Supporting Information

Publication: Nowell, L.H., Crawford, C.G., Gilliom, R.J., Nakagaki, N., Stone, W.W., Thelin, G.P., and Wolock, D.W., 2009, Regression models for explaining and predicting concentrations of organochlorine pesticides in fish from United States streams. Environmental Toxicology and Chemistry, v. 28, n. 6, pp. 1346-1358.

S1: Table S1. For organochlorine pesticides in the present study, pesticide use information and occurrence data in whole fish from the U.S. Geological Survey's National Water Quality Assessment (NAWQA) Program, 1992–2001.

S2: Table S2. Potential explanatory variables considered in developing regression models for organochlorine pesticides in whole fish.

S3: Methods for determining key source variables: Agricultural use intensity and termite-urban score.

S4: References cited in online Supporting Information S1, S2, and S3.

Table S1. For organochlorine pesticides in the present study, pesticide use information and occurrence data in whole fish from the U.S. Geological Survey's National Water Quality Assessment (NAWQA) Program, 1992–2001.[a]

Pesticide(s) applied	Primary agricultural crops in 1966 or 1971 (% of total crop use) [b]	Estimated historical use, in million lbs/year (Data year)		National detection frequency in NAWQA whole-fish samples (%) [d]				NAWQA fish samples exceeding high and low wildlife benchmarks (%) [e]
		Agriculture [b]	Termite control [c]	Agricultural	Urban	Mixed	Undeveloped	
Dieldrin + Aldrin	Corn (92%), fruit and nuts, miscellaneous vegetables	15.485 (1966)	1–1.5 (1985)	63.3	56.3	43.3	4.4	3–5
Chlordane	Corn (56%), apples (25%), nursery, greenhouse crops	1.89 (1966, 1971) [f]	10 (1980), 3–3.5 (1985)	51.1	79.2	61.9	14.2	0–2
DDT + TDE [g]	Cotton (66%), tobacco, fruit and nuts	29.9 (1966)	—	90.0	90.3	90.5	55.1	15–76

[a] Cited references are listed in Supporting Information S4.
[b] Pesticide use data (1966) are from Eichers et al. [1]; 1971 chlordane use data are from Andrilenas et al. [2].
[c] Quantitative estimates for the top termiticides nationally are available for 1980 [3] and 1985 [4], [Esworthy RF. 1987. Incremental benefit analysis—Restricted use of all pesticides registered for subterranean termite control. U.S. Environmental Protection Agency, Washington, DC]. Other urban uses were common, but quantitative use estimates are not available.
[d] Detection frequencies are from Appendix 7D of Gilliom et al. ([5]; http://water.usgs.gov/nawqa/pnsp/pubs/circ1291/appendix7/7d.html).
[e] Low and high end of range represent the percent of samples exceeding high and low wildlife benchmarks, respectively. Benchmarks are from Appendix 3B of Gilliom et al. ([5]; http://water.usgs.gov/nawqa/pnsp/pubs/circ1291/appendix3).
[f] Chlordane use data were for 1971, except in three states (AR, LA, MS), for which 1966 data were used because 1971 data were not available (see Supporting Information S3).
[g] TDE, tetrachlorodiphenylethane.

Table S2. Potential explanatory variables considered in developing regression models for organochlorine pesticides in whole fish. Names of variables selected in one or more extrapolation models are in bold italics. Names of variables considered, but not selected, are in regular italics. Variables are listed by phase of model development.

Type of variable	Variables [a]	Units [b]	Sources [c,d]
Phase I. Source Variables			
Past agricultural pesticide use	Historical ***agricultural use intensity*** of DDT plus TDE (1966), aldrin plus dieldrin (1966), and chlordane (1966, 1971)	lb/mi^2 of basin land	Derived from county-level estimates of agricultural pesticide use developed by Gail P. Thelin (USGS, Sacramento, CA, July 13, 2004, personal communication) and 1970s GIRAS agricultural land ([6], as enhanced by Price et al. [7]). (See S3 for detail.)
Land use, 1970s	*Residential; commercial and industrial; transportation; total urban;* ***cropland and pasture****; orchards and vineyards; total agriculture; total undeveloped;* ***forested land***	Percent of basin land	1970s GIRAS land use and land cover data ([6], as enhanced by Price et al.[7])
Other urban use surrogates	Mean ***population density***, 1990	People per km^2	Derived by Price [8] from 1990 Census of Population and Housing data [9]
	Termite-urban score, 1970s	—	Developed by Nowell et al. [10] from 1970s GIRAS urban land ([6], enhanced by Price et al.[7]) and subterranean termite density zones [11]. (See S3 for detail.)
Fish sample characteristics	***Fish lipid content***	Percent	Measured
	Elapsed *time* from 1966 to the sampling date	Years	Observed
Sediment-based dummy variables [e]	*Nondetections*; *<10 µg/kg*, *10 to <100 µg/kg*, and *≥100 µg/kg* dry weight	[f]—	Measured
	Sediment organic carbon in the associated sediment sample	g/kg	Measured

Type of variable	Variables [a]	Units [b]	Sources [c,d]
Phase I. Source Variables, cont.			
Sediment-based dummy variables, cont. [e]	*Fish-sediment date* (fish sampling date minus sediment sampling date)	Days	Observed
Phase II. Taxon Dummy Variables			
Fish sample characteristics	*Brown trout,* ***common carp****, largemouth bass,* ***largescale sucker****, longear sunfish, rock bass,* ***sculpins****,* ***white sucker***	[f]—	Observed
Phase III. Watershed Variables			
Physical characteristics	***Drainage basin area***	km^2	Computed from basin delineations developed in the 1990s by USGS scientists using a number of data sources at various map scales (Naomi Nakagaki, USGS, Sacramento, CA, September 1, 2003, personal communication)
	Mean basin *slope*	Percent	Computed from USGS National Elevation Data [12]
	Mean basin *aspect*	Degrees	Computed from USGS National Elevation Data [12]
	Site elevation; mean *basin elevation*	m	Computed from USGS National Elevation Data [12]
	Road density	km per km^2	Derived from Census Bureau 2000 TIGER roads [13]
Soil properties	Mean *AWC*	Fraction	From Wolock [14] from data in the STATSGO database [15]
	Mean *sand, silt,* and *clay* composition	Percent	From Wolock [14] from data in the STATSGO database [15]

Type of variable	Variables [a]	Units [b]	Sources [c,d]
Phase III. Watershed Variables, cont.			
Soil properties, cont.	*Well-drained* soils (classified as soil hydrologic groups A or B); *poorly drained* soils (classified as soil hydrologic groups C, D, or C/D)	Percent of basin soils	From the STATSGO database [15], as enhanced by Barbara C. Ruddy and William A. Battaglin (USGS, Denver, CO, October 3, 1997, personal communication), in which missing STATSGO soil hydrologic group values were populated based on soil characteristics described by Foth and Schafer [16]
	Mean ***soil organic matter*** content	Percent by weight	From Wolock [14] from data in the STATSGO database [15]
	Mean *soil permeability*	in/h	From Wolock [14] from data in the STATSGO database [15]
	Mean ***R factor*** (rainfall erosivity) from the Universal Soil Loss Equation	100s of ft-tonf-in per h-acre-yr	From mean annual R factor for 1971–2000 [17]
	Mean ***K factor*** (soil erodability) for the uppermost soil horizon, from the Universal Soil Loss Equation	—	Developed by David M. Wolock (USGS, Lawrence, KS, May 9, 2003, personal communication) from the STATSGO database [15], compiled for the uppermost soil horizon
Hydrologic parameters	*Dunne overland flow* (streamflow derived from runoff resulting from precipitation on saturated soil); *Horton overland flow* (streamflow derived from runoff resulting from precipitation exceeding soil infiltration rate)	Percent	From Wolock [18, 19], generated from TOPMODEL [20]
	Subsurface flow ***contact*** time (time required for precipitation to travel through subsurface flow paths before entering stream)	Days	Developed by David M. Wolock (USGS, Lawrence, KS, March 29, 2001, personal communication) using the USGS Digital Elevation Model [21] and soil characteristics [14]

Type of variable	Variables [a]	Units [b]	Sources [c,d]
Phase III. Watershed Variables, cont.			
Hydrologic parameters, cont.	*Topographic wetness* index	Natural log (m)	Derived by David M. Wolock (USGS, Lawrence, KS, March 29, 2001, personal communication) using the USGS Digital Elevation Model [21] and the single flow direction algorithm described in Wolock and McCabe [22]
	Mean annual *runoff*, 1951–1980	in	Created from Gebert et al. [23]
Climate	Mean annual *precipitation*, 1961–1990	mm	Climate data estimated using PRISM [24]
	Mean annual *temperature*, 1961–1990	°C	Climate data estimated using PRISM [24]
	Mean annual *PET*, 1961–1990	mm	Estimated by David M. Wolock (USGS, Lawrence, KS, March 29, 2001, personal communication) using PRISM temperature data [24] with the Hamon PET equation [25]
Agricultural management practices	***Conservation tillage***; [g] ***irrigated land***; [g] *artificial drainage*	Percent of basin land or agricul-tural land	From Wieczorek ([26–31], using data from the 1992 Natural Resources Inventory [32] and the USGS's National Land Cover Data Set 1992 [33]
Phase IV. Regional Dummy Variables			
Geographical regions [h]	*Appalachians;* ***Corn Belt****;* ***Delta States****; Lake States; Mountain States; Northeast; Northern Plains; Pacific States; Southeast;* ***Southern Plains***	[f]—	USDA Farm Production Regions [2]

[a] AWC, available water capacity; PET, potential evapotranspiration; TDE, tetrachlorodiphenylethane; µg/kg, micrograms per kilogram.
[b] ft-tonf-in per h-acre-yr, feet-long-ton-inches per hour-acre-year; g/kg, grams per kilogram; h, hour; in, inches; km, kilometer; km^2, square kilometer; lb/mi^2, pounds per square mile; m, meters; mm, millimeters; °C, degrees Celsius, —, none.
[c] Cited references are listed in Supporting Information S4. GIRAS, Geographic Information Retrieval and Analysis System; PET, potential evapotranspiration; PRISM, Parameter-Elevation Regressions on Independent Slopes Model; STATSGO, State Soil Geographic database; TIGER, Topologically Integrated Geographic Encoding and Referencing; TOPMODEL,

Topography-Based Watershed Model; USDA, U.S. Department of Agriculture; USGS, U.S. Geological Survey.

[d] Geographic Information System data sources available online are as follows:
[6], http://edc.usgs.gov/products/landcover/lulc.html
[7], http://pubs.usgs.gov/ds/2006/240
[8], http://water.usgs.gov/lookup/getspatial?uspopd90x10g
[12], http://ned.usgs.gov
[13], http://www.geolytics.com
[14], http://water.usgs.gov/lookup/getspatial?muid
[17], http://www.ocs.orst.edu/pub/maps/Precipitation/rfactor/U.S./us_rfactor_meta.html
[18], http://water.usgs.gov/lookup/getspatial?satof48
[19], http://water.usgs.gov/lookup/getspatial?ieof48
[23], http://water.usgs.gov/lookup/getspatial?runoff
[26], http://water.usgs.gov/lookup/getspatial?ofr041189329
[27], http://water.usgs.gov/lookup/getspatial?ofr041189it01
[28], http://water.usgs.gov/lookup/getspatial?ofr041189it02
[29], http://water.usgs.gov/lookup/getspatial?ofr041189it03
[30], http://water.usgs.gov/lookup/getspatial?ofr041189606
[31], http://water.usgs.gov/lookup/getspatial?ofr041189607

[e] Dummy variables for pesticide concentrations measured in associated sediment sample (only in sediment-based models).

[f] Each variable is populated by 1 if it applies to sample, otherwise by 0.

[g] The variables selected in the models were in units of percent of basin land.

[h] The USDA Farm Production Region containing the largest portion of basin (Fig. 1, inset).

Supporting Information S3. Methods for determining key source variables: Agricultural use intensity and termite-urban score.[a]

Agricultural Use Intensity

Historical *agricultural use intensity* in the basin was estimated for the following pesticides: aldrin and dieldrin combined (for use in the dieldrin model); chlordane (for chlordane-group models); and DDT and tetrachlorodiphenylethane (TDE) combined (for DDT-group models). Historical agricultural pesticide use at the county scale was estimated from regional application rates developed using U.S. Department of Agriculture (USDA) data on pesticide use by farmers in 1966 and (for chlordane) 1971 and county-level harvested crop acreage from the 1964 and 1969 Census of Agriculture (Gail P. Thelin, U.S. Geological Survey, July 13, 2004, personal communication).

Specifically, regional pesticide-use-by-crop coefficients were first developed for each pesticide on individual crops in each of the 10 USDA farm production regions by combining national and regional estimates of pesticide use in 1966 [1] and/or (for chlordane) in 1971 [2], with county-level harvested crop acreage from the 1964 or 1969 (depending on the crop) Census of Agriculture [34]. For chlordane, 1971 pesticide use data were used to represent maximum agricultural use levels (because USDA agricultural use estimates were three-fold higher in 1971 than 1966) in all states with available 1971 data. For three states (Arkansas, Louisiana, and Mississippi), no USDA data on agricultural use in 1971 were available; therefore, for these three states only, 1966 data were used. The regional pesticide-use-by-crop coefficients were used to calculate use estimates for each pesticide in all counties in the conterminous U.S. that have agricultural land, as determined from the Census of Agriculture [34]. The resulting historical agricultural use estimates for the conterminous United States are shown in Figure 2: aldrin plus dieldrin (Fig. 2A), chlordane (Fig. 2B), and DDT plus TDE (Fig. 2C).

The *agricultural use intensity* in each basin was then calculated for each pesticide and for each sampling site by combining the basin boundaries for the site with the above-described county-scale estimates of agricultural pesticide use and mapped agricultural land for the 1970s. Basin boundaries for each site were developed in the 1990s by U.S. Geological Survey (USGS) scientists using a number of data sources at various map scales (Naomi Nakagaki, USGS, Sacramento, CA, September 1, 2003, personal communication). The 1970s agricultural land-use information was derived from the USGS Land Use and Land Cover (LULC) data (([6]; http://edc.usgs.gov/products/landcover/lulc.html), as enhanced by Price et al. ([7]; http://pubs.usgs.gov/ds/2006/240)). The LULC data are also known as Geographic Information Retrieval and Analysis System (GIRAS) data, which refers to the digital file storage format and software system used to develop the LULC data set [35]. Distinct changes in *agricultural use intensity* sometimes occur along state boundaries (see Figs. 2A–C) because regional application rates are based on USDA farm production regions (Fig. 1, inset), which follow state boundaries.

Termite-Urban Score:

The *termite-urban score* was developed by Nowell et al. ([10]; http://pubs.usgs.gov/sir/2006/5020) as a surrogate to represent past termiticide use. Aldrin, dieldrin, and chlordane were used in subterranean termite control through the late

1980s [3, 4, 36, 37]. National use estimates are available for the top termiticides in 1980 [3] and 1985 [4], [Esworthy RF. 1987. Incremental benefit analysis—Restricted use of all pesticides registered for subterranean termite control. U.S. Environmental Protection Agency, Washington, DC] (see Supporting Information S1). However, quantitative termiticide use data are not available at the county or regional scales. In the present study, the *termite-urban score* was determined on the basis of the percentage of urban land use in the basin (1970s GIRAS data [6], as enhanced by Price et al. [7]) and the relative hazard of subterranean termite infestations [11]. Combined, these factors are expected to reflect the extent of past termiticide use, because termiticides were most likely applied to buildings in parts of the country where termites occurred. The national distributions of these two factors are shown in Figure 2D.

Specifically, the *termite-urban score* was calculated by multiplying the 1970s urban land (as a percentage of basin area) located within each of four zones of subterranean termite density by a weighting factor for that zone (1 for the none-to-slight zone, 2 for slight-to-moderate, 3 for moderate-to-heavy, and 4 for very heavy); and then summing the four weighted percentages. The *termite-urban score* in the present study corresponds to the weighted termite-urban score (1970s) in Nowell et al. [10].

[a] Variable names are in italics. Cited references are listed in Supporting Information S4.

Supporting Information S4. References cited in online Supporting Information S1, S2, and S3.

1. Eichers T, Andrilenas P, Blake H, Jenkins R, Fox A. 1970. Quantities of pesticides used by farmers in 1966. Agricultural Economic Report no. 179. U.S. Department of Agriculture, Economic Research Service, Washington, DC.

2. Andrilenas PA. 1974. Farmers' use of pesticides in 1971—Quantities. Agricultural Economic Report no. 252. U.S. Department of Agriculture, Economic Research Service, Washington, DC.

3. U.S. Environmental Protection Agency. 1983. Analysis of the risks and benefits of seven chemicals used for subterranean termite control. EPA-540/9-83-005. Washington, DC.

4. Nowell LH, Capel PD, Dileanis PD. 1999. *Pesticides in Stream Sediment and Aquatic Biota—Distribution, Trends, and Governing Factors*. Pesticides in the Hydrologic System series, Vol 4. CRC, Boca Raton, FL, USA.

5. Gilliom RJ, Barbash JE, Crawford CG, Hamilton PA, Martin JD, Nakagaki N, Nowell LH, Scott JC, Stackelberg PE, Thelin GP, Wolock DM. 2006. The quality of our nation's waters—Pesticides in the nation's streams and ground water, 1992–2001. Circular 1291. U.S. Geological Survey, Reston, VA.

6. U.S. Geological Survey. 1998. Land use and land cover digital data from 1:250,000- and 1:100,000-scale maps [digital files in GIRAS format]. Reston, VA.

7. Price CV, Nakagaki N, Hitt KJ, Clawges RM. 2007. Enhanced historical land-use and land-cover data sets of the U.S. Geological Survey. Data Series 240. U.S. Geological Survey, Reston, VA.

8. Price CV. 2003. 1990. Population density by block group for the conterminous United States, digital map. U.S. Geological Survey, Reston, VA.

9. U.S. Bureau of the Census. 1991. Census of population and housing, 1990. Public Law (P.L.) 94-171. Washington, DC.

10. Nowell LH, Crawford CG, Nakagaki N, Thelin GP, Wolock DW. 2006. Regression model for explaining and predicting concentrations of dieldrin in whole fish from United States streams. Scientific Investigations Report 2006-5020. U.S. Geological Survey, Reston, VA.

11. Beal RH, Mauldin JK, Jones SC. 1994. Subterranean termites: their prevention and control in buildings. *Home and Garden Bulletin* 64. U.S. Department of Agriculture, Washington, DC.

12. U.S. Geological Survey. 2003. National elevation data set (NED), digital data. Reston, VA.

13. Geolytics, Inc. 2004. Census CD 2000 and Street CD 2000. Geolytics, Inc., East Brunswick, NJ, USA.

14. Wolock DM. 1997. STATSGO soil characteristics for the conterminous United States. Digital map. Open-File Report 97-656. U.S. Geological Survey, Lawrence, KS.

15. Natural Resources Conservation Service. 1994. State soil geographic (STATSGO) database—data use information. Miscellaneous Publication no. 1492. U.S. Department of Agriculture, National Soil Survey Center, Lincoln, NE.

16. Foth HD, Schafer JW. 1980. *Soil Geography and Land Use.* John Wiley and Sons, Inc., New York, NY, USA.

17. Daly C, Taylor, GH. 2002. United States Mean Annual R-factor, 1971–2000, digital map. Oregon State University, Corvallis, OR, USA.

18. Wolock DM. 2003a. Saturation overland flow estimated by TOPMODEL for the conterminous United States. Open-File Report 03-264. U.S. Geological Survey, Reston, VA.

19. Wolock DM. 2003b. Infiltration-excess overland flow estimated by TOPMODEL for the conterminous United States. Open-File Report 03-310. U.S. Geological Survey, Reston, VA.

20. Wolock DM. 1993. Simulating the variable-source-area concept of streamflow generation with the watershed model TOPMODEL. Water Resources-Investigations Report 93-4124. U.S. Geological Survey, Reston, VA.

21. U.S. Geological Survey. 1993. Digital elevation models. *Data users guide 5*. Reston, VA.

22. Wolock DM, McCabe GJ Jr. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resour Res* 31(5): 1315–1324.

23. Gebert WA, Graczyk DJ, Krug WR. 1987. Average annual runoff in the United States, 1951–80. Hydrologic Investigations Atlas HA-710. U.S. Geological Survey, Madison, WI.

24. Daly C. 2006. Guidelines for assessing the suitability of spatial climate data sets. *Int J Climatol* 26:707–721.

25. Hamon WR. 1961. Estimating potential evapotranspiration. Proceedings of the American Society of Civil Engineers, *J Hydraulics Division* 87:107–120.

26. Wieczorek ME. 2004a. Conservation tillage systems on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

27. Wieczorek ME. 2004b. Gravity irrigation source on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

28. Wieczorek ME. 2004c. Pressure irrigation source on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

29. Wieczorek ME. 2004d. Gravity and pressure irrigation source on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

30. Wieczorek ME. 2004e. Subsurface drains on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

31. Wieczorek ME. 2004f. Surface drainage, field ditches on agricultural land in the conterminous United States. Ver. 1.1, U.S. Geological Survey digital map data. Open-File Report 2004-1189. U.S. Geological Survey, Baltimore, MD.

32. Natural Resources Conservation Service. 1995. 1992 National Resources Inventory. Soil Conservation Service National Cartography and GIS Center, Fort Worth, TX, USA.

33. Vogelmann JE, Howard SM, Yang L, Larson CR, Wylie BK, Van Driel N. 2001. Completion of the 1990's national land cover dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogramm Eng Remote Sensing* 67:650–662.

34. Economic Research Service. 1999. Historical census of agriculture data, 1949–1974, records of the Economic Research Service (record Group 354). Data compiled from Section 1, chap. 3–4 of the Census of Agriculture, Vol 1. National Archives, Washington, DC, USA.

35. Mitchell WW, Guptill SC, Anderson KE, Fegeas RG, Hallam CA. 1977. GIRAS—a geographic information system analysis system for handling land use and land cover data. Professional Paper 1059. U.S. Geological Survey, Reston, VA.

36. U.S. Environmental Protection Agency. 1992. National study of chemical residues in fish, volume II. EPA-823-R-92-008b. Office of Science and Technology, Washington, DC.

37. U.S. Environmental Protection Agency. 2000. Great Lakes binational toxics strategy: The Level I pesticides in the binational strategy. Section 3.0, Review of domestic use status by pesticide. Final Report. Great Lakes National Program Office, Chicago, IL.