies. For these chemicals, data availability in CAFE varied broadly, from 3 chemicals having only fate data to 52 chemicals having both fate and effects data (Table 2). For 23 chemicals, the fate module contained only physical properties, whereas for most chemicals (52) the effects module contained various amounts of toxicity data. For example, 16 chemicals had insufficient toxicity data to generate SSDs, whereas 20 chemicals had enough data to generate SSD for all 4 exposure durations.

Similarly, since 2003, the Emergency Response Division (NOAA) has responded to at least 2500 incidents involving either chemical or oil releases [24]. A total of 205 unique compounds were identified based on incidents reported as single chemicals or oils. Fate data were available in CAFE for 61% (86) of these chemicals, and toxicity data were available for 53% (74) and 45% (29) of chemicals and oils, respectively, with various amounts of data across compounds (Figure 5). Toxicity records varied from 1 to >1000 for individual chemicals, with 84% (62) of chemicals having ≤100 records. The amount of toxicity data for oils is limited and varied from 1 to over 165 records for individual oils, with 80% (23) of oils having ≤30 toxicity records. The number of unique species also varies, resulting in 61% of chemicals and 38% of oils having enough data to generate at least 1 SSD.

Practical application

Two spill incidents were used to demonstrate the practical application of fate and toxicity data in CAFE in informing potential environmental impacts. Modeled environmental concentrations at the point of release for the hypothetical releases of acrylonitrile and denatured ethanol showed concentrations rapidly declining over time (Figure 6). In the case of the acrylonitrile spill, fate data suggest that this chemical may slowly partition into air (Henry's law constant, 0.000138 atm-cu m/mole) and is not expected to partition into suspended solids or sediments (organic carbon partitioning coefficient [K_{OC}] 8.511). Concentrations within 180 m of the point of release were as high as 4520 mg/L within the first 10 min, rapidly declining to <1 mg/L within 2 h post spill. Given the short duration of the exposure, only the 24-h SSD was generated for this demonstration. Comparisons of modeled environmental concentration relative to the 24-h SSD indicated that most aquatic species in the 24-h SSD could be potentially affected during the first 35 min post spill, with modeled environmental concentrations falling below the HC5 (17 mg/L) at 45 min post spill. These assessments are assumed to be conservative, because the SSD used was developed using an exposure duration much longer than that of the exposure resulting from this hypothetical

Table 2. Data in the Chemical Aquatic Fate and Effects (CAFE) database[a]

			Data in CAFE					
					SSDs			
Chemical name	CAS no.	No. of incidents	Fate	Toxicity	24 h	48 h	72 h	96 h
Ethylene glycol	000107-21-1	1748	X	X	X	X		X
Sulfuric acid	007664-93-9	412	X	X				
Sodium hydroxide	001310-73-2	285	X	X				
Benzene concentrate	000071-43-2	225	X	X	X	X	X	X
Ammonia, anhydrous	007664-41-7	215	X	X	X	X	X	X
Hydrochloric acid	007647-01-0	164	X	X				
Sodium hypochlorite	007681-52-9	132	X	X	X	X	X	X
Chorine	007782-50-5	132	X	X	X	X	X	X
Styrene	000100-42-5	113	X	X	X	X	X	X
Toluene	000108-88-3	113	X	X	X	X	X	X
Mercury	007439-97-6	104	X	X	X	X	X	X
Phosphoric acid	007664-38-2	68	X	X				
Arsenic	007440-38-2	53	X	X				
Acetone	000067-64-1	49	X	X	X	X	X	X
Hydrogen sulfide	007783-06-4	34	X	X	X	X	X	X
Nitric acid	007697-37-2	33	X	X				
Vinyl chloride	000075-01-4	31	X					
Zinc bromide	007699-45-8	31	X	X				
Trichloroethylene	000079-01-6	27	X	X	X	X		X
Potassium hydroxide	001310-58-3	25	X	X				
Phenol	000108-95-2	24	X	X	X	X	X	X
Acetic acid	000064-19-7	23	X	X	X	X		X
Ethylene oxide	000075-21-8	23	X	X				
Ferric chloride	007705-08-0	22	X	X	X	X		X
Methyl methacrylate	000080-62-6	17	X	X	X	X		
Hydrofluoric acid	007664-39-3	17	X					
2-Propenenitrile	000107-13-1	16	X	X	X	X		X
Sulfuric acid, aluminum salt (3:2)	010043-01-3	16	X[b]	X	X	X	X	X
Hexane	000110-54-3	15	X	X				X
Sodium cyanide	000143-33-9	15	X	X	X	X	X	X
Sulfuric acid, iron (3+) salt (3:2)	010028-22-5	15	X[b]	X				
Hydrogen cyanide	000074-90-8	13	X	X				X
Vinyl acetate	000108-05-4	12	X	X	X	X		
Potassium permanganate	007722-64-7	12	X	X	X	X		X
Iron chloride	007758-94-3	12	X	X				
Tetrachloromethane	000056-23-5	11	X	X	X	X		X
Tetrachloroethene	000127-18-4	11	X	X	X	X	X	X
Sulfurous acid, monosodium salt	007631-90-5	11	X	X		X		X
Nitrous acid, sodium salt	007632-00-0	11	X	X	X	X	X	X
Naphthalene	000091-20-3	10	X	X	X	X	X	X
Ethylbenzene	000100-41-4	10	X	X	X	X	X	X
Cyclohexane	000110-82-7	10	X	X	X	X		X
Nitric acid, sodium salt (1:1)	007631-99-4	10	X	X	X	X	X	X
Acetic acid, zinc salt	000557-34-6	9	X	X	X			X
Lead acetate	000301-04-2	8	X	X	X	X	X	X
Pentachlorophenol	000087-86-5	7	X	X	X	X	X	X
Ethyl acetate	000141-78-6	7	X	X	X	X		
1,1′-Osybisethane	000060-29-7	6	X	X				
2-Propenoic acid	000079-10-7	6	X	X				
Acrolein	000107-02-8	6	X	X	X	X		X
Ethyl acrylate	000140-88-5	6	X	X				
Sodium hydrosulfide	016721-80-5	6	X	X				
2,2′Oxybisethanol	000111-46-6	5	X	X				
Zinc chloride	007646-85-7	5	X	X	X	X	X	X
Nitric oxide	010102-43-9	5	X					

[a]Including availability of species sensitivity distributions (SSDs), for chemicals with Chemical Abstracts Service (CAS) numbers involved in spill incidents between 2000 and 2014 reported to have reached water bodies (US National Response Center Database [10]). Only chemicals with at least 5 incidents are shown.
[b]Only physical properties were available.
SSDs = species sensitivity distributions.

spill. For acrylonitrile, estimated short acute exposures for 1 h, 2 h, 4 h, or 8 h are available in CAFE, with calculated HC5s ranging from 140 mg/L (1 h) to 34 mg/L (8 h). Based on these HC5 values, modeled environmental concentrations fall below the 1 h, 2 h, 4 h, and 8 h HC5 values at 22 min, 26 min, 30 min, and 36 min post spill, respectively.

In the case of the ethanol spill, fate data suggest that this chemical may slowly partition into air (Henry's law constant, 5×10^{-6} atm-cu m/mole) and is not expected to partition into suspended solids or sediments (K_{OC} 1.045). Concentrations within 20 m of the point of release were as high as 407 700 mg/L within the first h, declining to 1485 mg/L at 100 h post spill.

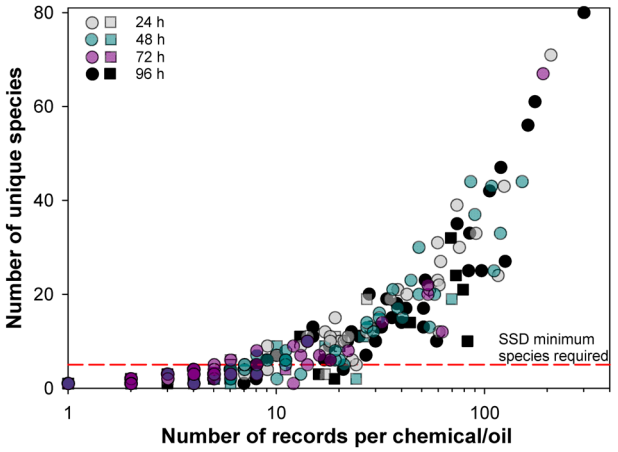

Figure 5. Number of toxicity records by exposure duration for each individual chemical (circles; 140) and oil (squares; 65) in the effects module of the Chemical Aquatic Fate and Effects database related to incidents requiring scientific support [24]. The dashed line represents the minimum number of species (5) required to generate SSDs.

Because of the relatively long exposure duration resulting from this hypothetical spill, 2 SSDs (24 h and 96 h) were generated for this demonstration. Comparisons of modeled environmental concentration relative to the 24-h SSD (least conservative) indicated that most aquatic species in the 24-h SSD could be potentially affected during the first 15 h post spill, with modeled environmental concentrations falling below the HC5 (2676 mg/L) at 60 h post spill. In contrast, comparisons of modeled environmental concentration relative to the 96-h SSD (most conservative) indicated that most aquatic species in the 96-h SSD could be potentially affected during the first 21 h post spill, with modeled environmental concentrations remaining above the HC5 (1323 mg/L) for nearly the entire modeled period (100 h).

Verification of hazard concentration values

Queries of data in CAFE allowed for comparisons of HC5 ($HC5_{CAFE}$) values for 100 chemicals. Comparisons were made between $HC5_{CAFE}$ values from each exposure duration SSD (24 h, 48 h, 72 h, and 96 h; 96, 97, 76, and 99 comparisons, respectively) relative to published HC5 values ($HC5_{Published}$)

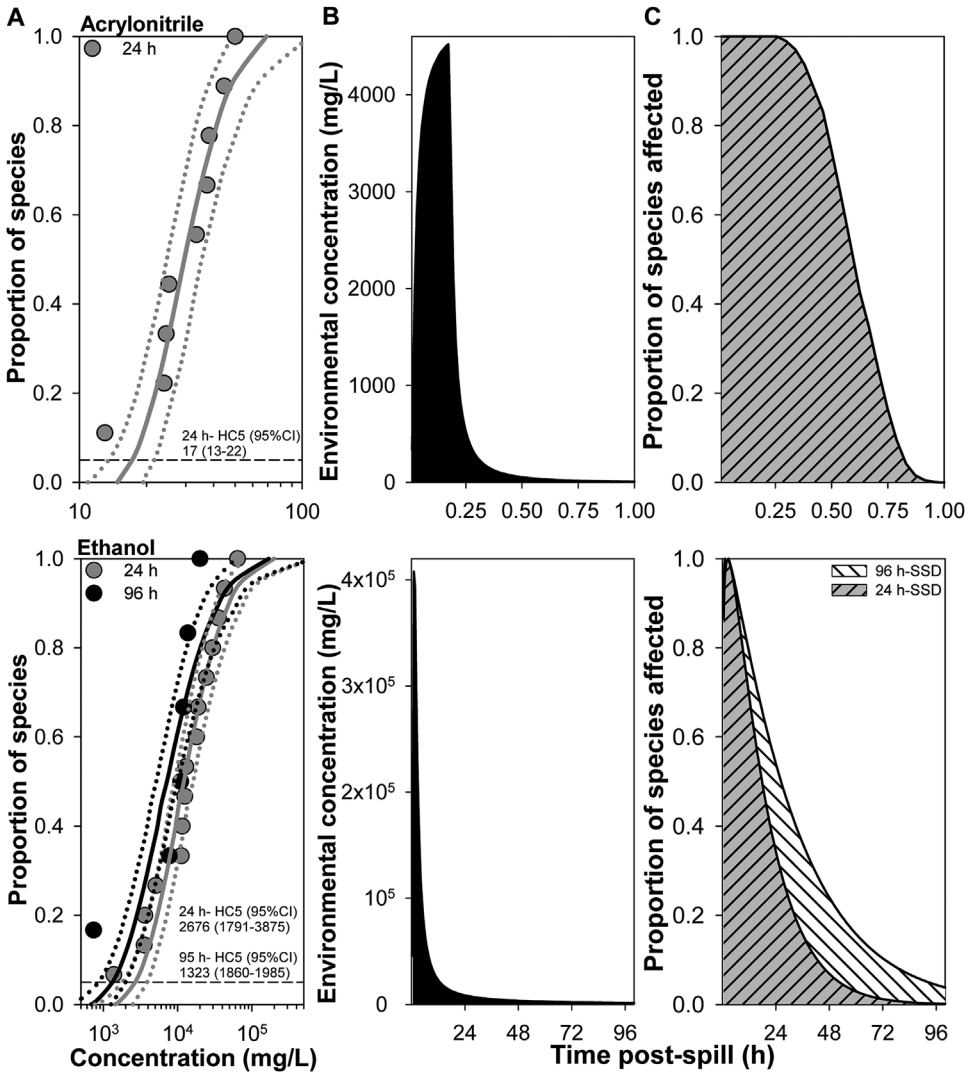

Figure 6. Practical application of toxicity data in the Chemical Aquatic Fate and Effects database to chemical spills (top: acrylonitrile; bottom: ethanol) in aquatic environments. Plots show (**A**) species sensitivity distributions (SSDs) for at least 1 exposure duration (h), (**B**) estimated environmental concentrations of the spilled chemical at the point of release as a function of time post release, and (**C**) proportion of species affected based on comparisons of environmental concentrations versus concentrations in the SSD. HC = hazard concentration; CI = confidence interval.

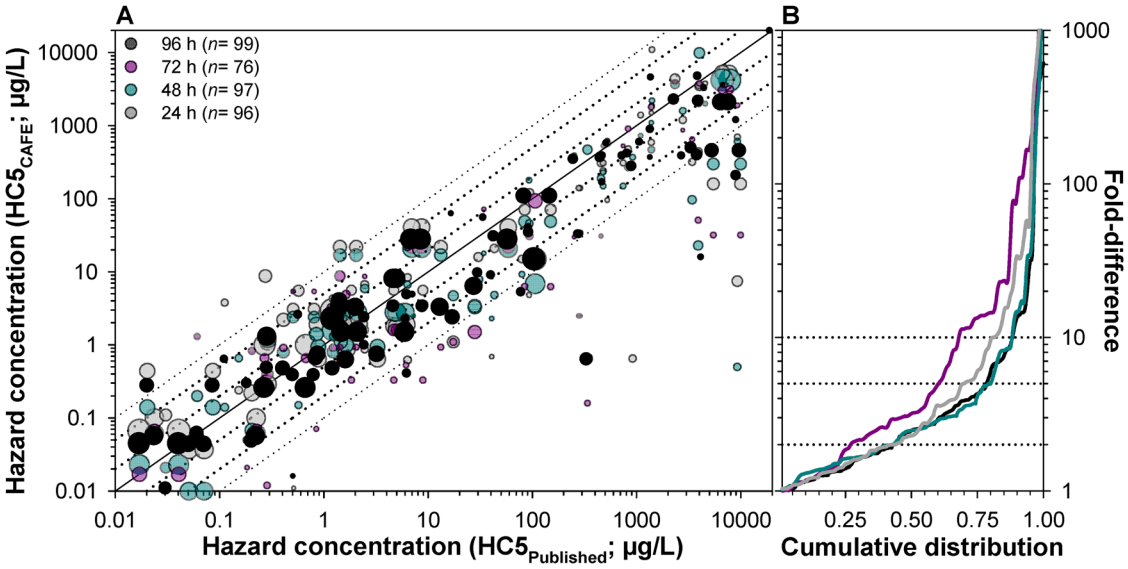

Figure 7. Comparison between the Chemical Aquatic Fate and Effects database hazard concentration ($HC5_{CAFE}$) values, by exposure duration, and published values ($HC5_{Published}$) [3,5]. (**A**) The circle size represents the number of species in each species sensitivity distribution generated in CAFE (min. = 5 species, max = 203 species). The solid line represents the 1:1 line (equal HC5 values), and the dashed lines represent the 2-, 5- and 10-fold differences between $HC5_{CAFE}$ and $HC5_{Published}$ values. (**B**) Cumulative distribution of comparisons based on fold-differences between $HC5_{CAFE}$ and $HC5_{Published}$ values, by exposure duration, with dashed horizontal lines representing 2-, 5-, and 10-fold differences.

from 2 sources [3,5]. Except for $HC5_{CAFE}$ values from 72-h exposures, 43%, ≤79%, and ≤88% of all $HC5_{CAFE}$ values were within a 2-, 5-, and 10-fold difference, respectively, of the $HC5_{Published}$ values (Figure 7). Larger deviations from the $HC5_{Published}$ values were noted across $HC5_{CAFE}$ values from 72-h exposures, which had the least amount of data and smaller numbers of species per SSD. The $HC5_{CAFE}$ values with smaller fold differences may be achievable by performing data queries in CAFE that match the characteristics of the data used to generate $HC5_{Published}$. However, these data were not readily available, limiting further comparisons.

DISCUSSION

The CAFE database was developed with the primary goal of facilitating access to fate and toxicity data in the event of a chemical spill. Some of the characteristics of CAFE, including summarization of chemical and environmental properties, interactive query windows, and automatic derivation of SSDs, make CAFE a tool with potential applications other than its originally intended use. For example, it has long been recognized that gathering toxicity data is one of the first steps required to perform effects characterization in ecological risk assessment [29]. Consequently, CAFE could also serve as the first stop of ecological risk assessors seeking to identify critical information including initial screening levels and sensitive species or life stages. Furthermore, the fate module of CAFE can provide initial information on the partitioning of a chemical across different media, as well as guidance on how to sample specific media. In addition, and as demonstrated in the present study with 2 hypothetical spills, calculation of the proportion of species affected could easily be achieved with the centralized toxicity data contained in CAFE by means of comparisons of SSDs with expected exposure concentrations. These types of assessments have been made in the past with other data sources [26,30,31]. In addition, as shown through the verification of HC5 values for 100 chemicals, CAFE can generate values comparable to those available in the literature [3,5]. Nearly half of all HC5 comparisons for all but 1 exposure duration (72 h) were within a 2-fold difference of the published values, with most comparisons being within a 5-fold difference. These fold-differences are within the range of those obtained in the derivation of HC5 values from ICE-based SSDs [3,5].

Although CAFE contains data for thousands of chemicals, one of the greatest challenges deals with data availability. The strengths of CAFE are mostly for chemicals of interest determined by the USEPA (e.g., pesticides, primary pollutants), because the primary data source is ECOTOX [18]. Within the context of chemical spills, chemicals of priority for inclusion in CAFE are those shipped in bulk and considered to be highly hazardous, which therefore have a high risk to be involved in spills of large volumes. Priority chemicals also include emerging chemicals of concern, including biodiesels and petroleum distillates (i.e., bitumen). Incorporation of additional chemical information requires identification of target chemicals, followed by gathering and extraction of information from published literature. Although prioritizing target chemicals has already been achieved by identifying the top 100 chemicals in the Computer-Aided Management of Emergency Operations (CAMEO) software [32] with limited data in CAFE, in most cases, gathering of toxicity data commonly occurs only after an incident has occurred. However, efforts are underway to augment the number of toxicity records for priority chemicals known to be data deficient. For instance, models based on interspecies correlations [3–5,33] or based on quantitative structure–activity relationships [34–36] could be integrated into CAFE to predict aquatic toxicity based on mathematical relationships.

The type of data that are generally most useful for the purpose of performing chemical spill assessments are those that are performed under short exposure duration (typically of a few hours), which account for the rapid dilution that generally occurs in the water column (i.e., spiked exposures). In addition, issues with acute toxicity data following exposure to physically and chemically dispersed oils (discussed elsewhere [37]),

including toxicity metrics reported as nominal concentrations and lack of environmentally realistic exposure conditions in most toxicity testing, also apply to data limitations in CAFE. As a result, most records in CAFE are classified as having a low applicability for the purpose of spill response. Despite these limitations, assessments of potential adverse impacts following a chemical spill can still be made with data in CAFE, which could be assumed to be conservative by using relatively longer exposure durations (i.e., from 24–96-h tests).

Despite the potential for widespread use of CAFE by the larger scientific community, there are a number of opportunities for improvements of the first version of the tool. One of the greatest challenges in developing CAFE dealt with the design of an interactive tool that was intuitive enough to enhance and facilitate its use. Although CAFE's design follows standard practices and principles [38] aimed at facilitating data accessibility by the user, large amounts of various types of data could hinder the flexibility of interactive tools. From its conception, CAFE's design has been driven by needs related to chemical spills, and consequently, future modifications would likely be driven by other data needs identified by the larger scientific community that also support NOAA's mission. In addition, current limitations of CAFE are reflective of the programming challenges of developing this type of tool. For example, goodness-of-fit tests are used to evaluate how well the logistic curve fits the empirical data [39,40], confidence intervals associated with the mean response are useful in assessing uncertainty [26,41], and different types of family distributions are available to generate SSDs [26,42]. Although these and other improvements could enhance the usability of CAFE, these specific types of tests and approaches have not been implemented because of programming challenges associated with the platform used to develop this tool. However, solutions are being explored to improve data visualization and confidence in the results generated in CAFE.

As shown in the present study, CAFE provides rapid and unrestricted access to centralized fate and effects data useful in characterizing the potential environmental partitioning and aquatic toxicity of chemicals of interest. Through this data repository, information can be readily obtained by decision-makers during spills, incidents, and other environmental assessments, facilitating real-time decisions. The database will likely have expanded capabilities in the future to suit the needs of the larger scientific community.

Acknowledgment—The CAFE development team thanks data providers, and acknowledge the special contributions of the US Environmental Protection Agency, the European Centre for Ecotoxicology and Toxicology of Chemicals, and the Syracuse Research Corporation. We are thankful to several individuals for their support during the development of CAFE and to those who provided valuable input during its testing phase. Their input was important in identifying opportunities for improvements. CAFE (Ver 1.1) is available for download, free of charge, from http://response.restoration.noaa.gov/cafe.

Disclaimer—Reference in CAFE to commercial products, processes, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the US government. The US government does not warrant or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, product, or process disclosed.

Data availability—Data sources are specifically stated in the body of the present study, including, for example, the USEPA ECOTOX database. For all other inquiries, please contact the corresponding author (abejarano@researchplanning.com).

REFERENCES

1. Budd R, Gonzales G, Fresquez P, Lopez E. 2004. The uptake and distribution of buried radionuclides by pocket gophers (*Thomomys bottae*). *J Toxicol Environ Health A* 39:611–625.
2. Fung A, O'rourke D. 2000. Reinventing environmental regulation from the grassroots up: Explaining and expanding the success of the toxics release inventory. *Environ Manage* 25:115–127.
3. 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.
4. Raimondo S, Mineau P, Barron M. 2007. Estimation of chemical toxicity to wildlife species using interspecies correlation models. *Environ Sci Technol* 41:5888–5894.
5. Awkerman JA, Raimondo S, Jackson CR, Barron MG. 2014. Augmenting aquatic species sensitivity distributions with interspecies toxicity estimation models. *Environ Toxicol Chem* 33:688–695.
6. Los Alamos National Laboratory. 2008. ECORISK database (Release 2.3). LA-UR-05-7424. Environmental Stewardship Division Environmental Remediation and Surveillance Program, Los Alamos, NM, USA.
7. Ryti RT, Markwiese J, Mirenda R, Soholt L. 2004. Preliminary remediation goals for terrestrial wildlife. *Hum Ecol Risk Assess* 10:437–450.
8. Hertwich EG, Mateles SF, Pease WS, McKone TE. 2001. Human toxicity potentials for life-cycle assessment and toxics release inventory risk screening. *Environ Toxicol Chem* 20:928–939.
9. Anton WRQ, Deltas G, Khanna M. 2004. Incentives for environmental self-regulation and implications for environmental performance. *J Environ Econ Manag* 48:632–654.
10. US Coast Guard. 2015. United States Coast Guard National Response Center. [cited 2015 July 6]. Available from: http://www.nrc.uscg.mil
11. Dowson P, Scrimshaw M, Nasir J, Bubb J, Lester J. 1996. The environmental impact of a chemical spill from a timber-treatment works on a lowland river system. *Water Environ J* 10:235–244.
12. Spazier E, Storch V, Braunbeck T. 1992. Cytopathology of spleen in eel *Anguilla anguilla* exposed to a chemical spill in the Rhine River. *Dis Aquat Organ* 14:1–22.
13. 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.
14. Syracuse Research Corporation. 2012. Physical properties database, PHYSPROP. North Syracuse, NY, USA. [cited 2015 July 6]. Available from: http://www.srcinc.com/what-we-do/product.aspx?id=133
15. US National Library of Medicine. 2015. Hazardous Substances Data Bank. Bethesda, MD, USA. [cited 2015 July 6]. Available from: http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB
16. US National Institute of Standards and Technology. 2014. NIST Standard Reference Database. Gaithersburg, MD, USA. [cited 2015 July 6]. Available from: http://www.nist.gov/srd/
17. US Environmental Protection Agency. 2012. Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. Washington, DC. [cited 2015 July 6]. Available from: http://www.epa.gov/opptintr/exposure/pubs/episuite.htm
18. US Environmental Protection Agency. 2012. ECOTOX. Release 4.0. Ecotoxicology Database. Duluth, MN, USA. [cited 2015 July 6]. Available from: www.epa.gov/ecotox
19. European Centre for Ecotoxicology and Toxicology of Chemicals. 2003. Aquatic hazard assessment II. Technical report 91. Brussels, Belgium. [cited 2015 July 6]. Available from: http://www.ecetoc.org/
20. Bejarano AC, Dahlin J. 2013. DTox: A quantitative database of the toxicity of dispersants and chemically dispersed oil, version 1.1 [computer software]. A project for the National Oceanic and Atmospheric Administration, and the University of New Hampshire's Coastal Response Research Center (contract 13-034). Research Planning, Columbia, SC, USA.
21. US National Oceanic and Atmospheric Administration, Emergency Response Division. 2015. Chemical Aquatic Fate and Effects (CAFE) Database: User Manual. Version 1.1. Office of Response and Restoration, Seattle, WA, USA.
22. US Integrated Taxonomic Information System. 2014. Integrated Taxonomic Information System on-line database. [cited 2015 July 6]. Available from: http://www.itis.gov
23. World Register of Marine Species Editorial Board. 2015. World Register of Marine Species. [cited 2015 July 6]. Available from: http://www.marinespecies.org

24. US National Oceanic and Atmospheric Administration Emergency Response Division. 2015. NOAA incident news. Office of Response and Restoration, National Ocean Service, Seattle WA, USA. [cited 2015 July 6]. Available from: http://incidentnews.noaa.gov/raw/index
25. US National Oceanic and Atmospheric Administration Emergency Response Division. 2015. River dilution calculator. Version 1.0.0 [computer software]. Office of Response and Restoration, Seattle, WA, USA.
26. Posthuma L, Suter GW, Traas TP. 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis, Boca Raton, FL, USA.
27. Bejarano AC, Farr JK. 2013. Development of short acute exposure hazard estimates: A tool for assessing the effects of chemical spills in aquatic environments. *Environ Toxicol Chem* 32:1918–1927.
28. US Environmental Protection Agency. 2015. Pesticides: Environmental effects. Technical overview of ecological risk assessment analysis phase: Ecological effects characterization. Washington, DC. [cited 2015 July 6]. Available from: http://www.epa.gov/oppefed1/ecorisk_ders/toera_analysis_eco.htm#Ecotox
29. Suter II GW. 2006. *Ecological Risk Assessment*. CRC, Boca Raton, FL, USA.
30. Solomon K, Giesy J, Jones P. 2000. Probabilistic risk assessment of agrochemicals in the environment. *Crop Prot* 19:649–655.
31. Bejarano AC, Mearns AJ. 2015. Improving environmental assessments by integrating species sensitivity distributions into environmental modeling: Examples with two hypothetical oil spills. *Mar Pollut Bull* 93:172–182.
32. US National Oceanic and Atmospheric Administration, Emergency Response Division. 2015. Computer-Aided Management of Emergency Operations (CAMEO) Chemical [computer software]. Office of Response and Restoration, Seattle, WA, USA.
33. Bejarano AC, Barron MG. 2014. Development and practical application of petroleum and dispersant interspecies correlation models for aquatic species. *Environ Sci Technol* 48:4564–4572.
34. Verhaar HJ, Van Leeuwen CJ, Hermens JL. 1992. Classifying environmental pollutants. *Chemosphere* 25:471–491.
35. Cronin MT, Dearden JC. 1995. QSAR in toxicology. 1. Prediction of aquatic toxicity. *Quant Struct Act Rel* 14:1–7.
36. 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.
37. Bejarano AC, Clark JR, Coelho GM. 2014. Issues and challenges with oil toxicity data and implications for their use in decision making: A quantitative review. *Environ Toxicol Chem* 33:732–742.
38. Gram C, Cockton G. 1996. *Design Principles for Interactive Software*. Chapman & Hall, London, UK.
39. D'Agostino RB. 1986. *Goodness-of-Fit Techniques*. CRC, Boca Raton, FL, USA.
40. Stephens MA. 1974. EDF statistics for goodness of fit and some comparisons. *J Am Stat Assoc* 69:730–737.
41. Aldenberg T, Slob W. 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotox Environmental Safe* 25:48–63.
42. Wheeler J, Grist E, Leung K, Morritt D, Crane M. 2002. Species sensitivity distributions: Data and model choice. *Mar Pollut Bull* 45:192–202.