Learned Discourses: Timely Scientific Opinions

Assessing the Reliability of Physico-Chemical Property Data (K_{ow}, S_w) for Hydrophobic Organic Compounds: DDT and DDE as a Case Study

Robert P. Eganhouse and James Pontolillo
U.S. Geological Survey, Reston, VA, USA
<eganhous@usgs.gov>

The fate of a chemical in the environment is primarily controlled by the physico-chemical properties of the substance, prevailing environmental conditions, and patterns of use (Mackay et al. 1997). Thus, accurate determination of physico-chemical properties is critical to the formulation of valid environmental models and assessments (Bennett et al. 2001). As litigation becomes an increasingly common aspect of environmental science, such studies may be challenged and even refuted. Recent legislation in the United States mandates that federal agencies must ensure the quality of data upon which regulatory action and policy decisions are founded (Reichhardt 2002). Some observers have suggested that because this legislation provides for unprecedented scrutiny of agency data quality, regulatory agencies might face a marked increase in the number of legal challenges.

Two of the most important physico-chemical properties relating to the environmental behavior of hydrophobic organic compounds are aqueous solubility (S_w) and the octanol-water partition coefficient (K_{ow}). Much of the interest in these properties comes from the fact that they can be used to estimate or predict other properties of more immediate environmental and ecotoxicological interest. For this reason, a primary focus of research in the last 20 years has been development of linear regression models in which S_w and K_{ow} (or more accurately, their logarithms) are correlated with parameters such as the organic carbon-normalized partition coefficient, bioconcentration and bioaccumulation factors, and indices of biodegradability or toxicity. These, in turn, can be incorporated into models that attempt to characterize the equilibrium distribution and transport rates of organic contaminants among environmental media or to predict impacts on biota. The predictions of such multimedia models have taken on greater immediacy with the recognition that certain persistent organic pollutants, such as DDT and DDE, are globally distributed and may be exerting adverse biological effects far from their original source areas (Wania and Mackay 1996).

Our interest in the K_{ow} and S_w values of DDT and DDE arose in the context of research we were conducting on the geochemical fate of these compounds in sediments of the Palos Verdes Shelf, California. We expected that studies reporting S_w and K_{ow} values for DDT and DDE would be well documented because DDT has a very long history, and its widespread application and biological effects are well known. Results of our initial literature survey, however, were unsatisfactory due to a large spread in data values. We decided to undertake an exhaustive review to identify all measured and estimated values of S_w and K_{ow} for these compounds. It soon became apparent that a number of serious problems severely limited use of the database and that our initial expectations had been naive. A complete description of the study discussed here can be found in a USGS report (Pontolillo and Eganhouse 2001) available in PDF format on the web (http://pubs.water.usgs.gov/wri01-4201/).

We examined more than 700 publications including databases, handbooks, review articles, and bibliographies for data, methodological details, and pertinent literature citations. The initial survey was augmented by computerized searches of the scientific literature using the Chemical Abstracts Search Service Index (CASSI) database and the World Wide Web using a variety of Internet search-engines. Several hundred potential articles were generated in this fashion. Each article was retrieved and scrutinized, and additional relevant leads were identified from the reference lists. The search progressed until no further leads could be identified (approximately 2 years).

Two problems are evident with the available database: egregious errors in reporting data and references, and poor data quality and/or inadequately documented procedures. The published literature is characterized by a preponderance of unnecessary data duplication. Numerous data and citation errors are also present. The percentage of original S_w and K_{ow} data for DDT and DDE in compilations has decreased with time, and in the most recent publications (1994-97) it composes only 6-26% of the reported data. Moreover, the largest recent compilation (Mackay et al. 1997) was found to contain only 68% and 30%, respectively, of the original S_w and K_{ow} data available at the time of its publication. Thus, widely used compilations are not only dominated by non-original data, they also are incomplete. The cumulative effect of these errors has been to obscure the extent and reliability of the original database such that deducing a reliable S_w or K_{ow} value would be difficult, if not impossible, for most users.

The verifiable, published original DDT/DDE S_w and K_{ow} data (ca. 1944-2001) consist of 62 S_w values (45 DDT, 17 DDE) and 100 log K_{ow} values (64 DDT, 36 DDE). The data values span 2-4 orders of magnitude, and there is little indication that the uncertainty in these properties has declined over the last 5 decades (Fig. 1). Consequently, the practice of using physico-chemical data without first carefully assessing its reliability is inappropriate. Unfortunately, most users do not have the time for such an undertaking and simply accept the recommendations offered in published tabulations. These recommendations are often based on examination of erroneous and incomplete data compilations such as those described previously. Without objective substantiating information, such recommendations would, thus, appear to be purely subjective. Even more troubling is the fact that such "recommended values" themselves can

Learned Discourses: Timely Scientific Opinions

fall within a range of more than two orders of magnitude (e.g. log K_{ow} DDT: 4.9-6.9; log K_{ow} DDE: 4.3-7.0). This makes the task of selecting a "recommended value" only marginally simpler (and no more reliable) than scrutinizing the database itself.

To assess the quality of original DDT and DDE S_w and K_{ow} data sources, we adapted criteria-based evaluation systems originally developed by Heller et al. (1994) and Kollig (1988). Examination of all original data sources revealed that 95-100% of the database literature is of poor or unevaluatable quality. The accuracy and reliability of the vast majority of the data are unknown due to inadequate documentation of the methods of determination. For example, estimates of precision have been reported for only 20% of experimental S_w data and only 10% of experimental K_{ow} data. Moreover, our research revealed that there has been

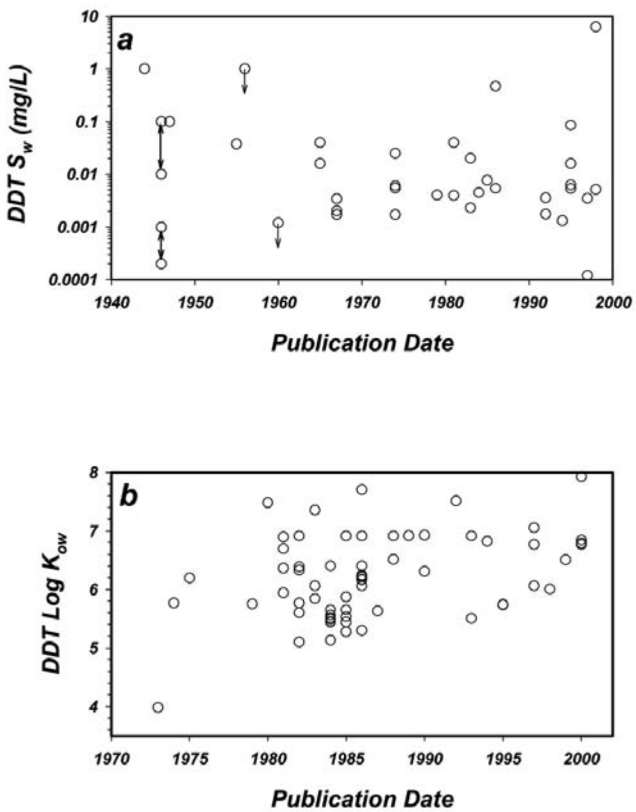

Figure 1. Original (a) S_w data (note log scale), and (b) log K_{ow} data for DDT (18-25°C) plotted as a function of publication date (modified from Pontolillo and Eganhouse, 2001). Double-headed arrows correspond to data published as a range of solubilities. Single-headed arrows indicate data published as maximum solubilities.

only one inter-laboratory comparison involving determination of the log K_{ow} values of DDT and DDE. Despite use of what is regarded as one of the more reliable direct experimental determination techniques (slow stir flask method) in this study, mean log K_{ow} values reported by the two laboratories for DDT differed by 0.6 units. We also examined some methodological trends and patterns in determination/estimation of S_w and K_{ow} using the DDT/DDE original database. The most striking observation was the fact that computational methods (for description see Pontolillo and Eganhouse 2001) have increasingly been substituted for experimental determination despite the fact that data used for model development and validation may be of unknown reliability.

Based on the prevalence of errors, the lack of methodological documentation, and unsatisfactory data quality, we have concluded that the reliability of the DDT/DDE S_w and K_{ow} database is questionable. The nature and extent of the errors are probably indicative of a more general problem in the literature of hydrophobic organic compounds. Under these circumstances, estimation of critical environmental parameters on the basis of S_w and K_{ow} is inadvisable because it will likely lead to incorrect environmental risk assessments. The current state of the database indicates that much greater efforts are needed to: 1) halt the proliferation of erroneous data and references, 2) initiate a coordinated program to develop improved methods of property determination and seek international standardization, 3) establish and maintain consistent reporting requirements for physicochemical property data, and 4) create a mechanism for archiving reliable data for widespread use in the scientific/regulatory community. Without corrective action, the trends we have observed can be expected to continue.

References

Bennett DH, Scheringer M, McKone TE, Hungerbühler K. 2001. Predicting long-range transport: A systematic evaluation of two multimedia transport models. *Environ Sci Technol* 35:1181-1189.

Heller SR, Bigwood DW, May WE. 1994. Expert systems for evaluating physico-chemical property values, 1. Aqueous solubility. *J Chem Inf Comput Sci* 34:627-636.

Kollig H. 1988. Criteria for evaluating the reliability of literature data on environmental process constants. *Toxicol Environ Chem* 17:287-311.

Mackay D, Shiu W-Y, Ma K-C. 1997. *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals, Volume 5-Pesticide Chemicals*. Lewis, Boca Raton, FL, USA.

Pontolillo J, Eganhouse RP. 2001. The search for reliable aqueous solubility (S_w) and octanol-water partition coefficient (K_{ow}) data for hydrophobic organic compounds: DDT and DDE as a case study. U.S. Geological Survey Water-Resources Investigations Report 01-4201, Reston, VA, USA.

Reichhardt T. 2002. Concern mounts as US agencies face challenges to data quality. *Nature* 416:249-250.

Wania F, Mackay D. 1996. Tracking the distribution of persistent organic pollutants. *Environ Sci Technol* 30: 390A-396A.